Nonlinear Processes
in Geophysics

Prognostic assumed-probability-density-function (distribution density function) approach: further generalization and demonstrations

Jun-Ichi Yano

CNRM UMR3589 (CNRS), Météo-France, Toulouse, France

Correspondence: Jun-Ichi Yano (jun-ichi.yano@cnrs.fr)

Abstract. A methodology for directly predicting the time evolution of the assumed parameters for distribution densities based on the Liouville equation, as proposed earlier, is extended to multidimensional cases and to cases in which the systems are constrained by integrals over a part of the variable range. The general formulation developed here is applicable to a wide range of problems, including the frequency distributions of subgrid-scale variables, hydrometeor size distributions, and probability distributions characterizing data uncertainties. The extended methodology is tested against a convective energy-cycle system and the Lorenz strange attractor. As a general tendency, the variance tends to collapse to a vanishing value over a finite time, regardless of the chosen assumed distribution form. This general tendency is likely due to a common cause, as the collapse of the variance is commonly found in ensemble-based data assimilation due to low dimensionality.

1 Introduction

A noble manner to characterize nonlinear systems is by predicting the evolution of the distributions of variables in space as well as the frequency distributions of variables at a single macroscopic point and as probabilities. Important applications in geophysics include subgrid-scale parameterizations based on the distribution density functions (DDFs: Sommeria and Deadorff, 1977; Mellor, 1977; Bougeault, 1981; Le-Treut and Li, 1991; Bechtold et al., 1992, 1995; Richard and Royer, 1993; Bony and Emanuel, 2001; Golaz et al., 2002; Tompkins, 2002), the particle-size distributions in cloud microphysics (cf. Khain et al., 2015; Khain and Pinsky, 2018), and the characterizations of data and model uncertainties by the probability density functions (PDFs) in data assimilations (Carrassi et al., 2018; Evensen et al., 2022). Readers are referred to Yano et al. (2018) for backgrounds of the present study in the numerical weather forecast. Yano et al. (2024), hereafter referred to as YLP, proposed a general methodology for evaluating the evolution of these distributions in an efficient manner. The essence of their approach may be called

"prognostic assumed-PDF (DDF)", as an extension of the classical assumed-PDF approaches (cf. Golaz et al., 2002).

The present study constitutes a sequel to the YLP study. Here, the uniqueness of this work lies in carrying out a very solid analysis of the performance by directly comparing the assumed-PDF results with direct numerical results from the Liouville equation using some simple dynamical systems. A basic, but general, robust formulation provided by YLP enables this kind of comparison: this is not possible with the current existing assumed-PDF schemes, as reviewed in the introduction of YLP, because these are formulated only in a case-by-case manner with specific applications in mind. Thus, the performance of those schemes cannot be tested by simple dynamical systems. The robustness of the proposed formulation is discussed extensively in YLP.

The goal of the present paper is to test this formulation using more advanced cases. In YLP, only one-variable cases under relatively simple constraints (i.e., output conditions) were considered. The purpose of the present work is to (1) further generalize this methodology to multidimensional systems with more general constraints and (2) present further demonstrative cases. The present study is considered one of

the first steps in making this new novel approach for computing distributions prognostically into operations. Some difficulties encountered are also discussed with respect to this ultimate goal.

The presentation of this work will be developed in the following manner: Sect. 2 briefly summarizes the formulation presented in YLP, whose details and extensive references are to be consulted in the original paper, and generalizes it to cases in which (1) constraints are defined over limited integral ranges and (2) different assumed distribution forms are assumed over the different domains; Sect. 2 generalizes one-variable results for the assumed-PDF formulation into the multivariable systems in a general manner, without specifying a PDF form; and Sect. 3 shows (in a more concrete manner) how these general formulations can be used when a Gaussian distribution with two variables is assumed. This example also suggests how reductions with different assumed-PDF forms can be proceeded. The general formulations presented in Sects. 2 and 3 are applied to the two dynamical systems in Sects. 4 and 5: (i) the two-variable convective energy-cycle system introduced by Yano and Plant (2012) (Sect. 4) and (ii) the three-variable system of the Lorenz strange attractor (Sect. 5). Three different assumed-PDF forms are considered for each system, and the results are discussed. Section 6 summarizes the results and presents further discussions.

2 General formulation

2.1 Basic formulation

The basic formulation presented in YLP is summarized first; however, the reader is referred to the original paper for details. Let us assume a dynamical system with a single variable, ϕ, and that a distribution of ϕ can be approximately represented by an assumed form

$$p = p(\phi; \lambda_0, \lambda_1, \ldots, \lambda_N), \tag{1}$$

which is characterized by N parameters, λ_j $(j = 1, \ldots, N)$. We also separately introduce a normalization factor, λ_0, that satisfies a relation of $p \propto \lambda_0$. It follows that

$$\frac{\partial p}{\partial \lambda_0} = \frac{p}{\lambda_0}. \tag{2}$$

The distribution Eq. (1) is normalized by

$$\int p \, \mathrm{d}\phi = 1. \tag{3}$$

Here and in the following, an unspecified integral range may be taken from $-\infty$ to $+\infty$ for many of the physical variables, but some physical variables are semi-positive definite (e.g., temperature and mixing ratios). In the latter case, the integral range above must be from 0 to $+\infty$.

The evolution of this distribution, p, is governed by the Liouville equation,

$$\frac{\partial p}{\partial t} = -\frac{\partial p S}{\partial \phi}, \tag{4}$$

when the dynamical system is defined by $\dot{\phi} = S$. From the time derivative of Eq. (3),

$$\frac{\dot{\lambda}_0}{\lambda_0} = -\sum_{i=1}^{N} \left[\int \frac{\partial p}{\partial \lambda_i} \mathrm{d}\phi \right] \dot{\lambda}_i. \tag{5}$$

By inserting Eq. (1) into Eq. (4), weighting it by σ_l $(l = 1, \ldots, N)$, and integrating it over the full variable range, we obtain a final expression for the prognostic equation for the distribution parameters, $\{\lambda_j\}$:

$$\sum_{j=1}^{N} \dot{\lambda}_j \left[\int \sigma_l \frac{\partial p}{\partial \lambda_j} \mathrm{d}\phi - \int \sigma_l p \, \mathrm{d}\phi \int \frac{\partial p}{\partial \lambda_j} \mathrm{d}\phi \right] = \int p S \frac{\partial \sigma_l}{\partial \phi} \mathrm{d}\phi \tag{6}$$

for $l = 1, \ldots, N$. We can see from Eq. (6) that the weights, σ_l, are most conveniently chosen in such a manner that

$$\langle \sigma_l \rangle = \int p \sigma_l \mathrm{d}\phi \quad \text{and} \quad l = 1, \ldots, N \tag{7}$$

constitute the constraints for this distribution: YLP suggest choosing these constraints (Eq. 7) to be the outputs that are required in a host model, and they call it the "output-controlled distribution principle". As the left-hand side is equal to $\mathrm{d}\langle \sigma_l \rangle / \mathrm{d}t$, Eq. (6) can predict these constraints in a self-consistent manner under an assumed form of Eq. (1). The core part of the derivation of this formulation and further discussion on the meaning of the weight σ_l are presented fully in Sect. 5.1 of YLP.

In the following two subsections, this basic formulation is generalized into two cases: first, constraints are defined over limited integral ranges of the distribution variable, ϕ; second, as a consequence, distributions take different forms over those limited ranges. This generalization is important in many operational applications; for example, in cloud schemes, various conditional statistics, including the cloud fraction, must be evaluated by restricting the integrals above saturation in the total-water distribution.

2.2 When the constraints are defined over limited integral ranges

Here, we generalize the above basic formulation to cases in which the different constraints are introduced over two different ranges of the distribution variables. More specifically, we assume that the distribution variable, ϕ, is defined over $[-\infty, +\infty]$, and the constraints are introduced in analogy with Eq. (6), although differently over the two ranges, $[-\infty, 0]$ and $[0, +\infty]$:

$$\langle \sigma_l \rangle_+ \equiv \int_0^{+\infty} F_l^+(p, \phi) \mathrm{d}\phi = C_l^+, \quad l = 1, \ldots, N^+, \tag{8a}$$

$$\langle \sigma_l \rangle_- \equiv \int_{-\infty}^{0} F_l^-(p, \phi) \mathrm{d}\phi = C_l^-, \quad l = 1, \ldots, N^-, \tag{8b}$$

as a generalization of Eq. (7). In the above expressions, $\sigma_l^{\pm} = F_l^{\pm}/p_{\pm}$ by following the output-constrained distribution principle proposed in YLP. Here, in the following abstract examples, there is no host model looking for outputs from a system defining a distribution. Thus, we will simply set those constraints to be the means and variances, as well as covariances, of the system, with priority given to defining the former under a given number of PDF parameters.

In this case, the variational principle to maximize the information entropy with the given constraints (cf. Sect. 3.1.1 of YLP) is given by the following:

$$\delta \left[-\int_{-\infty}^{+\infty} p \log p \, \mathrm{d}\phi - \sum_{l=0}^{N^+} \lambda_l^{+\prime} \int_{0}^{+\infty} F_l^+(p, \phi) \mathrm{d}\phi \right.$$
$$\left. -\sum_{l=0}^{N^-} \lambda_l^{-\prime} \int_{-\infty}^{0} F_l^-(p, \phi) \mathrm{d}\phi \right] = 0. \tag{9}$$

It then reduces to the following:

$$\int_{0}^{+\infty} \left(\log p + \sum_{l=0}^{N^+} \lambda_l^+ \frac{\partial F_l^+}{\partial p} \right) \delta p \, \mathrm{d}\phi$$
$$+ \int_{-\infty}^{0} \left(\log p + \sum_{l=0}^{N^-} \lambda_l^- \frac{\partial F_l^-}{\partial p} \right) \delta p \, \mathrm{d}\phi = 0.$$

Thus, the most likely distribution under the constraints (Eq. 8a, b) is

$$p = \begin{cases} \lambda_0^+ \exp\left[-\sum_{l=1}^{N^+} \lambda_l^+ \frac{\partial F_l^+}{\partial p} \right], & \phi > 0 \\ \lambda_0^- \exp\left[-\sum_{l=1}^{N^-} \lambda_l^- \frac{\partial F_l^-}{\partial p} \right], & \phi < 0. \end{cases} \tag{10}$$

Note that $\lambda_0^+ = \lambda_0^-$, due to the continuity of p, when $\partial F_l^-/\partial p = 0$ at $\phi = 0$.

2.3 When the distribution takes different forms in two different domains

In the last subsection, the constraints have been generalized to a case in which they are defined over two limited ranges of the distribution variable, ϕ; see Eq. (8a) and (8b). As it turns out, the most likely distribution (Eq. 10) under these constraints also takes different forms over these two ranges. Consequently, the formulation for predicting the given PDF parameters must also be generalized to such cases: this subsection addresses this issue.

By following the last subsection, we divide the variable range into the two subdomains $[-\infty, 0]$ and $[0, +\infty]$ and assume that the distribution takes different forms over those two subdomains. Thus,

$$p = \begin{cases} p^+(\phi, \lambda_0^+, \lambda_1^+, \ldots, \lambda_N^+), & \phi > 0 \\ p^-(\phi, \lambda_0^-, \lambda_1^-, \ldots, \lambda_N^-), & \phi < 0, \end{cases} \tag{11}$$

assuming $N^+ = N^- = N$ for now. By continuity, at $\phi = 0$, $p^+ = p^-$. Again, we assume that the first parameters, $\lambda_0^{\pm}$, are the normalization factors; thus,

$$\frac{\partial p}{\partial \lambda_0^+} = \frac{p}{\lambda_0^+}, \phi > 0, \tag{12a}$$

$$\frac{\partial p}{\partial \lambda_0^-} = \frac{p}{\lambda_0^-}, \phi < 0. \tag{12b}$$

Especially when $p|_{\phi=\pm 0} = \lambda_0^{\pm}$, as is the case with the result from Eq. (10), it follows that $\lambda_0^+ = \lambda_0^- (= \lambda_0)$. Here, the assumed form of Eq. (11) follows from the constraints of the form in Eq. (8a) and (8b).

By applying the time derivative to the normalization condition,

$$\int_{-\infty}^{+\infty} p \, \mathrm{d}\phi = 1,$$

we find

$$\frac{\partial}{\partial t} \int_{0}^{+\infty} p \, \mathrm{d}\phi + \frac{\partial}{\partial t} \int_{-\infty}^{0} p \, \mathrm{d}\phi = 0, \tag{13}$$

which reduces to

$$\frac{\dot{\lambda}_0}{\lambda_0} = -\sum_{i=1}^{N} \left[\dot{\lambda}_i^+ \int_{0}^{+\infty} \frac{\partial p}{\partial \lambda_i^+} \mathrm{d}\phi + \dot{\lambda}_i^- \int_{-\infty}^{0} \frac{\partial p}{\partial \lambda_i^-} \mathrm{d}\phi \right]. \tag{14}$$

Here, $\lambda_0^+ = \lambda_0^- \equiv \lambda_0$ and $p_+ + p_- = 1$, where

$$p_+ = \int_{0}^{+\infty} p \, \mathrm{d}\phi, \quad p_- = \int_{-\infty}^{0} p \, \mathrm{d}\phi.$$

After substituting the assumed-PDF form in Eq. (11) and applying the chain rules on the time derivative in the Liouville equation (i.e., Eq. 4), it reduces to the following:

$$p \frac{\dot{\lambda}_0}{\lambda_0} + \sum_{i=1}^{N} \frac{\partial p}{\partial \lambda_i^+} \dot{\lambda}_i^+ + \frac{\partial}{\partial \phi}(pS) = 0, \phi > 0, \tag{15a}$$

$$p \frac{\dot{\lambda}_0}{\lambda_0} + \sum_{i=1}^{N} \frac{\partial p}{\partial \lambda_i^-} \dot{\lambda}_i^- + \frac{\partial}{\partial \phi}(pS) = 0, \phi < 0. \tag{15b}$$

By applying weighted integrals on both and then further substituting Eq. (14) into the above, we obtain a pair of equations for predicting the PDF parameters:

$$
\sum_{i=1}^{N} \left\{ \dot{\lambda}_i^+ \left[\int_0^{+\infty} \sigma_l^+ \frac{\partial p}{\partial \lambda_i^+} d\phi - \int_0^{+\infty} \sigma_l^+ p d\phi \int_0^{+\infty} \frac{\partial p}{\partial \lambda_i^+} d\phi \right] \right.
$$
$$
\left. - \dot{\lambda}_i^- \int_0^{+\infty} \sigma_l^+ p d\phi \int_{-\infty}^0 \frac{\partial p}{\partial \lambda_i^-} d\phi \right\}
$$
$$
+ \int_0^{+\infty} \sigma_l^+ \frac{\partial}{\partial \phi}(pS) d\phi = 0, \tag{16a}
$$

$$
\sum_{i=1}^{N} \left\{ -\dot{\lambda}_i^+ \int_{-\infty}^0 \sigma_l^- p d\phi \int_0^{+\infty} \frac{\partial p}{\partial \lambda_i^+} d\phi \right.
$$
$$
\left. + \dot{\lambda}_i^- \left[\int_{-\infty}^0 \sigma_l^- \frac{\partial p}{\partial \lambda_i^-} d\phi - \int_{-\infty}^0 \sigma_l^- p d\phi \int_{-\infty}^0 \frac{\partial p}{\partial \lambda_i^-} d\phi \right] \right\}
$$
$$
+ \int_{-\infty}^0 \sigma_l^- \frac{\partial}{\partial \phi}(pS) d\phi = 0, \tag{16b}
$$

for $l = 1, \ldots, N$.

2.4 Multidimensional case (I)

The formulations introduced in the last subsections are now further generalized into a multidimensional case, in which the distribution depends on more than one variable. This variable set is treated as a vector, $\boldsymbol{x}$, with the first component corresponding to x. Here, note a change in the notation for the distribution variable from ϕ to $\boldsymbol{x}$ in the following.

As in the last two subsections, we assume that a given system is constrained differently depending on the sign of x. More precisely, we assume that there are N constraints that are defined differently for the different sign of x as well as M additional constraints that are defined to be independent of the sign of x. Consequently, the weights to be introduced are as follows: $\sigma_l^\pm = F_l^\pm/p_\pm$, for $l = 1, \ldots, N$, and $\sigma_l = F_l/p$, for $l = N+1, \ldots, N+M$, following the notation in Eq. (8a) and (8b). It also follows that the corresponding distribution is defined differently depending on the sign of x in the following manner:

$$
p = \begin{cases} p^+(\boldsymbol{x}, \lambda_0^+, \lambda_1^+, \ldots, \lambda_N^+, \lambda_{N+1}, \ldots, \lambda_{N+M}), & x > 0 \\ p^-(\boldsymbol{x}, \lambda_0^-, \lambda_1^-, \ldots, \lambda_N^-, \lambda_{N+1}, \ldots, \lambda_{N+M}), & x < 0. \end{cases}
$$

Here, the first N parameters, $\{\lambda_l^\pm\}$ ($l = 1, \ldots, N$), take different definitions for the positive and negative sides of x, whereas the last M parameters, $\{\lambda_{N+l}\}$ ($l = 1, \ldots, M$), are assumed to be common.

However, for simplicity, we drop the superscript "$\pm$" on p for now. Thus, repeating the same procedure as in the last

subsection, we obtain a pair of prognostic equations for the PDF parameters:

$$
\sum_{i=1}^{N} \left\{ \dot{\lambda}_i^+ \left[\int_+ \sigma_l \frac{\partial p}{\partial \lambda_i^+} d\boldsymbol{x} - \int_+ \sigma_l p d\boldsymbol{x} \int_+ \frac{\partial p}{\partial \lambda_i^+} d\boldsymbol{x} \right] \right.
$$
$$
\left. - \dot{\lambda}_i^- \int_+ \sigma_l p d\boldsymbol{x} \int_- \frac{\partial p}{\partial \lambda_i^-} d\boldsymbol{x} \right\}
$$
$$
+ \sum_{i=1}^{M} \dot{\lambda}_{N+i} \left[\int_+ \sigma_l \frac{\partial p}{\partial \lambda_{N+i}} d\boldsymbol{x} - \int_+ \sigma_l p d\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_{N+i}} d\boldsymbol{x} \right]
$$
$$
+ \int_+ \sigma_l \frac{\partial}{\partial x}(pS) d\boldsymbol{x} = 0, \tag{17a}
$$

$$
\sum_{i=1}^{N} \left\{ -\dot{\lambda}_i^+ \int_- \sigma_l p d\boldsymbol{x} \int_+ \frac{\partial p}{\partial \lambda_i^+} d\boldsymbol{x} + \dot{\lambda}_i^- \left[\int_- \sigma_l \frac{\partial p}{\partial \lambda_i^-} d\boldsymbol{x} \right. \right.
$$
$$
\left. \left. - \int_- \sigma_l p d\boldsymbol{x} \int_- \frac{\partial p}{\partial \lambda_i^-} d\boldsymbol{x} \right] \right\}
$$
$$
+ \sum_{i=1}^{M} \dot{\lambda}_{N+i} \left[\int_- \sigma_l \frac{\partial p}{\partial \lambda_{N+i}} d\boldsymbol{x} \right.
$$
$$
\left. - \int_- \sigma_l p d\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_{N+i}} d\boldsymbol{x} \right] + \int_- \sigma_l \frac{\partial}{\partial x}(pS) d\boldsymbol{x} = 0, \tag{17b}
$$

where the $\int_\pm$ terms suggest integrals over $x > 0$ and $x < 0$, respectively, depending on the sign.

Taking the sum of the two, we obtain the following:

$$
\sum_{i=1}^{N} \left\{ \dot{\lambda}_i^+ \left[\int_+ \sigma_l \frac{\partial p}{\partial \lambda_i^+} d\boldsymbol{x} - \int \sigma_l p d\boldsymbol{x} \int_+ \frac{\partial p}{\partial \lambda_i^+} d\boldsymbol{x} \right] \right.
$$
$$
\left. + \dot{\lambda}_i^- \left[\int_- \sigma_l \frac{\partial p}{\partial \lambda_i^-} d\boldsymbol{x} - \int \sigma_l p d\boldsymbol{x} \int_- \frac{\partial p}{\partial \lambda_i^-} d\boldsymbol{x} \right] \right\}
$$
$$
+ \sum_{i=1}^{M} \dot{\lambda}_{N+i} \left[\int \sigma_l \frac{\partial p}{\partial \lambda_{N+i}} d\boldsymbol{x} - \int \sigma_l p d\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_{N+i}} d\boldsymbol{x} \right]
$$
$$
+ \int \sigma_l \frac{\partial}{\partial \phi}(pS) d\boldsymbol{x} = 0. \tag{18a}
$$

Especially when σ_l does not depend on the sign of x (i.e., for $l = N+1, \ldots, N+M$) and the distribution is separable with these two distributions variables, the first N-sum disappears in Eq. (18a) and

$$
\sum_{i=1}^{M} \dot{\lambda}_{N+i} \left[\int \sigma_l \frac{\partial p}{\partial \lambda_{N+i}} d\boldsymbol{x} - \int \sigma_l p d\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_{N+i}} d\boldsymbol{x} \right]
$$
$$
+ \int \sigma_l \frac{\partial}{\partial \phi}(pS) d\boldsymbol{x} = 0. \tag{18b}
$$

Thus, the last M parameters, λ_{N+l} ($l = 1, \ldots, M$), can be predicted by a single set of equations (i.e., Eq. 18b). Note also

that the prognostic equations for the PDF parameters reduce to Eq. (18b) when all of the constraints are defined over the full range so that $N = 0$.

2.5 Multidimensional case (II)

In this subsection, we further divide the y direction into two subdomains depending on the sign of y. In this case, the weights to be introduced are as follows: $\sigma_l^{\pm\pm} = F_l^{\pm\pm}/p_{\pm\pm}$, for $l = 1, \ldots, N$, and $\sigma_l = F_l/p$, for $l = N+1, \ldots, N+M$, following similar notation to that in Eq. (8a) and (8b). Here, the double superscripts "$\pm\pm$" are introduced to indicate the sides in both the x and y directions. Otherwise, the conventions for the definitions of the $N+M$ PDF parameters remain the same. Thus, the first N weights take four different definitions depending on the signs of x and y. It also follows that we further subdivide the domain in the y direction; thus,

$$
p = \begin{cases}
p^{++}(\boldsymbol{x}, \lambda_0^{++}, \lambda_1^{++}, \ldots, \lambda_N^{++}, \lambda_{N+1}, \ldots, \lambda_{N+M}), & x > 0, y > 0 \\
p^{+-}(\boldsymbol{x}, \lambda_0^{+-}, \lambda_1^{+-}, \ldots, \lambda_N^{+-}, \lambda_{N+1}, \ldots, \lambda_{N+M}), & x > 0, y < 0 \\
p^{-+}(\boldsymbol{x}, \lambda_0^{-+}, \lambda_1^{-+}, \ldots, \lambda_N^{-+}, \lambda_{N+1}, \ldots, \lambda_{N+M}), & x < 0, y > 0 \\
p^{--}(\boldsymbol{x}, \lambda_0^{--}, \lambda_1^{--}, \ldots, \lambda_N^{--}, \lambda_{N+1}, \ldots, \lambda_{N+M}), & x < 0, y < 0.
\end{cases}
$$

These PDF parameters can be predicted by the following equations:

$$
\sum_{i=1}^{N} \left\{ \dot\lambda_i^{++} \left[\int_{++} \sigma \frac{\partial p}{\partial \lambda_i^{++}} \mathrm{d}\boldsymbol{x} - \int_{++} \sigma p \mathrm{d}\boldsymbol{x} \int_{++} \frac{\partial p}{\partial \lambda_i^{++}} \mathrm{d}\boldsymbol{x} \right] \right.
$$
$$
- \dot\lambda_i^{+-} \int_{++} \sigma p \mathrm{d}\boldsymbol{x} \int_{+-} \frac{\partial p}{\partial \lambda_i^{+-}} \mathrm{d}\boldsymbol{x} - \dot\lambda_i^{-+} \int_{++} \sigma p \mathrm{d}\boldsymbol{x} \int_{-+} \frac{\partial p}{\partial \lambda_i^{-+}} \mathrm{d}\boldsymbol{x}
$$
$$
\left. - \dot\lambda_i^{--} \int_{++} \sigma p \mathrm{d}\boldsymbol{x} \int_{--} \frac{\partial p}{\partial \lambda_i^{--}} \mathrm{d}\boldsymbol{x} \right\} + \sum_{i=1}^{M} \dot\lambda_{N+i} \left[\int_{++} \sigma \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right.
$$
$$
\left. - \int_{++} \sigma p \mathrm{d}\boldsymbol{x} \int_{++} \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right] + \int_{++} \sigma \frac{\partial}{\partial \phi}(pS) \mathrm{d}\boldsymbol{x} = 0, \tag{19a}
$$

$$
\sum_{i=1}^{N} \left\{ -\dot\lambda_i^{++} \int_{+-} \sigma p \mathrm{d}\boldsymbol{x} \int_{++} \frac{\partial p}{\partial \lambda_i^{++}} \mathrm{d}\boldsymbol{x} + \dot\lambda_i^{+-} \left[\int_{+-} \sigma \frac{\partial p}{\partial \lambda_i^{+-}} \mathrm{d}\boldsymbol{x} \right. \right.
$$
$$
\left. - \int_{+-} \sigma p \mathrm{d}\boldsymbol{x} \int_{+-} \frac{\partial p}{\partial \lambda_i^{+-}} \mathrm{d}\boldsymbol{x} \right] - \dot\lambda_i^{-+} \int_{+-} \sigma p \mathrm{d}\boldsymbol{x} \int_{-+} \frac{\partial p}{\partial \lambda_i^{-+}} \mathrm{d}\boldsymbol{x}
$$
$$
\left. - \dot\lambda_i^{--} \int_{+-} \sigma p \mathrm{d}\boldsymbol{x} \int_{--} \frac{\partial p}{\partial \lambda_i^{--}} \mathrm{d}\boldsymbol{x} \right\} + \sum_{i=1}^{M} \dot\lambda_{N+i} \left[\int_{+-} \sigma \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right]
$$
$$
- \int_{+-} \sigma p \mathrm{d}\boldsymbol{x} \int_{+-} \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right] + \int_{+-} \sigma \frac{\partial}{\partial \phi}(pS) \mathrm{d}\boldsymbol{x} = 0, \tag{19b}
$$

$$
\sum_{i=1}^{N} \left\{ -\dot\lambda_i^{++} \int_{-+} \sigma p \mathrm{d}\boldsymbol{x} \int_{++} \frac{\partial p}{\partial \lambda_i^{++}} \mathrm{d}\boldsymbol{x} - \dot\lambda_i^{+-} \int_{-+} \sigma p \mathrm{d}\boldsymbol{x} \int_{+-} \frac{\partial p}{\partial \lambda_i^{+-}} \mathrm{d}\boldsymbol{x} \right.
$$
$$
+ \dot\lambda_i^{-+} \left[\int_{-+} \sigma \frac{\partial p}{\partial \lambda_i^{-+}} \mathrm{d}\boldsymbol{x} - \int_{-+} \sigma p \mathrm{d}\boldsymbol{x} \int_{-+} \frac{\partial p}{\partial \lambda_i^{-+}} \mathrm{d}\boldsymbol{x} \right]
$$
$$
\left. - \dot\lambda_i^{--} \int_{-+} \sigma p \mathrm{d}\boldsymbol{x} \int_{--} \frac{\partial p}{\partial \lambda_i^{--}} \mathrm{d}\boldsymbol{x} \right\} + \sum_{i=1}^{M} \dot\lambda_{N+i} \left[\int_{-+} \sigma \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right]
$$
$$
- \int_{-+} \sigma p \mathrm{d}\boldsymbol{x} \int_{-+} \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right] + \int_{-+} \sigma \frac{\partial}{\partial \phi}(pS) \mathrm{d}\boldsymbol{x} = 0, \tag{19c}
$$

$$
\sum_{i=1}^{N} \left\{ -\dot\lambda_i^{++} \int_{--} \sigma p \mathrm{d}\boldsymbol{x} \int_{++} \frac{\partial p}{\partial \lambda_i^{++}} \mathrm{d}\boldsymbol{x} - \dot\lambda_i^{+-} \int_{--} \sigma p \mathrm{d}\boldsymbol{x} \int_{+-} \frac{\partial p}{\partial \lambda_i^{+-}} \mathrm{d}\boldsymbol{x} \right.
$$
$$
- \dot\lambda_i^{-+} \int_{--} \sigma p \mathrm{d}\boldsymbol{x} \int_{-+} \frac{\partial p}{\partial \lambda_i^{-+}} \mathrm{d}\boldsymbol{x} + \dot\lambda_i^{--} \left[\int_{--} \sigma \frac{\partial p}{\partial \lambda_i^{--}} \mathrm{d}\boldsymbol{x} \right.
$$
$$
\left. \left. - \int_{--} \sigma p \mathrm{d}\boldsymbol{x} \int_{--} \frac{\partial p}{\partial \lambda_i^{--}} \mathrm{d}\boldsymbol{x} \right] \right\} + \sum_{i=1}^{M} \dot\lambda_{N+i} \left[\int_{--} \sigma \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right]
$$
$$
- \int_{--} \sigma p \mathrm{d}\boldsymbol{x} \int_{--} \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right] + \int_{--} \sigma \frac{\partial}{\partial \phi}(pS) \mathrm{d}\boldsymbol{x} = 0. \tag{19d}
$$

By further taking the sum of the four, we again obtain extra constraints:

$$
\sum_{i=1}^{N} \left\{ \dot\lambda_i^{+} \left[\int_{++} \sigma \frac{\partial p}{\partial \lambda_i^{++}} \mathrm{d}\boldsymbol{x} - \int_{++} \sigma p \mathrm{d}\boldsymbol{x} \int_{++} \frac{\partial p}{\partial \lambda_i^{++}} \mathrm{d}\boldsymbol{x} \right] \right.
$$
$$
+ \left[\int_{+-} \sigma \frac{\partial p}{\partial \lambda_i^{+-}} \mathrm{d}\boldsymbol{x} - \int_{+-} \sigma p \mathrm{d}\boldsymbol{x} \int_{+-} \frac{\partial p}{\partial \lambda_i^{+-}} \mathrm{d}\boldsymbol{x} \right]
$$
$$
+ \dot\lambda_i^{-+} \left[\int_{-+} \sigma \frac{\partial p}{\partial \lambda_i^{-+}} \mathrm{d}\boldsymbol{x} - \int_{-+} \sigma p \mathrm{d}\boldsymbol{x} \int_{-+} \frac{\partial p}{\partial \lambda_i^{-+}} \mathrm{d}\boldsymbol{x} \right]
$$
$$
\left. + \dot\lambda_i^{--} \left[\int_{--} \sigma \frac{\partial p}{\partial \lambda_i^{--}} \mathrm{d}\boldsymbol{x} - \int_{--} \sigma p \mathrm{d}\boldsymbol{x} \int_{--} \frac{\partial p}{\partial \lambda_i^{--}} \mathrm{d}\boldsymbol{x} \right] \right\}
$$
$$
+ \sum_{i=1}^{M} \dot\lambda_{N+i} \left[\int \sigma \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} - \int \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right]
$$
$$
+ \int \sigma \frac{\partial}{\partial \phi}(pS) \mathrm{d}\boldsymbol{x} = 0. \tag{20a}
$$

Especially when σ does not depend on the direction distinguishing the subscripts $\pm$ in the integrals above, the first N-sum disappears in Eq. (20a) for the separable distributions; therefore, we obtain the following:

$$
\sum_{i=1}^{M} \dot\lambda_{N+i} \left[\int \sigma \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} - \int \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_{N+i}} \mathrm{d}\boldsymbol{x} \right]
$$
$$
+ \int \sigma \frac{\partial}{\partial \phi}(pS) \mathrm{d}\boldsymbol{x} = 0. \tag{20b}
$$

The formulations in these two last subsections will be applied in more specific cases considered in Sects. 4 and 5.

3   Example: Gaussian distribution with two variables

As a specific example of a multidimensional case, we consider a two-dimensional Gaussian distribution in this section, setting $\boldsymbol{x} = (x, y)$ and $\boldsymbol{S} = (S_x, S_y)$:

$$p = p_0 \exp[-\lambda_1(x - \overline{x})^2 - \lambda_2(y - \overline{y})^2 - \lambda_3(x - \overline{x})(y - \overline{y})] \quad (21)$$

(cf. Golaz et al., 2002; Larson and Golaz, 2005). From the normalization condition,

$$p_0 = \frac{(\lambda_1\lambda_2 + \lambda_3^2/4)^{1/2}}{\pi}, \quad (22)$$

as derived in Appendix A1.

For the distribution (Eq. 21), Eq. (18b) reduces to the following:

$$\sum_{i=1}^{3} \left[\int\int \sigma \frac{\partial p}{\partial \lambda_i} dx dy - \int\int \sigma p \, dx dy \int\int \frac{\partial p}{\partial \lambda_i} dx dy \right] \dot{\lambda}_i$$

$$+ \sum_{i=1}^{2} \left[\int\int \sigma \frac{\partial p}{\partial \overline{x}_i} dx dy - \int\int \sigma p \, dx dy \int\int \frac{\partial p}{\partial \overline{x}_i} dx dy \right] \dot{\overline{x}}_i$$

$$+ \int\int \sigma \nabla \cdot p \boldsymbol{S} \, dx dy = 0. \quad (23)$$

Here, the integral range is kept implicit, and the weight σ is given without a subscript for simplicity. By noting the relations

$$\frac{\partial p}{\partial \lambda_1} = -(x - \overline{x})^2 p,$$

$$\frac{\partial p}{\partial \lambda_2} = -(y - \overline{y})^2 p,$$

$$\frac{\partial p}{\partial \lambda_3} = -(x - \overline{x})(y - \overline{y}) p,$$

$$\frac{\partial p}{\partial \overline{x}} = [2\lambda_1(x - \overline{x}) + \lambda_3(y - \overline{y})] p,$$

$$\frac{\partial p}{\partial \overline{y}} = [2\lambda_2(y - \overline{y}) + \lambda_3(x - \overline{x})] p,$$

Eq. (23) further reduces to the following:

$$\dot{\lambda}_1 [-\langle (x - \overline{x})^2 \sigma \rangle + \langle \sigma \rangle \langle (x - \overline{x})^2 \rangle]$$

$$+ \dot{\lambda}_2 [-\langle (y - \overline{y})^2 \sigma \rangle + \langle \sigma \rangle \langle (y - \overline{y})^2 \rangle]$$

$$+ \dot{\lambda}_3 [-\langle (x - \overline{x})(y - \overline{y}) \sigma \rangle + \langle \sigma \rangle \langle (x - \overline{x})(y - \overline{y}) \rangle$$

$$+ \dot{\overline{x}} [2\lambda_1 \langle (x - \overline{x})\sigma \rangle + \lambda_3 \langle (y - \overline{y})\sigma \rangle]$$

$$+ \dot{\overline{y}} [2\lambda_2 \langle (y - \overline{y})\sigma \rangle + \lambda_3 \langle (x - \overline{x})\sigma \rangle] = \langle \boldsymbol{S} \cdot \nabla \sigma \rangle, \quad (24)$$

where $\langle \ \rangle$ suggests the phase-space average.

To proceed further, we note that the moments are given by

$$\langle x - \overline{x} \rangle = \langle y - \overline{y} \rangle = 0, \quad (25a)$$

$$\langle (x - \overline{x})^2 \rangle = \frac{1}{2\lambda_1} \left(1 - \frac{\lambda_3^2}{4\lambda_1\lambda_2} \right)^{-1}, \quad (25b)$$

$$\langle (y - \overline{y})^2 \rangle = \frac{1}{2\lambda_2} \left(1 - \frac{\lambda_3^2}{4\lambda_1\lambda_2} \right)^{-1}, \quad (25c)$$

$$\langle (x - \overline{x})(y - \overline{y}) \rangle = -\frac{\lambda_3}{4\lambda_1\lambda_2} \left(1 - \frac{\lambda_3^2}{4\lambda_1\lambda_2} \right)^{-1}, \quad (25d)$$

$$\langle (x - \overline{x})^3 \rangle = \langle (y - \overline{y})^3 \rangle = \langle (x - \overline{x})^2(y - \overline{y}) \rangle$$

$$= \langle (x - \overline{x})(y - \overline{y})^2 \rangle = 0, \quad (25e)$$

$$\langle (x - \overline{x})^4 \rangle = \frac{3}{4\lambda_1^2} \left(1 - \frac{\kappa}{4} \right)^{-2}, \quad (25f)$$

$$\langle (y - \overline{y})^4 \rangle = \frac{3}{4\lambda_2^2} \left(1 - \frac{\kappa}{4} \right)^{-2}, \quad (25g)$$

$$\langle (x - \overline{x})^2(y - \overline{y})^2 \rangle = \frac{1}{4\lambda_1\lambda_2} \left(1 + \frac{\kappa}{2} \right) \left(1 - \frac{\kappa}{4} \right)^{-2}, \quad (25h)$$

$$\langle (x - \overline{x})^3(y - \overline{y}) \rangle = -\frac{3\lambda_3}{8\lambda_1^2\lambda_2} \left(1 - \frac{\kappa}{4} \right)^{-2}, \quad (25i)$$

$$\langle (x - \overline{x})(y - \overline{y})^3 \rangle = -\frac{3\lambda_3}{8\lambda_1\lambda_2^2} \left(1 - \frac{\kappa}{4} \right)^{-2}, \quad (25j)$$

as derived in Appendix A2. Here,

$$\kappa = \frac{\lambda_3^2}{\lambda_1\lambda_2}, \quad (25k)$$

and the condition $4 > \kappa$ is required to ensure positive variances. To obtain the results for the third moments, we also need to assume $\lambda_1\lambda_2(4\lambda_1\lambda_2 - 3\lambda_3^2)^2 - \lambda_3^6 \neq 0$. The reader is referred to the Appendix for the derivations.

By using Eq. (25a)–(25j) for the moments, we apply varying weights, σ, on Eq. (24), focusing on the means, as already remarked. By setting $\sigma = x - \overline{x}$ and $\sigma = y - \overline{y}$, respectively, we obtain

$$\dot{\overline{x}} = \langle S_x \rangle, \quad (26a)$$

$$\dot{\overline{y}} = \langle S_y \rangle. \quad (26b)$$

The physical meaning of the results are clear: we recover the mean equations for the evolution.

Next, we set $\sigma = (x - \overline{x})^2$, $\sigma = (y - \overline{y})^2$, and $\sigma = (x - \overline{x})(y - \overline{y})$ in Eq. (24). This leads to the following:

$$-\frac{\dot{\lambda}_1}{\lambda_1} - \frac{\kappa}{4}\frac{\dot{\lambda}_2}{\lambda_2} + \frac{\kappa}{2}\frac{\dot{\lambda}_3}{\lambda_3} = 4\lambda_1 \left(1 - \frac{\kappa}{4} \right)^2 \langle (x - \overline{x})S_x \rangle, \quad (27a)$$

$$-\frac{\kappa}{4}\frac{\dot{\lambda}_1}{\lambda_1} - \frac{\dot{\lambda}_2}{\lambda_2} + \frac{\kappa}{2}\frac{\dot{\lambda}_3}{\lambda_3} = 4\lambda_2 \left(1 - \frac{\kappa}{4} \right)^2 \langle (y - \overline{y})S_y \rangle, \quad (27b)$$

$$\frac{\dot{\lambda}_1}{\lambda_1} + \frac{\dot{\lambda}_2}{\lambda_2} - \left(1 + \frac{\kappa}{4} \right)\frac{\dot{\lambda}_3}{\lambda_3} = \frac{4\lambda_3}{\kappa} \left(1 - \frac{\kappa}{4} \right)^2 \langle (y - \overline{y})S_x$$

$$+ (x - \overline{x})S_y \rangle, \quad (27c)$$

where κ has been defined by Eq. (25k). Equation (27a), (27b), and (27c) can be rewritten into three separate prognostic equations for λ_j $(j = 1, 2, 3)$ by a matrix inversion of the left-hand side.

To see this last procedure, we may rewrite Eq. (27a), (27b), and (27c) as follows:

$$\alpha_{11}\frac{\dot{\lambda}_1}{\lambda_1} + \alpha_{12}\frac{\dot{\lambda}_2}{\lambda_2} + \alpha_{13}\frac{\dot{\lambda}_3}{\lambda_3} = f_1,$$

$$\alpha_{21}\frac{\dot{\lambda}_1}{\lambda_1} + \alpha_{22}\frac{\dot{\lambda}_2}{\lambda_2} + \alpha_{23}\frac{\dot{\lambda}_3}{\lambda_3} = f_2,$$

$$\alpha_{31}\frac{\dot{\lambda}_1}{\lambda_1} + \alpha_{32}\frac{\dot{\lambda}_2}{\lambda_2} + \alpha_{33}\frac{\dot{\lambda}_3}{\lambda_3} = f_3.$$

The matrix inversion of the above leads to the following:

$$\frac{\dot{\lambda}_1}{\lambda_1} = \frac{1}{D}[(\alpha_{22}\alpha_{33} - \alpha_{23}\alpha_{32})f_1 + (-\alpha_{12}\alpha_{33} + \alpha_{13}\alpha_{32})f_2$$
$$+ (\alpha_{12}\alpha_{23} - \alpha_{13}\alpha_{22})f_3], \tag{28a}$$

$$\frac{\dot{\lambda}_2}{\lambda_2} = \frac{1}{D}[(\alpha_{23}\alpha_{31} - \alpha_{21}\alpha_{33})f_1 + (-\alpha_{13}\alpha_{31} + \alpha_{11}\alpha_{33})f_2$$
$$+ (\alpha_{13}\alpha_{21} - \alpha_{11}\alpha_{23})f_3], \tag{28b}$$

$$\frac{\dot{\lambda}_3}{\lambda_3} = \frac{1}{D}[(\alpha_{21}\alpha_{32} - \alpha_{22}\alpha_{31})f_1 + (-\alpha_{11}\alpha_{32} + \alpha_{12}\alpha_{31})f_2$$
$$+ (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21})f_3], \tag{28c}$$

where

$$D = \alpha_{11}(\alpha_{22}\alpha_{33} - \alpha_{23}\alpha_{32}) + \alpha_{12}(-\alpha_{21}\alpha_{33} + \alpha_{23}\alpha_{31})$$
$$+ \alpha_{13}(\alpha_{21}\alpha_{32} - \alpha_{22}\alpha_{31})$$
$$= \alpha_{21}(\alpha_{32}\alpha_{13} - \alpha_{33}\alpha_{12}) + \alpha_{22}(-\alpha_{31}\alpha_{13} + \alpha_{33}\alpha_{11})$$
$$+ \alpha_{23}(\alpha_{31}\alpha_{12} - \alpha_{32}\alpha_{11})$$
$$= \alpha_{31}(\alpha_{12}\alpha_{23} - \alpha_{13}\alpha_{22}) + \alpha_{32}(-\alpha_{11}\alpha_{23} + \alpha_{13}\alpha_{21})$$
$$+ \alpha_{33}(\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}). \tag{28d}$$

In the following two sections, we apply the general formulations developed over the last two sections to two specific systems: (i) the convective energy-cycle system, introduced by Yano and Plant (2012) in Sect. 4, and (ii) the Lorenz (1963) strange-attractor system in Sect. 5. The derived prognostic equations for the PDF parameters are integrated by the fourth-order Runge–Kutta method with a time step depending the assumed-PDF model.

4 Convective energy-cycle system (Yano and Plant, 2012)

The convective energy-cycle system introduced by Yano and Plant (2012) is presented in a nondimensional form as follows:

$$\dot{x} = xy, \tag{29a}$$
$$\dot{y} = -x + 1, \tag{29b}$$

where x is the convective kinetic energy (mass flux) and y is the cloud work function (a measure of potential energy). The equilibrium is always at $(x, y) = (1, 0)$ and $x > 0$.

An explicit calculation of the evolution of the initial uncertainty distribution with this system has already been performed by Yano and Ouchter (2017) using the Liouville

equation. In Fig. 1 of this work, their result is reproduced in a different format compared with the original paper. Here, the initial condition is a very localized Gaussian distribution, as shown in Fig. 1a. Note that the characteristics of the subsequent evolution are qualitatively different from those of the assumed-PDF forms introduced in the following. The main challenge of the assumed-PDF approach is, nevertheless, to predict the bulk statistics of the system fairly accurately.

4.1 Combination of the gamma and the Gaussian distributions (Model I)

As the system is semi-infinite in the x direction, whereas y extends to infinity to both sides (cf. Yano and Plant, 2012), the most natural choice for the assumed-PDF form for this system, following the argument in YLP, is to adopt a gamma distribution in the x direction and a Gaussian distribution in the y direction. This is referred as Model I in the following. Thus,

$$p(x, y) = p_1(x)p_2(y), \tag{30}$$

where

$$p_1(x) = p_{10}x^\mu \exp(-\lambda_1 x), \quad p_{10} = \frac{\lambda_1^{\mu+1}}{\Gamma(\mu+1)}, \tag{31a}$$

$$p_2(y) = p_{20}\exp[-\lambda_2(y - \overline{y})^2], \quad p_{20} = (\lambda_2/\pi)^{1/2} \tag{31b}$$

(cf. Appendix A1). In general, the Liouville equation leads to Eq. (6); in this case, this reduces to the following:

$$\dot{\mu}\left[\iint \sigma\frac{\partial p}{\partial\mu}dxdy - \iint \sigma p\,dxdy\iint \frac{\partial p}{\partial\mu}dxdy\right]$$
$$+ \dot{\overline{y}}\left[\iint \sigma\frac{\partial p}{\partial\overline{y}}dxdy - \iint \sigma p\,dxdy\iint \frac{\partial p}{\partial\overline{y}}dxdy\right]$$
$$+ \sum_{i=1}^{2}\dot{\lambda}_i\left[\iint \sigma\frac{\partial p}{\partial\lambda_i}dxdy - \iint \sigma p\,dxdy\iint \frac{\partial p}{\partial\lambda_i}dxdy\right]$$
$$= \iint p\mathbf{S}\cdot\nabla\sigma\,dxdy, \tag{32}$$

where

$$\frac{\partial p}{\partial\mu} = p\log x, \tag{33a}$$

$$\frac{\partial p}{\partial\overline{y}} = 2\lambda_2(y - \overline{y})p, \tag{33b}$$

$$\frac{\partial p}{\partial\lambda_1} = -px, \tag{33c}$$

$$\frac{\partial p}{\partial\lambda_2} = -(y - \overline{y})^2 p. \tag{33d}$$

By substituting Eq. (33a), (33b), (33c), and (33d) into Eq. (32), we obtain the following:

$$\dot{\mu}[\langle\sigma\log x\rangle - \langle\sigma\rangle\langle\log x\rangle] + 2\lambda_2\dot{\overline{y}}[\langle\sigma(y - \overline{y})\rangle - \langle\sigma\rangle\langle y - \overline{y}\rangle]$$
$$- \dot{\lambda}_1[\langle\sigma x\rangle - \langle\sigma\rangle\langle x\rangle] - \dot{\lambda}_2[\langle\sigma(y - \overline{y})^2\rangle - \langle\sigma\rangle\langle(y - \overline{y})^2\rangle]$$
$$= \langle\mathbf{S}\cdot\nabla\sigma\rangle. \tag{34}$$
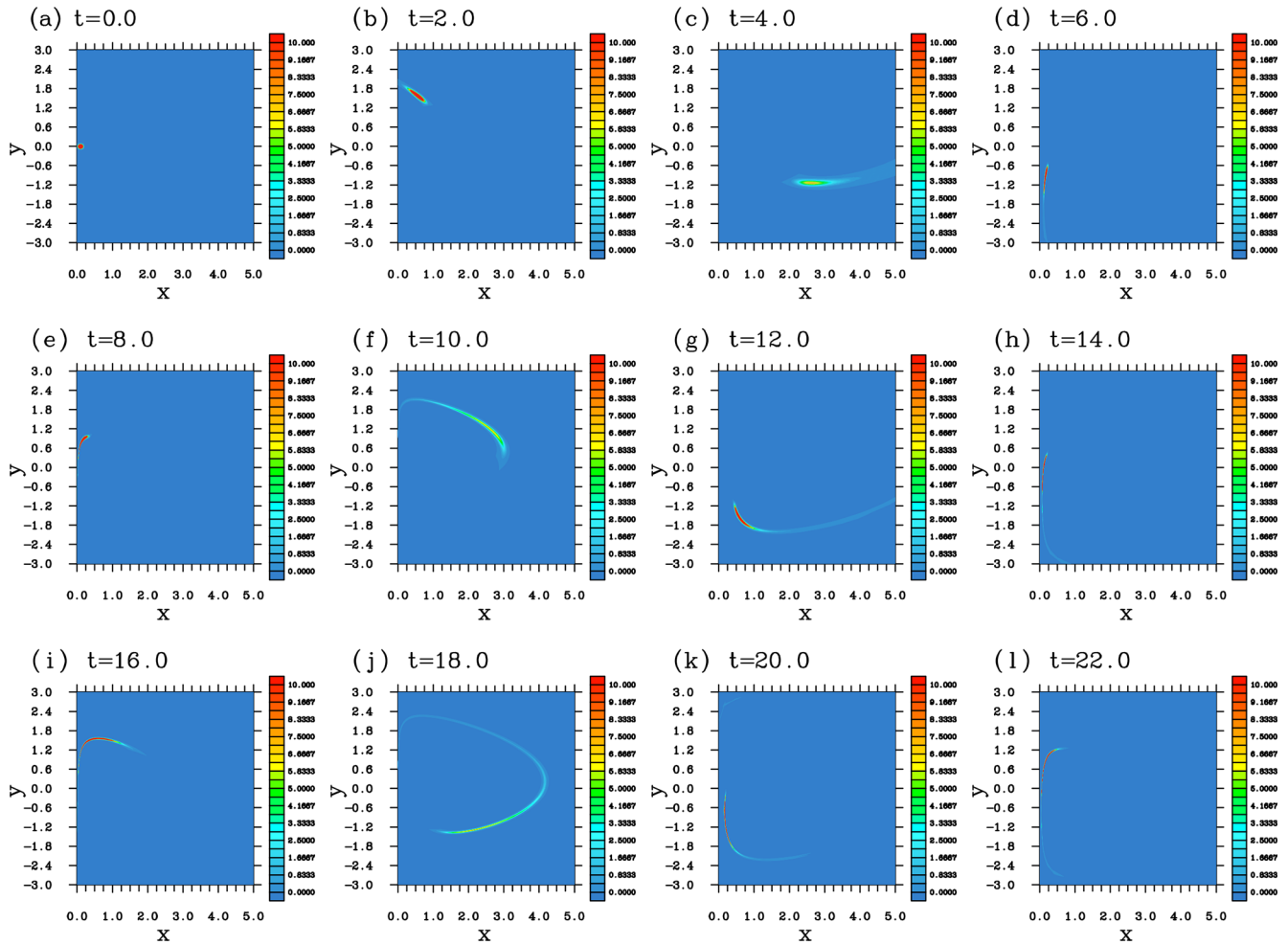
Figure 1. Snapshots of the evolution of the PDF with the convective energy cycle using a direct numerical integration of the Liouville equation.

Recall that $\langle\ \rangle$ suggests the phase-space average.

We are going to derive prognostic equations for the four parameters, μ, $\bar{y}$, and λ_i $(i = 1, 2)$ from Eq. (34) by choosing four different weights, σ. Here, the first two only depend on x, whereas the last two only depend on y. The first two choices are $\sigma = x$ and $\sigma = x^2$. Substitutions of these weights into Eq. (34) lead to the following:

$$\frac{\dot{\mu}}{\lambda_1} - (\mu+1)\frac{\dot{\lambda}_1}{\lambda_1^2} = \langle S_x \rangle, \tag{35a}$$

$$\frac{2\mu+3}{\lambda_1^2}\dot{\mu} - \frac{2(\mu+2)(\mu+1)}{\lambda_1^3}\dot{\lambda}_1 = 2\langle xS_x \rangle, \tag{35b}$$

noting $\langle\sigma(y-\bar{y})\rangle - \langle\sigma\rangle\langle y-\bar{y}\rangle = 0$ and $\langle\sigma(y-\bar{y})^2\rangle = \langle\sigma\rangle\langle(y-\bar{y})^2\rangle = 0$ for both $\sigma = x$ and $\sigma = x^2$. We have also noted the following relations:

$$\langle x\log x \rangle = \frac{1}{\lambda_1}[(\mu+1)\langle\log x\rangle + 1],$$

$$\langle x^2\log x \rangle = \frac{1}{\lambda_1^2}[(\mu+1)(\mu+2)\langle\log x\rangle + 2\mu + 3].$$

By combining Eq. (35a) and (35b), we further obtain stand-alone prognostic equations for these two individual parameters:

$$\dot{\mu} = 2[(\mu+2)\langle S_x\rangle - \lambda_1\langle xS_x\rangle], \tag{35c}$$

$$\dot{\lambda}_1 = \frac{\lambda_1^2}{\mu+1}[(2\mu+3)\langle S_x\rangle - 2\lambda_1\langle xS_x\rangle]. \tag{35d}$$

Here,

$$S_x = xy, \tag{36a}$$

$$\langle S_x \rangle = \langle xy \rangle = \langle x \rangle\langle y \rangle = \frac{\mu+1}{\lambda_1}\bar{y}, \tag{36b}$$

$$\langle xS_x \rangle = \langle x^2y \rangle = \langle x^2\rangle\langle y\rangle = \frac{(\mu+2)(\mu+1)}{\lambda_1^2}\bar{y}. \tag{36c}$$

To derive these three results (Eq. 36a, b, c), we have used the following relations:

$$\langle x \rangle = \frac{\mu + 1}{\lambda_1}, \quad \langle x^2 \rangle = \frac{(\mu + 2)(\mu + 1)}{\lambda_1^2}, \quad \text{and} \quad \langle y \rangle = \overline{y}.$$

Using these three relations (Eq. 36a, b, c), the right-hand side of Eq. (35c) vanishes and

$$\dot{\mu} = 0, \tag{37a}$$

whereas Eq. (35d) reduces to

$$\dot{\lambda}_1 = -\lambda_1 \overline{y}. \tag{37b}$$

As for the other two weights depending only on y, we choose $\sigma = y - \overline{y}$ and $\sigma = (y - \overline{y})^2$. In reductions, we invoke the following relations:

$$\langle y - \overline{y} \rangle = \langle (y - \overline{y})^3 \rangle = 0, \quad \langle (y - \overline{y})^2 \rangle = \frac{1}{2\lambda_2},$$

$$\langle (y - \overline{y})^4 \rangle = \frac{3}{4\lambda_2^2}.$$

The final results are as follows:

$$\dot{\overline{y}} = 1 - \frac{\mu + 1}{\lambda_1}, \tag{37c}$$

$$\dot{\lambda}_2 = 0. \tag{37d}$$

Thus, Eq. (37a), (37b), (37c), and (37d) dictate the evolution of the distribution (Eq. 30), in which the two parameters, μ and λ_2, turn out to be constants of time. On the other hand, by combining Eq. (37b) and (37c), we find

$$-\frac{d\lambda_1}{\lambda_1 \overline{y}} = \frac{d\overline{y}}{1 - \frac{\mu + 1}{\lambda_1}} = dt.$$

The first two terms can be rearranged to the following:

$$-\frac{1}{\lambda_1} \left(1 - \frac{\mu + 1}{\lambda_1} \right) d\lambda_1 = \overline{y} d\overline{y},$$

which can readily be integrated, and the trajectory $(\overline{x}, \overline{y})$ of the system (noting that $\overline{x} = (\mu + 1)/\lambda_1$) is found as follows:

$$2\overline{x} - 2\log\overline{x} + \overline{y}^2 = R, \tag{38}$$

where R is a constant, noting that μ is constant in time based on Eq. (37a). By comparing this final expression with Eq. (25) of Yano and Plant (2012), we find that the mean trajectory is identical to that of the solution of the system (Eq. 29a, b). Note that this is not necessarily the case. In fact, the numerical result in Fig. 2 shows that a full solution presents a damping circular trajectory in the phase space of $(\overline{x}, \overline{y})$ towards the equilibrium $(1, 0)$. This aspect is simply not captured by the given assumed PDF. Thus, a remedy to it is to be sought.
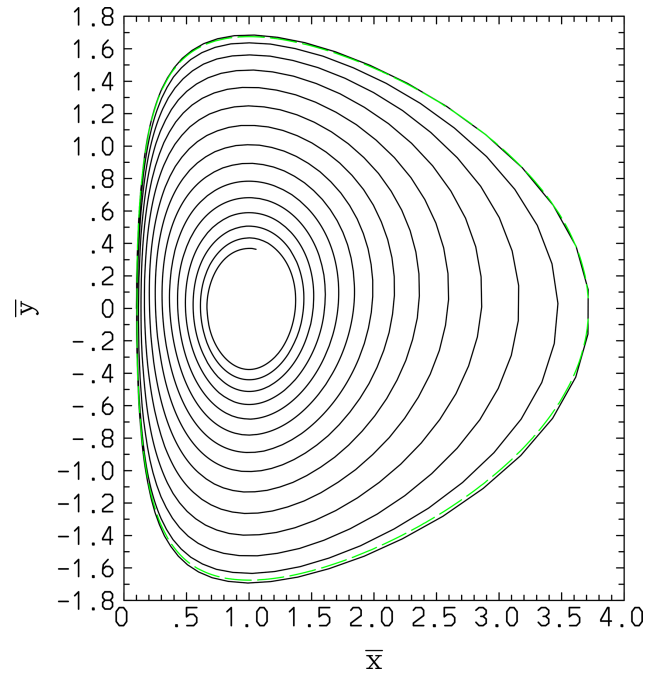

Figure 2. Trajectory of $(\overline{x}, \overline{y})$ as directly obtained from the Liouville equation (solid black lines) and by the assumed-PDF (Eqs. 30, 31a, b) and Model I (dashed green line) results. Here, $\mu = 10.2271805$ and $R = 2.19206810$ in Eq. (38). This assumed-PDF solution simply takes a closed orbit, failing to reproduce the damping tendency of the actual PDF evolution.

Figure 3 shows further statistical quantifications of the performance of Model I. Here, the adopted time step is $\Delta t = 1 \times 10^{-2}$. As already remarked, this model predicts a simple periodic cycle for the mean values and fails to reproduce the damping tendency of the actual PDF evolution (Fig. 3a). The same follows for the x variance (Fig. 3b), whereas Model I does not predict the evolution of the y variance – the latter is constant in time according to Eq. (37d).

An obvious defect of this assumed form is traced to the lack of correlation between the two variables, and the key nonlinear contribution, $\langle x' y' \rangle$, drops out from the set of evolution equations, where $x' = x - \langle x \rangle$ and $y' = y - \langle y \rangle$.

4.2 Two-dimensional Gaussian distribution (Model II)

Model I (Eq. 30) in the last subsection fails to predict the dissipating tendency with respect to the mean and the amplifying tendency with respect to the variance, as seen in Figs. 2 and 3. This reason for this is traced to the lack of a correlation between the two dependent variables, x and y, in the distribution; thus, the problem of the prediction of the mean becomes identical to that of the original dynamical system.

The most straightforward modification to overcome this issue is to modify the distribution as follows:

$$p(x, y) = p_0 x^\mu \exp[-\lambda_1 x - \lambda_2 (y - \overline{y})^2 - \lambda_3 x (y - \overline{y})].$$
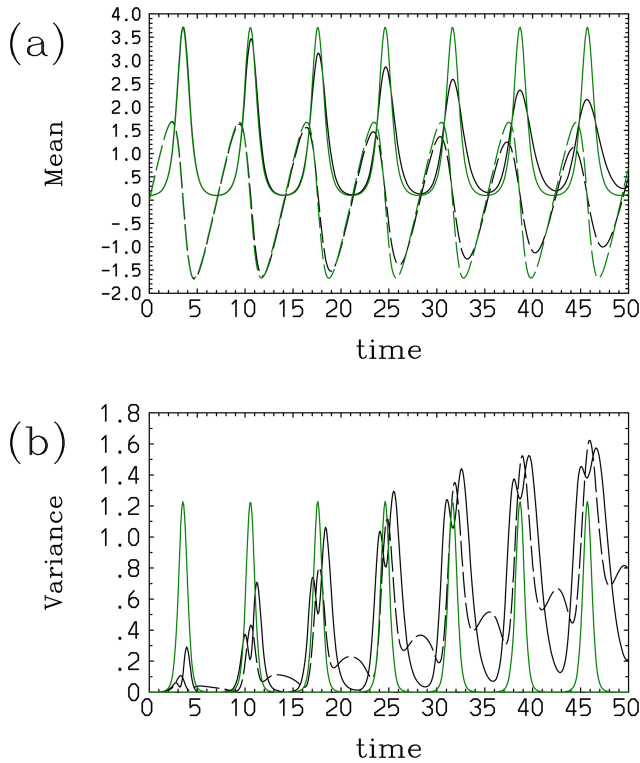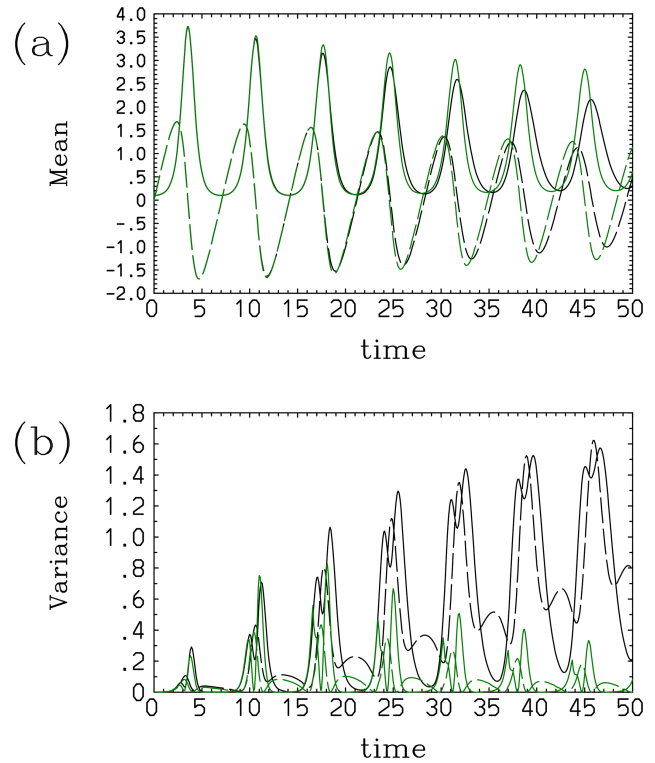
Figure 3. Comparisons of the statistics of the convective energy cycle between results with a direct computation of the Liouville equation (black) and that based on the assumed-PDF method (Model I, green): **(a)** means, $\overline{x}$ (solid line) and $\overline{y}$ (long-dashed line), and **(b)** variances, $\langle (x - \overline{x})^2 \rangle$ (solid line) and $\langle (y - \overline{y})^2 \rangle$ (long-dashed line).

However, this assumed PDF has an unfavorable feature in that the resulting integrals become impossible to perform analytically. The requirement for numerical integrals increases the computational cost and, therefore, effectively negates the advantage of the assumed-PDF approach: an integral over an infinite domain at each time step, as required, is substantially more numerically expensive than just predicting a few PDF parameters.

To avoid this difficulty, we instead adopt a Gaussian distribution with the two variables considered in Sect. 3 (Eq. 21): this model is referred as Model II. A minor disadvantage with this assumed PDF is that a distribution can spread to $x < 0$. However, this disadvantage has no serious consequence as long as we focus our attention on the basic statistics (the mean and variance, as well as correlation) and the mean of x remains positive (i.e., $\overline{x} > 0$).

By applying the expression of the source term for this system (Eq. 29a, b) to general formulas in Sec. 3, we obtain the following prognostic equation set for the assumed-PDF parameters:

$$\dot{\overline{x}} = \overline{x}\,\overline{y} - \frac{\kappa}{4\lambda_3}\left(1 - \frac{\kappa}{4}\right)^{-1}, \tag{39a}$$

$$\dot{\overline{y}} = -\overline{x} + 1, \tag{39b}$$

Figure 4. Comparisons of the statistics of the convective energy cycle between results with a direct computation of the Liouville equation (black) and that based on the assumed-PDF method (Model II, green): **(a)** mean, $\overline{x}$ (solid line) and $\overline{y}$ (long-dashed line), and **(b)** variance, $\langle (x - \overline{x})^2 \rangle$ (solid line) and $\langle (y - \overline{y})^2 \rangle$ (long-dashed line).

$$-\frac{\dot{\lambda}_1}{\lambda_1} - \frac{\kappa}{4}\frac{\dot{\lambda}_2}{\lambda_2} + \frac{\kappa}{2}\frac{\dot{\lambda}_3}{\lambda_3} = \left(1 - \frac{\kappa}{4}\right)\left(-\frac{\lambda_3}{\lambda_2}\overline{x} + 2\overline{y}\right), \tag{39c}$$

$$-\frac{\kappa}{4}\frac{\dot{\lambda}_1}{\lambda_1} - \frac{\dot{\lambda}_2}{\lambda_2} + \frac{\kappa}{2}\frac{\dot{\lambda}_3}{\lambda_3} = \left(1 - \frac{\kappa}{4}\right)\frac{\lambda_3}{\lambda_1}, \tag{39d}$$

$$\frac{\dot{\lambda}_1}{\lambda_1} + \frac{\dot{\lambda}_2}{\lambda_2} - \left(1 + \frac{\kappa}{4}\right)\frac{\dot{\lambda}_3}{\lambda_3} = \left(1 - \frac{\kappa}{4}\right)\left(-\frac{2\lambda_2}{\lambda_3} + \frac{2\lambda_1}{\lambda_3}\overline{x} - \overline{y}\right). \tag{39e}$$

Recall that κ is defined by Eq. (25k):

$$\kappa = \frac{\lambda_3^2}{\lambda_1\lambda_2}. \tag{40}$$

The initial condition in this case is set equal to that of the full Liouville run, which is also initialized with a Gaussian distribution.

Characteristics of the evolution of the system with Model II obtained with the time step of $\Delta t = 1 \times 10^{-4}$ are shown in Fig. 4: the mean values (panel a) decay as in the case with the explicit Liouville run. The agreement is almost perfect up to the end of the third cycle, but a difference gradually becomes noticeable. A periodic increase in the variance (panel b) is also predicted up to the second cycle. However, after the end of the third cycle, the variances predicted by

Model II rapidly decrease with time. The model is numerically unstable and blows up after $t = 60$ with $\Delta t = 10^{-2}$. Euler time-stepping is also attempted: it crashes at $t = 23.5$. These behaviors can be understood, to a good extent, simply by inspecting the obtained prediction equations. Even using Model II, the evolution of $\overline{y}$ (Eq. 39b) still follows that of y, replacing the terms by the averages, as is also the case with Model I. However, an extra term is found for the evolution equation for $\overline{x}$ (Eq. 39a), which can work as a damping term whenever $1 - \kappa/4 > 0$ is satisfied, as expected from the full Liouville simulation. Due to this dissipating tendency of $\overline{x}$, $\overline{y}$ also dissipates with time to the extent that the former dissipates, as seen from the long-dashed curves in Fig. 4a.

The condition $1 - \kappa/4 > 0$ is indeed satisfied initially with a Gaussian distribution leading to $\kappa = 0$, with $\lambda_3 = 0$. However, the distribution evolves with increasing κ with time, approaching $\kappa = 4$. Thus, this dissipative term also becomes smaller with time, and (at a certain point) $\overline{x}$ no longer dissipates as effectively as it does in the full Liouville solution, as seen from the solid curves in Fig. 4a. The variances also grow with time as long as $1 - \kappa/4 > 0$, following Eq. (39c), (39d), and (39e). However, as remarked, due to the tendency of $\kappa \to 4$, this growing tendency only continues over the first three convective cycles; the variances then begin to diminish with time, as seen in Fig. 4b. These results suggest that adding a cross term (i.e., $\lambda_3 \neq 0$) in the distribution is not sufficient to reproduce statistical tendencies for an extensive duration of the simulation, especially because the variances tend to collapse after initial realistic tendencies with growing variances.

4.3 Alternative possibility

As an alternative possibility for the distributions in Sect. 4.2, the form

$$p(x, y) = \begin{cases} p_0 x^\mu \exp(-\lambda_1 x - \lambda_2^+ y) & y > 0 \\ p_0 x^\mu \exp(-\lambda_1 x + \lambda_2^- y) & y < 0 \end{cases}$$

is also considered. As it turns out, the evolution of the mean values, $\overline{x}$ and $\overline{y}$, remains identical to that of Model I without any damping tendency. A slight improvement is that the standard deviation in y, in this case, evolves with

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle y'^2 \rangle = \langle y \rangle \langle S_y \rangle.$$

Due to the fact that only limited improvements are expected, this case is not actually attempted. It transpires that it is crucial to include a dependence on xy in the distribution in order to successfully predict the damping tendency of mean values, as is the case with Model II.

5 The Lorenz (1963) system

The system proposed by Lorenz (1963) is given by the following:

$$\dot{x} = -Px + Py, \tag{41a}$$
$$\dot{y} = -xz + rx - y, \tag{41b}$$
$$\dot{z} = xy - bz. \tag{41c}$$

Here, we assume the following standard parameters: $b = 8/3$, the Rayleigh number, $r = 28$, and the Prandtl number, $P = 10$. The system consists of three unstable steady solutions (i.e., fixed points): two of them corresponding to steady convection are found at $(x, y, z) = (\pm 6\sqrt{2}, \pm 6\sqrt{2}, 27)$, while another is found at $(x, y, z) = (0, 0, 0)$.

First, as a reference, the evolution of the PDF is computed by directly integrating the Liouville equation. Here, the adopted numerics are identical to those used in Yano and Phillips (2016) for the Fokker–Planck equation, except that a treatment for the diffusion term arising from stochasticity is missing in the present case. The initial condition is a Gaussian distribution centered at $(0, 1, 0)$ with variances of 12.5 in all three directions. The initial center point is identical to the initial condition adopted by Lorenz (1963). An identical run has also been repeated by replacing the initial center of the Gaussian distribution to the origin, $(0, 0, 0)$, of the system. Some remarks on this latter case will also be added in the following.

Figures 5 and 6 show that evolution of the PDF is highly non-Gaussian on both the y–z and x–y planes. On both planes, the distribution splits into two peaks over time up to $t = 1$, corresponding to two unstable fixed points of the system. Distributions around these two peaks gradually diffuse with time as the strange attractor fully develops. Recall that the Lorenz system is chaotic; thus, the individual trajectories of solutions remain nonstationary throughout their evolution. However, as expected from the evolution of the distribution, the statistics of the system gradually converge to a stationary state after an initial transient period. Note that this is a major difference from a fully turbulent system to a low-dimensional chaotic system: in the former case, the statistics can also remain nonstationary throughout the evolution of the system. As will be demonstrated in the following, this aspect turns out to be the hardest to reproduce using an assumed-PDF form with only a few parameters, as a governing equation system predicting these PDF parameters can itself become a nonlinear chaotic system. This tendency most clearly emerges with Model I in the following.

In the following assumed-PDF demonstrations, a time step of $\Delta t = 10^{-4}$ has been adopted by default. Runs have also been repeated with $\Delta t = 2 \times 10^{-4}$, but no modification of the results has been found.
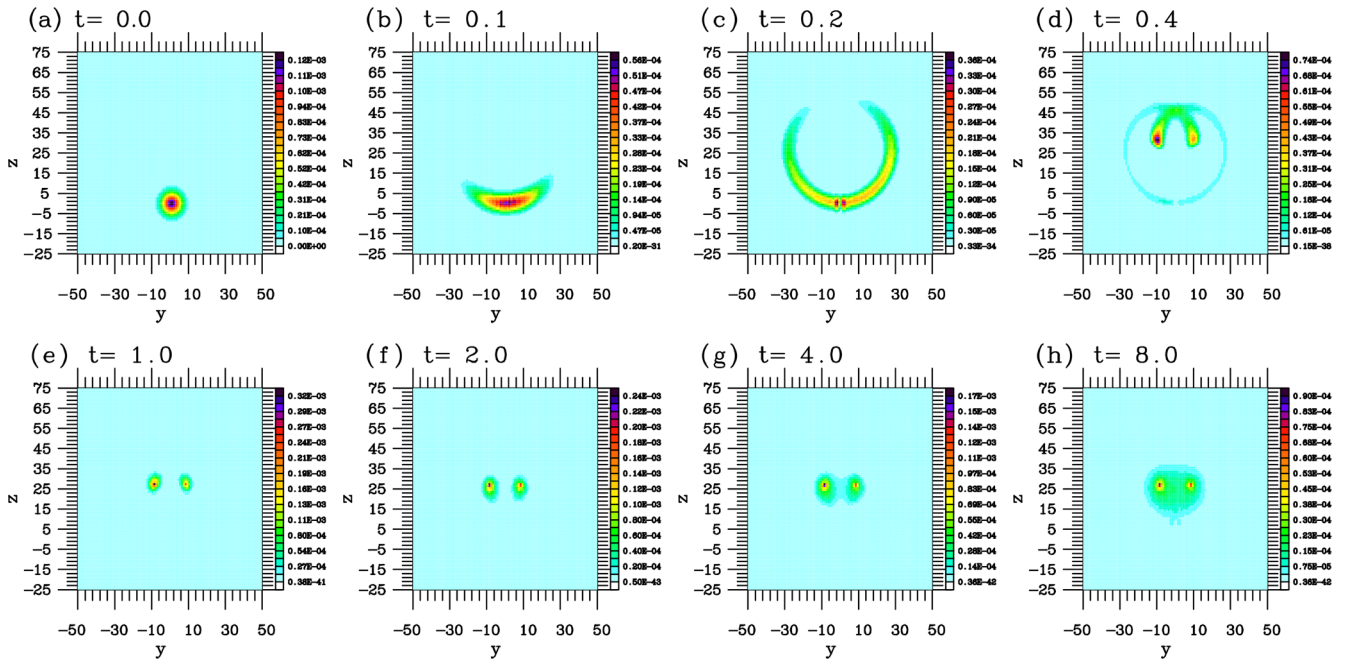
Figure 5. Time evolution of the PDF for the Lorenz system on the y–z plane as directly predicted by the Liouville equation.
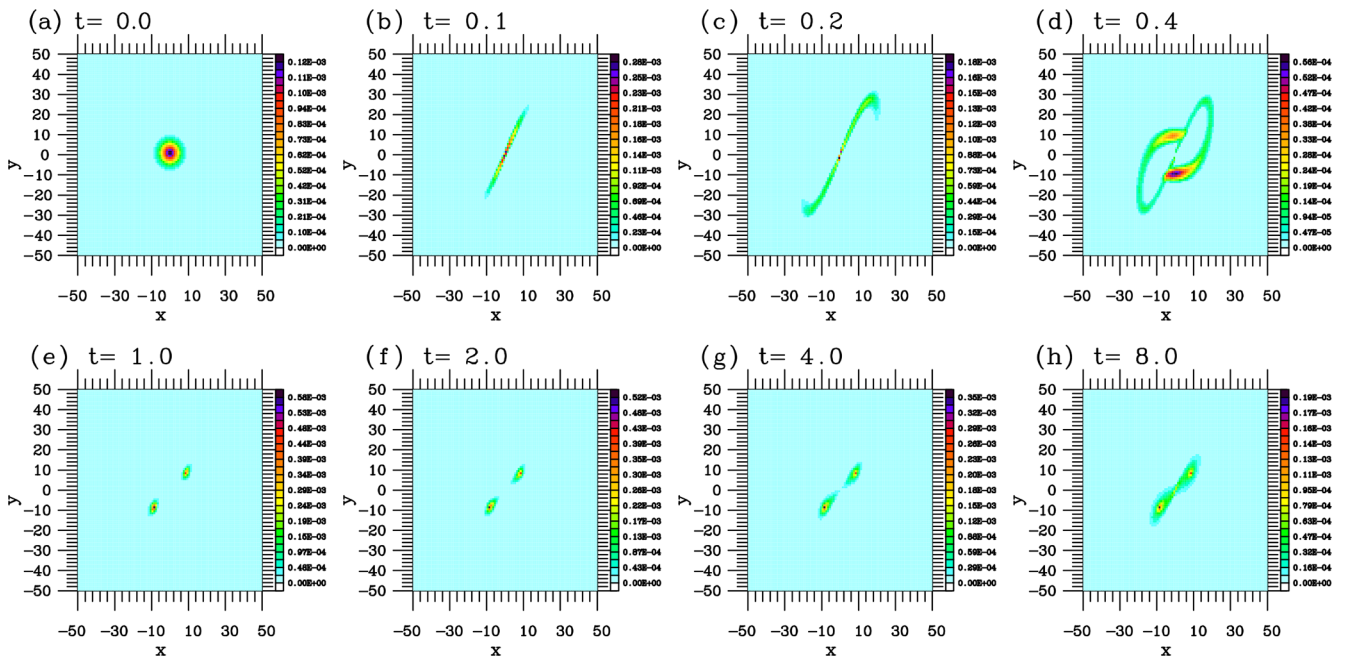

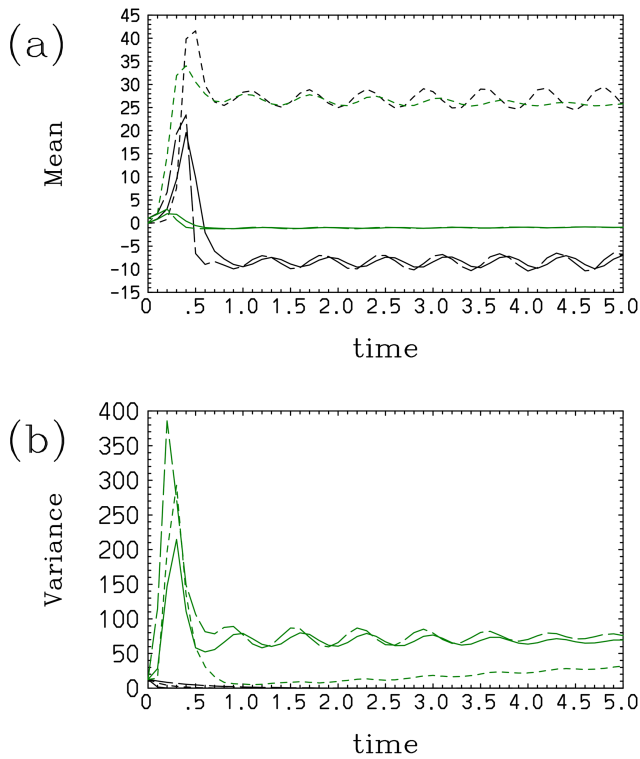**Figure 6.** The same as Fig. 5 but on the x–y plane.

Figure 7. The **(a)** mean and **(b)** variance of the Lorenz system with Model I: the model results (black) and results obtained by direct prediction with the Liouville equation (green). Here (and in the following plots), the curves are defined for the x, y, and z components using solid lines, long-dashed lines, and short-dashed lines, respectively.

5.1 The Lorenz (1963) system: Gaussian distribution (Model I)

As the first model (Model I), we consider a three-dimensional Gaussian distribution without correlations between the dependent variables, x, y, and z:

$$p(x, y, z) = p_1(x)p_2(y)p_3(z), \tag{42}$$

where

$$p_1(x) = p_{10} \exp[-\lambda_1(x - \overline{x})^2], \tag{43a}$$
$$p_2(y) = p_{20} \exp[-\lambda_2(y - \overline{y})^2], \tag{43b}$$
$$p_3(z) = p_{30} \exp[-\lambda_3(z - \overline{z})^2]. \tag{43c}$$

Here, the normalization conditions are as follows:

$$p_{j0} = (\lambda_j/\pi)^{1/2}, \tag{44}$$

where $j = 1, 2, 3$.

From the general formulation (Eq. 6), we obtain the following:

$$\sum_{j=1}^{3} \left\{ \dot{\lambda}_j \left[\int \sigma \frac{\partial p}{\partial \lambda_j} \mathrm{d}\boldsymbol{x} - \int \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_j} \mathrm{d}\boldsymbol{x} \right] \right.$$
$$\left. + \dot{\overline{x}}_j \left[\int \sigma \frac{\partial p}{\partial \overline{x}_j} \mathrm{d}\boldsymbol{x} - \int \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \overline{x}_j} \mathrm{d}\boldsymbol{x} \right] \right\}$$
$$= \int p \boldsymbol{S} \cdot \frac{\partial \sigma}{\partial \boldsymbol{x}} \mathrm{d}\boldsymbol{x}, \tag{45}$$

where $\boldsymbol{x} = (x_1, x_2, x_3) = (x, y, z)$ and σ is a weight. Moreover, we find the relations

$$\partial p/\partial \lambda_j = -(x_j - \overline{x}_j)^2 p, \tag{46a}$$
$$\partial p/\partial \overline{x}_j = 2\lambda_j(x_j - \overline{x}_j)p. \tag{46b}$$

By substituting them into Eq. (45), we find the following:

$$\sum_{j=1}^{3} \left\{ \dot{\lambda}_j \left[-\langle \sigma(x_j - \overline{x}_j)^2 \rangle + \langle \sigma \rangle \langle (x_j - \overline{x}_j)^2 \rangle \right] \right.$$
$$\left. + 2\lambda_j \dot{\overline{x}}_j \langle \sigma(x_j - \overline{x}_j) \rangle \right\} = \left\langle \boldsymbol{S} \cdot \frac{\partial \sigma}{\partial \boldsymbol{x}} \right\rangle. \tag{47}$$

We choose the weights $\sigma = x_j - \overline{x}_j$ and $\sigma = (x_j - \overline{x}_j)^2$ ($j = 1, 2, 3$), and Eq. (47) then reduces to the following:

$$\dot{\overline{x}}_j = \langle S_j \rangle, \tag{48a}$$
$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{\lambda_j} \right) = 4\langle (x_j - \overline{x}_j)S_j \rangle, \tag{48b}$$

for $j = 1, 2, 3$. From Eq. (41a), (41b), and (41c), the source terms are defined by

$$S_x = -Px + Py,$$
$$S_y = -xz + rx - y,$$
$$S_z = xy - bz.$$

Here,

$$\langle S_j \rangle = S_j(\overline{\boldsymbol{x}}), \tag{49}$$

and we also find the following by direct manipulations:

$$\langle (x - \overline{x})S_x \rangle = -P/2\lambda_1, \tag{50a}$$
$$\langle (y - \overline{y})S_y \rangle = -1/2\lambda_2, \tag{50b}$$
$$\langle (z - \overline{z})S_z \rangle = -b/2\lambda_3. \tag{50c}$$

Model I predicts the evolution of $\overline{z}$ (short-dashed line) reasonably, but $\overline{x}$ (solid line) and $\overline{y}$ (long-dashed line) somehow settle to one of two unstable fixed points (to the negative side, shown in black), without suggesting an alternative possibility: equal probability among these two fixed points makes the mean values, $\overline{x}$ and $\overline{y}$, close to zero in the explicit simulation (green line in Fig. 7a). Model I also fails to predict
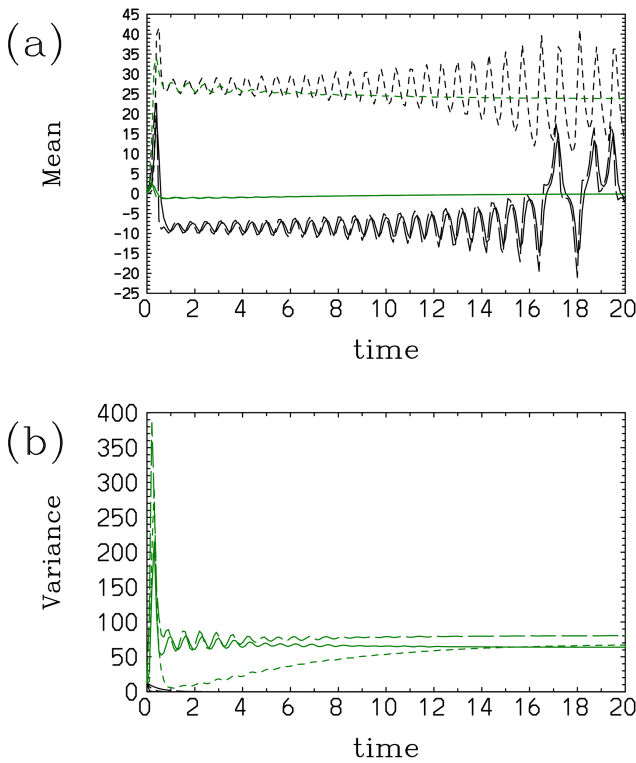
Figure 8. The same as Fig. 7 but for a longer duration.

an increase in the variances at an initial phase, as seen in the explicit simulation (green line); rather, the variances simply decay rapidly (black line in Fig. 7b). The latter behavior with respect to the variances has already been seen in Eq. (48b) and the definitions (Eq. 50a, b, c) of the right-hand-side forcing term: as the parameters, λ_j ($j = 1, 2, 3$), are defined to be positive definite, the variance can only decay with time.

Over a longer time, the aforementioned chaotic nature of the system emerges, as shown in Fig. 8. Note especially that the equation set for the means is identical to the original Lorenz attractor system; thus, the means never converge to a statistical equilibrium, as realized in a direct computation of the Liouville equation. Instead, after $t = 5$, the means gradually begin to oscillate around the tentatively settled unstable fixed points. A first transition happens a little after $t = 16$, over which $\overline{x}$ and $\overline{y}$ transit to a nonperiodic oscillation around the origin.

5.2 The Lorenz (1963) system: semi-exponential in the y direction (Model II)

The attempt in the last subsection to constrain the system in terms of the domain-averaged statistics failed to capture the tendency of the system to settle around two unstable fixed points; rather, the assumed PDF tends to settle to only one of the two fixed points. As a measure to alleviate this tendency, we constrain the PDF in terms of the averages over subdo-

mains of the system in this subsection. Using Model II, we now more specifically constrain the system as follows:

$$\overline{y}^+ = \langle y \rangle_+ \equiv \int_0^{+\infty} y\, p_2 \mathrm{d}y \Big/ \int_0^{+\infty} p_2 \mathrm{d}y,$$

$$\overline{y}^- = \langle y \rangle_- \equiv \int_{-\infty}^{0} y\, p_2 \mathrm{d}y \Big/ \int_{-\infty}^{0} p_2 \mathrm{d}y.$$

These constraints suggest a semi-exponential distribution in y direction:

$$p_2 = \begin{cases} p_{20}^+ \exp(-\lambda_2^+ y), & y > 0 \\ p_{20}^- \exp(\lambda_2^- y), & y < 0; \end{cases} \tag{51}$$

at $y = 0$,

$$p_{20}^+ = p_{20}^- = p_{20}. \tag{52}$$

Here, p_1 and p_3 in Eq. (42) remain the same as in Model I in Sect. 5.1.

From the normalization condition

$$\int_{-\infty}^{\infty} p_2 \mathrm{d}y = 1,$$

we find

$$p_{20} = (1/\lambda_2^+ + 1/\lambda_2^-)^{-1}. \tag{53}$$

Furthermore, let

$$p_+ \equiv \int_0^{+\infty} p\, \mathrm{d}y = \frac{p_{20}}{\lambda_2^+}, \tag{54a}$$

$$p_- \equiv \int_{-\infty}^{0} p\, \mathrm{d}y = \frac{p_{20}}{\lambda_2^-}. \tag{54b}$$

Note further,

$$\frac{\partial p}{\partial \lambda_2^+} = -yp,\ y > 0; \tag{55a}$$

$$\frac{\partial p}{\partial \lambda_2^-} = yp,\ y < 0; \tag{55b}$$

and

$$\overline{y}_+ = \int_0^{+\infty} yp\,\mathrm{d}y \int_0^{+\infty} p\,\mathrm{d}y = \frac{1}{\lambda_2^+}, \tag{56a}$$

$$\overline{y}_- = \int_{-\infty}^{0} yp\,\mathrm{d}y \int_{-\infty}^{0} p\,\mathrm{d}y = -\frac{1}{\lambda_2^-}, \tag{56b}$$

$$\overline{y} \equiv \int_{-\infty}^{+\infty} yp\mathrm{d}y = \frac{1}{\lambda_2^+} - \frac{1}{\lambda_2^-}. \tag{56c}$$

The prognostic equations for λ_1, $\lambda_2^{\pm}$, λ_3, and $\overline{x}_j$ $(j = 1, 3)$ are as follows:

$$\sum_{j=1,3} \left\{ \dot{\lambda}_j \left[\int_+ \sigma \frac{\partial p}{\partial \lambda_j} \mathrm{d}\boldsymbol{x} - \int_+ \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_j} \mathrm{d}\boldsymbol{x} \right] \right.$$
$$\left. + \dot{\overline{x}}_j \left[\int_+ \sigma \frac{\partial p}{\partial \overline{x}_j} \mathrm{d}\boldsymbol{x} - \int_+ \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \overline{x}_j} \mathrm{d}\boldsymbol{x} \right] \right\}$$
$$+ \dot{\lambda}_2^+ \left[\int_+ \sigma \frac{\partial p}{\partial \lambda_2^+} \mathrm{d}\boldsymbol{x} - \int_+ \sigma p \mathrm{d}\boldsymbol{x} \int_+ \frac{\partial p}{\partial \lambda_2^+} \mathrm{d}\boldsymbol{x} \right]$$
$$- \dot{\lambda}_2^- \int_+ \sigma p \mathrm{d}\boldsymbol{x} \int_- \frac{\partial p}{\partial \lambda_2^-} \mathrm{d}\boldsymbol{x} + \int_+ \sigma \frac{\partial}{\partial \boldsymbol{x}} (p\boldsymbol{S}) \mathrm{d}\boldsymbol{x} = 0, \tag{57a}$$

$$\sum_{j=1,3} \left\{ \dot{\lambda}_j \left[\int_- \sigma \frac{\partial p}{\partial \lambda_j} \mathrm{d}\boldsymbol{x} - \int_- \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_j} \mathrm{d}\boldsymbol{x} \right] \right.$$
$$\left. + \dot{\overline{x}}_j \left[\int_- \sigma \frac{\partial p}{\partial \overline{x}_j} \mathrm{d}\boldsymbol{x} - \int_- \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \overline{x}_j} \mathrm{d}\boldsymbol{x} \right] \right\}$$
$$- \dot{\lambda}_2^+ \int_- \sigma p \mathrm{d}\boldsymbol{x} \int_+ \frac{\partial p}{\partial \lambda_2^+} \mathrm{d}\boldsymbol{x}$$
$$+ \dot{\lambda}_2^- \left[\int_- \sigma \frac{\partial p}{\partial \lambda_2^-} \mathrm{d}\boldsymbol{x} - \int_- \sigma p \mathrm{d}\boldsymbol{x} \int \frac{\partial p}{\partial \lambda_2^-} \mathrm{d}\boldsymbol{x} \right]$$
$$+ \int_- \sigma \frac{\partial}{\partial \boldsymbol{x}} (p\boldsymbol{S}) \mathrm{d}\boldsymbol{x} = 0. \tag{57b}$$

Also recall that

$$\frac{\partial p}{\partial \lambda_j} = -(x_j - \overline{x}_j)^2 p,$$
$$\frac{\partial p}{\partial \overline{x}_j} = 2\lambda_j(x_j - \overline{x}_j)p,$$

for $j = 1, 3$, and further refer to Eq. (55a) and (55b). By substituting these expressions into Eq. (57a) and (57b),

$$\sum_{j=1,3} \dot{\lambda}_j [-\langle \sigma(x_j - \overline{x}_j)^2 \rangle_+ + \langle \sigma \rangle_+ \langle (x_j - \overline{x}_j)^2 \rangle]$$
$$+ 2\lambda_j \dot{\overline{x}}_j [\langle \sigma(x_j - \overline{x}_j) \rangle_+ - \langle \sigma \rangle_+ \langle x_j - \overline{x}_j \rangle]\}$$
$$+ \dot{\lambda}_2^+ [-\langle \sigma y \rangle_+ + p_+ \langle \sigma \rangle_+ \langle y \rangle_+] - p_- \dot{\lambda}_2^- \langle \sigma \rangle_+ \langle y \rangle_-$$
$$= \langle \boldsymbol{S} \cdot \nabla \sigma \rangle_+ + \int\int \sigma p S_y|_{y=0} \mathrm{d}x \mathrm{d}z / p_+, \tag{58a}$$

$$\sum_{j=1,3} \dot{\lambda}_j [-\langle \sigma(x_j - \overline{x}_j)^2 \rangle_- + \langle \sigma \rangle_- \langle (x_j - \overline{x}_j)^2 \rangle]$$
$$+ 2\lambda_j \dot{\overline{x}}_j [\langle \sigma(x_j - \overline{x}_j) \rangle_- - \langle \sigma \rangle_- \langle x_j - \overline{x}_j \rangle]\}$$
$$+ p_+ \dot{\lambda}_2^+ \langle \sigma \rangle_- \langle y \rangle_+ + \dot{\lambda}_2^- [\langle \sigma y \rangle_- - p_- \langle \sigma \rangle_- \langle y \rangle_-]$$
$$= \langle \boldsymbol{S} \cdot \nabla \sigma \rangle_- - \int\int \sigma p S_y|_{y=0} \mathrm{d}x \mathrm{d}z / p_-. \tag{58b}$$

When σ depends on only x or z,

$$\langle \sigma y \rangle_{\pm} - p_{\pm} \langle \sigma \rangle_{\pm} \langle y \rangle_{\pm} = 0.$$

The following four types of constraints (i–iv) are considered:

– (i) $\sigma = x_i - \overline{x}_i$ $(j = 1, 3)$.

Here, we note $\langle \sigma \rangle_+ = \langle \sigma \rangle_- = \langle \sigma \rangle$ as well as

$$\langle x_i - \overline{x}_i \rangle = \langle (x_i - \overline{x}_i)^3 \rangle = 0,$$
$$\langle (x_i - \overline{x}_i)^2 \rangle - \langle x_i - \overline{x}_i \rangle^2 = \langle (x_i - \overline{x}_i)^2 \rangle = 1/2\lambda_i,$$
$$\langle (x_i - \overline{x}_i)^4 \rangle = 3/4\lambda_i^2,$$

for $i = 1$ and 3. Then, we obtain

$$\dot{\overline{x}}_i = \langle S_i \rangle_+ + \int\int \sigma p S_y|_{y=0} \mathrm{d}x \mathrm{d}z / p_+, \tag{59a}$$
$$\dot{\overline{x}}_i = \langle S_i \rangle_- - \int\int \sigma p S_y|_{y=0} \mathrm{d}x \mathrm{d}z / p_-. \tag{59b}$$

By taking a weighted sum of these expressions,

$$\dot{\overline{x}}_i = \langle S_i \rangle. \tag{59c}$$

Equation (59c) predicts the evolution of $\overline{x}_i$ with $i = 1$, 3.

– (ii) $\sigma = (x_i - \overline{x}_i)^2$ $(j = 1, 3)$. It is noted that

$$\langle (x_i - \overline{x}_i)^4 \rangle - \langle (x_i - \overline{x}_i)^2 \rangle^2 = 1/2\lambda_i^2,$$

and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{\lambda_i} \right) - \frac{p_-}{\lambda_i} (\overline{y}_+ \dot{\lambda}_2^+ + \overline{y}_- \dot{\lambda}_2^-)$$
$$= 4\langle (x - \overline{x}_i)S_i \rangle_+ + 2 \int\int \sigma p S_y|_{y=0} \mathrm{d}x \mathrm{d}z / p_+, \tag{60a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{\lambda_i} \right) + \frac{p_+}{\lambda_i} (\overline{y}_+ \dot{\lambda}_2^+ + \overline{y}_- \dot{\lambda}_2^-)$$
$$= 4\langle (x - \overline{x}_i)S_i \rangle_- - 2 \int\int \sigma p S_y|_{y=0} \mathrm{d}x \mathrm{d}z / p_-, \tag{60b}$$

for $i = 1$, 3. By taking the weighted sum of the two expressions,

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{1}{\lambda_i} \right) = 4\langle (x - \overline{x}_i)S_i \rangle. \tag{60c}$$

Equation (60c) predicts the evolution of λ_i with $i = 1, 3$.

Next, note that the sum over $j = 1, 3$ drops out when σ depends only on y:

– (iii) $\sigma = 1$. Both Eq. (58a) and (58b) reduce to

$$\dot{\lambda}_2^+/\lambda_2^+ - \dot{\lambda}_2^-/\lambda_2^- = -(\lambda_2^+ + \lambda_2^-)\langle S_y' \rangle, \tag{61}$$

where

$$\langle S_y' \rangle = (r - \overline{z})\overline{x}.$$

Note that the system is overconstrained, when the $\sigma = 1$ (i.e., Eq. 61) condition is used along with the results with $\sigma = y$ (i.e., Eq. 62c, d). Thus, it is better not to consider the constraint (Eq. 61).

– (iv) $\sigma = y$. Thus,

$$(2 - p_+)\overline{y}_+^2 \dot{\lambda}_2^+ + p_- \overline{y}_- \overline{y}_+ \dot{\lambda}_2^- = -\langle S_y' \rangle + \overline{y}_+, \tag{62a}$$

$$p_+ \overline{y}_- \overline{y}_+ \dot{\lambda}_2^+ + (2 - p_-)\overline{y}_-^2 \dot{\lambda}_2^- = \langle S_y' \rangle - \overline{y}_-. \tag{62b}$$

From Eq. (62a, b), we further obtain the following:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\lambda_2^+}\right) = -\frac{1}{\lambda_2^+} + \frac{\langle S_y' \rangle}{2}, \tag{62c}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\lambda_2^-}\right) = -\frac{1}{\lambda_2^-} - \frac{\langle S_y' \rangle}{2}. \tag{62d}$$

For better numerical stability, the time integration is performed in terms of $1/\lambda_2^\pm$, rather than $\lambda_2^\pm$. Note that both $1/\lambda_2^\pm$ terms present a damping tendency due to the first terms on the right-hand side. Moreover, note that asymmetry arises from the second terms; thus, when $1/\lambda_2^+$ tends to grow, $1/\lambda_2^-$ tends to decay, and vice versa: the general tendency with this pair is qualitatively consistent with a case in which the probability of being on the positive and negative sides of y, $p_\pm$, follow the evolution tendency of the individual phase-space particle. The governing equations for the parameters for the distribution in the x and z directions remain unchanged, being constrained by conditions (i) and (ii) above.

Model II shows improvement over Model I with respect to reasonably predicting the y variance (long-dashed line in Fig. 9b): it grows but with a delay of about 0.5 compared with the actual evolution directly predicted using the Liouville equation. However, as expected, prediction of the behavior of the variances of the x and z components is as unsuccessful as for Model I. The performance of x (solid line) and y (long-dashed line) for the mean (Fig. 9a) remains the same as for Model I. The failure to predict $\overline{y}$ correctly, in spite of a modification of the distribution in the y direction, is attributed to the fact that the prediction of $\overline{y}_+$ (solid line)

decays to zero fairly rapidly after a relatively successful initial prediction until $t \simeq 1$ (Fig. 9c). On the other hand, the prediction of $\overline{y}_-$ (long-dashed line) is reasonable. Here, we recall the asymmetry in the initial condition.

When the run is initiated from the origin, the model behavior is even worse: both $\overline{y}_+$ and $\overline{y}_-$ monotonously decay towards zero with a timescale of $t \simeq 2$ (not shown). Thus, the initial condition symmetric to $y = 0$ somehow worsens the model behavior, rather than improving it; this is presumably because the first term in the right-hand sides of Eq. (62c) and (62d) dominate throughout the experiment.

5.3 The Lorenz (1963) system (Model III)

As a further extension of Model II in the last subsection, we now also constrain the system by $\overline{x}_\pm$. In this case, p_1 takes the following form:

$$p_1 = \begin{cases} p_{10} \exp(-\lambda_1^+ x), & x > 0 \\ p_{10} \exp(\lambda_1^- x), & x < 0. \end{cases} \tag{63}$$

Derivation of the equations for the assumed-PDF coefficients also proceeds in a similar manner to that outlined in the last subsection; the only change is that the $\lambda_1^\pm$ terms are now predicated by

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\lambda_1^+}\right) = -\frac{P}{\lambda_1^+} + \frac{P}{2}\overline{y}, \tag{64a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\lambda_1^-}\right) = -\frac{P}{\lambda_1^-} - \frac{P}{2}\overline{y}. \tag{64b}$$

The major improvement with this modification is better performance with respect to the x variance (solid line in Fig. 10b): it no longer decays out rapidly; rather, it remains about one-half of the actual value. On the other hand, the performance of the z variance (short-dashed line in Fig. 10b) does not change. As a rather intriguing modification, both $\overline{x}$ (solid line) and $\overline{y}$ (long-dashed line) somehow settled to a positive unstable fixed point in this case (Fig. 10a).

Even after the attempts with these three assumed PDFs, a successful representation of the statistics associated with a split of the distribution into two peaks is still a key remaining challenge. However, as it currently stands, it is not clear which parameter should be added to satisfy this challenge.

6 Summary and discussions

A general methodology for solving the time evolution of assumed-PDF parameters based on the Liouville equation has been proposed by Yano et al. (2024), with this prior study referred to as YLP in this work. The present paper extends this study in several aspects. First, it has been generalized (Sect. 2.2) for cases in which the constraints are defined by limited integral ranges. This generalization is a critical first
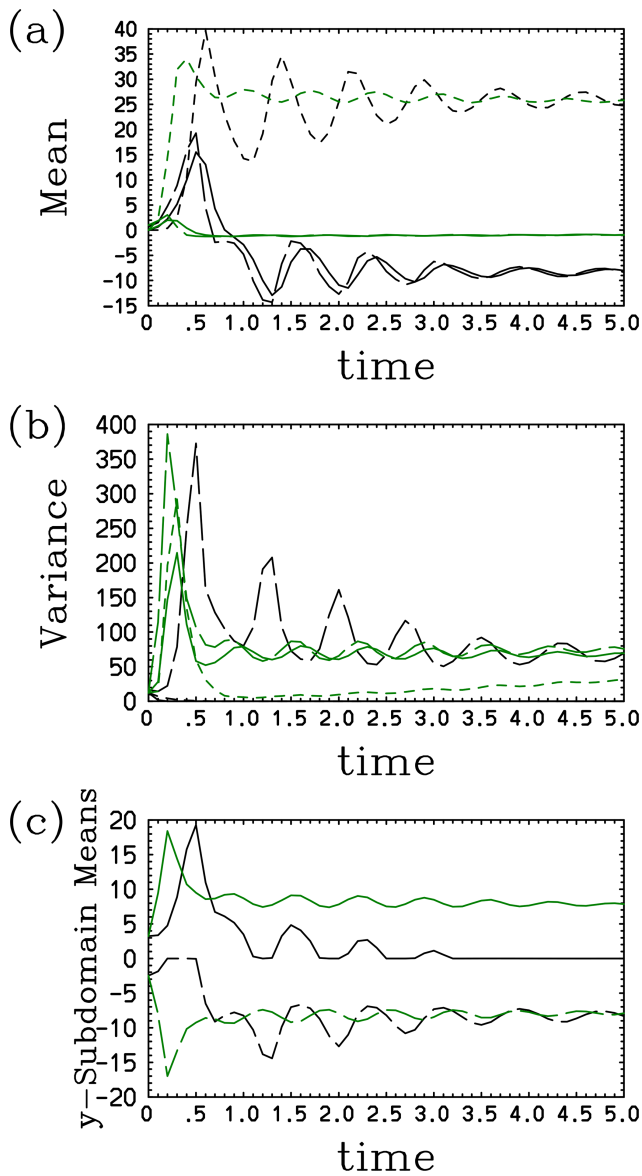
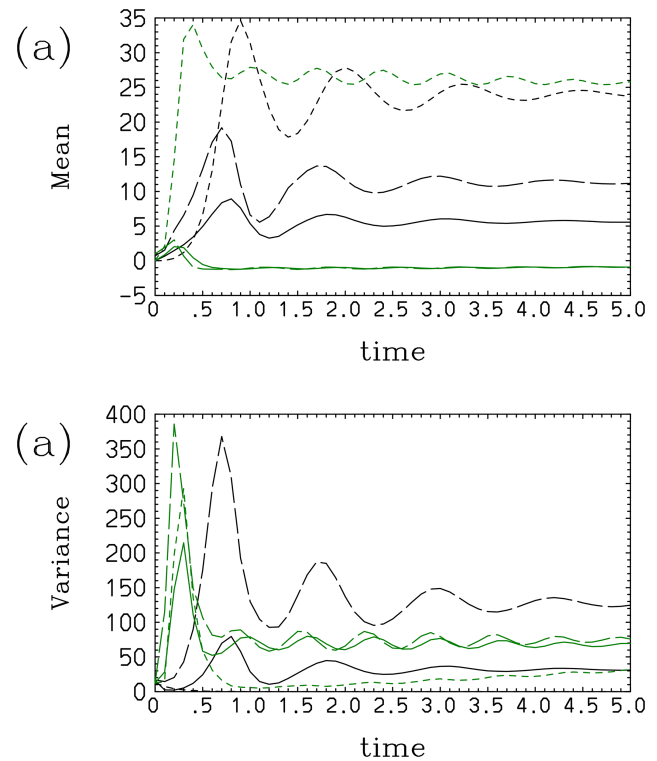**Figure 9.** The (a) mean, (b) variance, and (c) $\bar{y}_+$ (solid line) and $\bar{y}_-$ (long-dashed line) of the Lorenz system with Model II: the model results (black) and results obtained by direct prediction with the Liouville equation (green). In panels (a) and (b), the solid, long-dashed, and short-dashed curves denote the respective x, y, and z components.

Figure 10. The (a) mean and (b) variance of the Lorenz system with Model III: the model results (black) and results obtained by direct prediction with the Liouville equation (green).

step, for example, in developing a cloud scheme based on this formulation. As a result, the assumed-PDF forms also take different forms over different subdomains; thus, hereafter, the formulation for the prognostic equations for the PDF parameters has also been generalized accordingly (Sect. 2.3). Finally, the formulation has been explicitly generalized into multidimensional cases (Sect. 2.4 and 2.5). These further generalized formulations have been applied to two simple dynamical systems (Sects. 4 and 5, respectively): a convec-

tive energy-cycle system proposed by Yano and Plant (2012) and the Lorenz (1963) strange-attractor system.

Here, it may be worthwhile reiterating the originality of the present study, as extensive work already exists within the framework of the so-called assumed-PDF approach. In fact, as emphasized in YLP, more or less all of the existing formulations for the distribution problems fall into this category. However, this work is original in that it attempts to predict the evolution of a distribution in a self-consistent manner and to verify the performance using simple dynamical systems. Such an effort does not exist in the literature, as the existing assumed-PDF schemes, both with respect to subgrid-scale representations and cloud microphysics, are developed case by case with ad hoc closures, without generality. Thus, it is simply not possible to perform verifications of those schemes using simple dynamical systems.

The performance of the assumed-PDF formulations has been directly compared with the results from the direct time integration of the Liouville equation for both systems. By adopting simple dynamical systems with up to three dependent variables, such direct time integration becomes feasible. Unfortunately, both testing cases tend to suffer from the same tendencies: regardless of a specific choice, statistics (means and variances) predicted by an assumed-PDF approach gradually noticeably deviate from the exact results predicted by

the Liouville equation. Furthermore, some statistics, for example, the sign-dependent conditional averages, $\overline{x}_\pm$ and $\overline{y}_\pm$, in the Lorenz system turn out to be rather difficult to predict properly: in general, only one of the sign-dependent mean pair, $(\overline{x}_\pm, \overline{y}_\pm)$, is predicted properly, and the other of the pair simply settles to a vanishing value. In this respect, the assumed-PDF approach does not provide a "magic recipe".

One may judge that the obtained results are not overly promising and, perhaps, even a failure. However, it should be emphasized that rather difficult cases are used as test cases: for both systems, the initial Gaussian distribution rapidly evolves into a qualitatively totally different form. One must also consider the more basic fact that the assumed PDF attempts something almost impossible: performing an accurate prediction of a distribution by only using a limited number of parameters. For this reason, one should consider the obtained results an important demonstration of the fundamental difficulties associated with assumed-PDF methods in general, not only with the particular approach adopted herein. Arguably, the performance is rather impressive considering the fact that the PDF forms assumed are also qualitatively very different from the actual distributions predicted by the Liouville equation. The real question still to be to be answered is how well the more standard assumed-PDF approaches perform with the same systems. In that manner, the advantage of the present assumed-PDF approach may be better established.

The study has also suggested that the output-constrained distribution principle, proposed by YLP, may not be sufficient to decide an assumed-PDF form: to ensure good performance of a prediction of the distribution statistics, assumed-PDF forms must be constrained by something more than just those required as outputs for host models. In the present study, we have assumed that those constraints should be averages and variances. The present exploratory study has suggested that it is also crucial to evaluate the mean evolution of actual forcing terms of a system accurately; thus, they must also be added as a constraint. In both of the models considered herein, the term xy is crucial in prediction; thus, this correlation term must also be properly predicted. It has been shown that adding a constraint of $\langle xy \rangle$ can improve the predictions by assumed-PDF forms, although not necessarily in a satisfactory manner.

Prediction of the PDF of the Lorenz system is inherently difficult due to the fact that the solution tends to be clustered around the two unstable fixed points. The assumed-PDF, with all of the cases considered so far, have always failed to predict one of those two tendencies under conditional averages, $\overline{x}_\pm$ and $\overline{y}_\pm$. A further possibility to pursue is to re-initialize a prediction at a middle point, in the spirit of data assimilation, and to examine whether this difficulty is overcome by this procedure.

Some numerical issues have also been revealed by the present study. In some cases, exponential parameters in a distribution can vary to extremes that cause overflow and underflow problems in computations. To avoid this issue, some exponential parameters have been predicted in their logarithmic form to ensure better numerical stability. However, these basic procedures have not always been able to resolve the stability issues: in the case of Model III for the Lorenz system (Sect. 5.3), a rather small nondimensional time step of 10^{-4} has been necessary to run a prediction for long enough, but it still ultimately crashes. The numerical stability of this case must more closely be investigated in its own right.

As a whole, the present study reveals the inherent difficulties associated with accurately predicting a distribution with a limited number of distribution parameters. In particular, we have faced the universal tendency of the variance of the distribution to decay with time, when it must increase, regardless of the choice of the assumed distribution form. This behavior is reminiscent of the variance collapse identified as a major problem in data assimilation when it is performed with an ensemble. In the latter case, it is often found that the probability weight assigned to each ensemble member collapses close to zero except for a single member close to the average (i.e., weight collapse; see van Leeuwen, 2003, and Poterjoy, 2016); thus, the prior estimate of the variance also collapses as a consequence (cf. Snyder et al., 2008; the reader is referred to van Leeuwen, 2009, for more background). Although the cause of this problem is usually attributed to a relatively small ensemble size and the tendency of the ensemble to collapse into a "stable" state, the present study suggests that this tendency is more universal, and it can happen whenever a highly truncated representation of a distribution is adopted, even with an approach considering a full dimension of the phase space, as in the present case.

These ensemble-based assimilation studies, in turn, have attempted various remedies to alleviate this tendency: the simplest is to inflate the variance by a certain factor with time to prevent its collapse (Anderson and Anderson, 1999; Anderson, 2001). More generally, resampling approaches can, at least, partially delay the collapse of the weights (Snyder et al., 2008; Anderson, 2001). However, it appears that neither of the approaches is directly applicable to the assumed-PDF formulation in the present study. First, the inflation is nothing other than adding an extra adjustable parameter, which can only be chosen when an exact PDF evolution is known a priori. Resampling approaches do not work by simply not adopting an ensemble formulation.

The most feasible solution to solve the variance collapse under the assume-PDF approaches would be to include a feedback of the truncated distribution parameters in the form of a parameterization, in a similar manner to that with which the effects of higher moments are represented by certain hypotheses in the turbulence-closure models (Mellor, 1973; Mellor and Yamada, 1974). However, parameterization of the higher-order assumed-PDF parameters is a completely new frontier, and much investment is required before we can propose any specific solution.

Appendix A: Integrals of the two-dimensional Gaussian distribution

A1 Normalization

The normalization condition for a distribution with two variables, (x, y), is given by

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p\, \mathrm{d}x\mathrm{d}y = 1.$$

Recall the Gaussian integral formula

$$\int_{-\infty}^{+\infty} e^{-\lambda\xi^2}\mathrm{d}\xi = \left(\frac{\pi}{\lambda}\right)^{1/2}.$$

For casting the integral in x in this form, we rearrange a part of the exponent as follows:

$$\lambda_1(x-\overline{x})^2 + \lambda_3(x-\overline{x})(y-\overline{y})$$
$$= \lambda_1\left[x - \left\{\overline{x} - \frac{\lambda_3}{2\lambda_1}(y-\overline{y})\right\}\right]^2 - \lambda_1\left[\overline{x} - \frac{\lambda_3}{2\lambda_1}(y-\overline{y})\right]^2$$
$$+ \lambda_1\overline{x}^2 - \lambda_3(y-\overline{y})\overline{x}.$$

Thus,

$$\int_{-\infty}^{+\infty} \frac{p}{p_0}\mathrm{d}x = \left(\frac{\pi}{\lambda_1}\right)^{1/2}\exp\left[-\lambda_2(y-\overline{y})^2\right.$$
$$+\lambda_1\left\{\overline{x} - \frac{\lambda_3}{2\lambda_1}(y-\overline{y})\right\}^2 - \lambda_1\overline{x}^2$$
$$\left.+\lambda_3(y-\overline{y})\overline{x}\right]. \tag{A1}$$

The remaining exponent can be rearranged as

$$\lambda_2(y-\overline{y})^2 - \lambda_1\left[\overline{x} - \frac{\lambda_3}{2\lambda_1}(y-\overline{y})\right]^2 + \lambda_1\overline{x}^2$$
$$- \lambda_3(y-\overline{y})\overline{x} = \left(\lambda_2 + \frac{\lambda_3^2}{4\lambda_1}\right)(y-\overline{y})^2.$$

Thus, a further integral of Eq. (A1) in y leads to

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \frac{p}{p_0}\mathrm{d}x\mathrm{d}y = \left(\frac{\pi}{\lambda_1}\right)^{1/2}\left(\frac{\pi}{\lambda_2+\lambda_3^2/4\lambda_1}\right)^{1/2}$$
$$= \frac{\pi}{(\lambda_1\lambda_2 + \lambda_3^2/4)^{1/2}}.$$

The final expression proves Eq. (31a), (31b), and (31c).

A2 Moments

The moments given by Eq. (25a)–(25j) are derived using relationships obtained by taking the differentiation of the Gaussian distribution in a recursive manner:

$$\frac{\partial}{\partial x}\left[e^{-\{\lambda_1(x-\overline{x})^2+\lambda_3(x-\overline{x})(y-\overline{y})\}}\right]$$
$$= [2\lambda_1(x-\overline{x})+\lambda_3(y-\overline{y})]e^{-[\lambda_1(x-\overline{x})^2+\lambda_3(x-\overline{x})(y-\overline{y})]}.$$

The integral of the right-hand side leads to

$$\int_{-\infty}^{+\infty} [2\lambda_1(x-\overline{x})+\lambda_3(y-\overline{y})]e^{-[\lambda_1(x-\overline{x})^2+\lambda_3(x-\overline{x})(y-\overline{y})]}\mathrm{d}x$$
$$= -e^{-[\lambda_1(x-\overline{x})^2+\lambda_3(x-\overline{x})(y-\overline{y})]}|_{-\infty}^{+\infty} = 0. \tag{A2}$$

This relation immediately finds the following:

$$2\lambda_1\langle x-\overline{x}\rangle + \lambda_3\langle y-\overline{y}\rangle = 0. \tag{A3a}$$

Likewise, by symmetry

$$2\lambda_2\langle y-\overline{y}\rangle + \lambda_3\langle x-\overline{x}\rangle = 0. \tag{A3b}$$

Solving these expressions together leads to Eq. (25a), provided $4\lambda_1\lambda_2 - \lambda_3^2 \neq 0$.

To generalize a type of relation like Eq. (A3a) and (A3b), we note that Eq. (A2) is still valid by multiplying a weight, σ, by any function of y. Letting $\sigma = y - \overline{y}$,

$$2\lambda_1\langle(x-\overline{x})(y-\overline{y})\rangle + \lambda_3\langle(y-\overline{y})^2\rangle = 0. \tag{A4a}$$

By symmetry,

$$2\lambda_2\langle(x-\overline{x})(y-\overline{y})\rangle + \lambda_3\langle(x-\overline{x})^2\rangle = 0. \tag{A4b}$$

By combining these expressions together, we obtain the following relations:

$$\langle(x-\overline{x})(y-\overline{y})\rangle = -\frac{\lambda_3}{2\lambda_2}\langle(x-\overline{x})^2\rangle = -\frac{\lambda_3}{2\lambda_1}\langle(y-\overline{y})^2\rangle, \tag{A5a}$$

$$\langle(x-\overline{x})^2\rangle = -\frac{2\lambda_2}{\lambda_3}\langle(x-\overline{x})(y-\overline{y})\rangle, \tag{A5b}$$

$$\langle(y-\overline{y})^2\rangle = -\frac{2\lambda_1}{\lambda_3}\langle(x-\overline{x})(y-\overline{y})\rangle. \tag{A5c}$$

Equation (A2) can further be used to derive similar expressions for higher-moment integrals:

$$\int_{-\infty}^{+\infty} [2\lambda_1(x-\overline{x})+\lambda_3(y-\overline{y})]^n e^{-[\lambda_1(x-\overline{x})^2+\lambda_3(x-\overline{x})(y-\overline{y})]}\mathrm{d}x, \tag{A6}$$

where n is an arbitrary integral. We first set $n = 2$ and then obtain the following by a partial integral:

$$\int_{-\infty}^{+\infty} [2\lambda_1(x-\overline{x})+\lambda_3(y-\overline{y})]^2 e^{-[\lambda_1(x-\overline{x})^2+\lambda_3(x-\overline{x})(y-\overline{y})]}\mathrm{d}x$$

$$= -[2\lambda_1(x-\overline{x}) + \lambda_3(y-\overline{y})]e^{-[\lambda_1(x-\overline{x})^2 + \lambda_3(x-\overline{x})(y-\overline{y})]}\Big|_{-\infty}^{+\infty}$$
$$+ 2\lambda_1\int_{-\infty}^{+\infty} e^{-[\lambda_1(x-\overline{x})^2 + \lambda_3(x-\overline{x})(y-\overline{y})]}dx.$$

Thus,

$$\langle[2\lambda_1(x-\overline{x}) + \lambda_3(y-\overline{y})]^2\rangle$$
$$= -\int_{-\infty}^{+\infty}[2\lambda_1(x-\overline{x}) + \lambda_3(y-\overline{y})]p\big|_{-\infty}^{+\infty}dy + 2\lambda_1.$$

As the partial integral vanishes, by expanding the left-hand side, we obtain

$$4\lambda_1^2\langle(x-\overline{x})^2\rangle + \lambda_3^2\langle(y-\overline{y})^2\rangle$$
$$+ 4\lambda_1\lambda_3\langle(x-\overline{x})(y-\overline{y})\rangle = 2\lambda_1. \tag{A7a}$$

By symmetry, we also obtain

$$4\lambda_2^2\langle(y-\overline{y})^2\rangle + \lambda_3^2\langle(x-\overline{x})^2\rangle$$
$$+ 4\lambda_2\lambda_3\langle(x-\overline{x})(y-\overline{y})\rangle = 2\lambda_2. \tag{A7b}$$

By substituting Eq. (A5b) and (A5c) into the above, we obtain Eq. (25d). Its substitution back to Eq. (A5b) and (A5c) leads to Eq. (25b) and (25c), respectively.

To obtain the expressions for the third moments, we now set $n = 3$ in Eq. (A6). As this whole integral vanishes, by expanding this integral, we obtain the following:

$$8\lambda_1^3\langle(x-\overline{x})^3\rangle + \lambda_3^3\langle(y-\overline{y})^3\rangle + 12\lambda_1^2\lambda_3\langle(x-\overline{x})^2(y-\overline{y})\rangle$$
$$+ 6\lambda_1\lambda_3^2\langle(x-\overline{x})(y-\overline{y})^2\rangle = 0. \tag{A8a}$$

By symmetry,

$$8\lambda_2^3\langle(y-\overline{y})^3\rangle + \lambda_3^3\langle(x-\overline{x})^3\rangle + 12\lambda_2^2\lambda_3\langle(x-\overline{x})(y-\overline{y})^2\rangle$$
$$+ 6\lambda_2\lambda_3^2\langle(x-\overline{x})^2(y-\overline{y})\rangle = 0. \tag{A8b}$$

As an extension of Eq. (A3a) and (A3b), we obtain the following:

$$2\lambda_1\langle(x-\overline{x})(y-\overline{y})^2\rangle + \lambda_3\langle(y-\overline{y})^3\rangle = 0,$$
$$2\lambda_2\langle(x-\overline{x})^2(y-\overline{y})\rangle + \lambda_3\langle(x-\overline{x})^3\rangle = 0.$$

They lead to the following:

$$\langle(x-\overline{x})^3\rangle = -\frac{2\lambda_2}{\lambda_3}\langle(x-\overline{x})^2(y-\overline{y})\rangle, \tag{A9a}$$

$$\langle(y-\overline{y})^3\rangle = -\frac{2\lambda_1}{\lambda_3}\langle(x-\overline{x})(y-\overline{y})^2\rangle. \tag{A9b}$$

By substituting Eq. (A9a) and (A9b) into Eq. (A8a) and (A8b), we obtain

$$\lambda_1(4\lambda_1\lambda_2 - 3\lambda_3^2)\langle(x-\overline{x})^2(y-\overline{y})\rangle - \lambda_3^3\langle(x-\overline{x})(y-\overline{y})^2\rangle = 0,$$

$$\lambda_2(4\lambda_1\lambda_2 - 3\lambda_3^2)\langle(x-\overline{x})(y-\overline{y})^2\rangle - \lambda_3^3\langle(x-\overline{x})^2(y-\overline{y})\rangle = 0.$$

By solving them for $\langle(x-\overline{x})^2(y-\overline{y})\rangle$ and $\langle(x-\overline{x})(y-\overline{y})^2\rangle$ and also substituting this result into Eq. (A9a) and (A9b), we obtain Eq. (25e).

Finally, to obtain the results for the fourth moments, we set $n = 4$ in Eq. (A6), which leads to the following:

$$16\lambda_1^4\langle(x-\overline{x})^4\rangle + \lambda_3^4\langle(y-\overline{y})^4\rangle$$
$$+ 32\lambda_1^3\lambda_3\langle(x-\overline{x})^3(y-\overline{y})\rangle$$
$$+ 8\lambda_1\lambda_3^3\langle(x-\overline{x})(y-\overline{y})^3\rangle$$
$$+ 24\lambda_1^2\lambda_3^2\langle(x-\overline{x})^2(y-\overline{y})^2\rangle = 12\lambda_1^2. \tag{A10a}$$

By symmetry,

$$16\lambda_2^4\langle(y-\overline{y})^4\rangle + \lambda_3^4\langle(x-\overline{x})^4\rangle$$
$$+ 32\lambda_2^3\lambda_3\langle(x-\overline{x})(y-\overline{y})^3\rangle$$
$$+ 8\lambda_2\lambda_3^3\langle(x-\overline{x})^3(y-\overline{y})\rangle$$
$$+ 24\lambda_2^2\lambda_3^2\langle(x-\overline{x})^2(y-\overline{y})^2\rangle = 12\lambda_2^2. \tag{A10b}$$

As an extension of Eq. (A3a) and (A3b), we obtain

$$2\lambda_1\langle(x-\overline{x})(y-\overline{y})^3\rangle + \lambda_3\langle(y-\overline{y})^4\rangle = 0,$$
$$2\lambda_2\langle(x-\overline{x})^3(y-\overline{y})\rangle + \lambda_3\langle(x-\overline{x})^4\rangle = 0,$$

which lead to

$$\langle(x-\overline{x})(y-\overline{y})^3\rangle = -\frac{\lambda_3}{2\lambda_1}\langle(y-\overline{y})^4\rangle, \tag{A11a}$$

$$\langle(x-\overline{x})^3(y-\overline{y})\rangle = -\frac{\lambda_3}{2\lambda_2}\langle(x-\overline{x})^4\rangle. \tag{A11b}$$

The results of Eq. (A7a) and (A7b) are extended by applying weights of $(y-\overline{y})^2$ and $(x-\overline{x})^2$, respectively, and we obtain the following:

$$4\lambda_1^2\langle(x-\overline{x})^2(y-\overline{y})^2\rangle + \lambda_3^2\langle(y-\overline{y})^4\rangle$$
$$+ 4\lambda_1\lambda_3\langle(x-\overline{x})(y-\overline{y})^3\rangle = \frac{\lambda_1}{\lambda_2}\left(1-\frac{\kappa}{4}\right)^{-1},$$

$$4\lambda_2^2\langle(x-\overline{x})^2(y-\overline{y})^2\rangle + \lambda_3^2\langle(x-\overline{x})^4\rangle$$
$$+ 4\lambda_2\lambda_3\langle(x-\overline{x})^3(y-\overline{y})\rangle = \frac{\lambda_2}{\lambda_1}\left(1-\frac{\kappa}{4}\right)^{-1}.$$

By substituting Eq. (A11a) and (A11b) into the above, we obtain the following:

$$\langle(x-\overline{x})^4\rangle = \frac{4\lambda_2^2}{\lambda_3^2}\langle(x-\overline{x})^2(y-\overline{y})^2\rangle - \frac{\lambda_2}{\lambda_1\lambda_3^2}\left(1-\frac{\kappa}{4}\right)^{-1}, \tag{A12a}$$

$$\langle(y-\overline{y})^4\rangle = \frac{4\lambda_1^2}{\lambda_3^2}\langle(x-\overline{x})^2(y-\overline{y})^2\rangle - \frac{\lambda_1}{\lambda_2\lambda_3^2}\left(1-\frac{\kappa}{4}\right)^{-1}. \tag{A12b}$$

We first substitute Eq. (A11a) and (A11b) into Eq. (A10a) and (A10b) and then substitute Eq. (A12a) and (A12b) into the latter. This gives Eq. (25h). Substituting Eq. (25h) back into Eq. (A12a) and (A12b) gives Eq. (25f) and (25g), and further substitutions of them into Eq. (A11a) and (A11b) lead to Eq. (25i) and (25j).

References

Anderson, J. L.: An ensemble adjustment Kalman filter for data assimilation, Mon. Weather Rev., 129, 2084–2903, 2001.

Anderson, J. L. and Anderson, S. L.: A Monte-Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts, Mon. Weather Rev., 127, 2741–2758, 1999.

Bechtold, P., Fravalo, C., and Pinty, J. P.: A model of marine boundary-layer cloudness for mesoscale applications, J. Atmos. Sci., 49, 1723–1744, 1992.

Bechtold, P., Cuijpers, J. W. M., Mascart, P., and Trouilhet, P.: Modeling of trade-wind cumuli with a low-order turbulence model – Towards a unified description of Cu and Sc clouds in meteorological models, J. Atmos. Sci., 52, 455–463, 1995.

Bony, S. and Emanuel, K. A.: A parameterization of the cloudness associated with cumulus convection: Evaluation using TOGA COARE data, J. Atmos. Sci., 58, 3158–3183, 2001.

Bougeault, P.: Modeling the trade-wind cumulus boundary-layer. Part I: Testing the ensemble cloud relations against numerical data, J. Atmos. Sci., 38, 2414–2428, 1981.

Carrassi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, Climatic Change, 9, e535, https://doi.org/10.1002/wcc.535, 2018.

Evensen, G., Vossepoel, F. C., and van Leeuwen, P. J.: Data Assimilation Fundamentals, Springer, Berlin, 246 pp., ISBN-10 3030967085, 2022.

Golaz, J.-C., Larson, V. E., and Cotton, W. R.: A PDF-based model for boundary layer clouds. Part I: Method and model description, J. Atmos. Sci., 59, 3540–3551, 2002.

Khain, A. P. and Pinsky, M.: Physical Processes in Clouds and Cloud Modeling, Cambridge University Press, Cambridge, 626 pp., ISBN 0521767431, 9780521767439, 2018.

Khain, A. P., Beheng, K. D., Heymsfield, A., Korolev, A., Krichak, S. O., Levin, Z., Pinsky, M., Phillips, V., Prabhakaran, T., Teller, A., van den Heever, S. C., and Yano, J. I.: Representation of microphysical processes in cloud-resolving models: spectral (bin) microphysics vs. bulk-microphysics, Rev. Geophys., 53, 247–322, https://doi.org/10.1002/2014RG000468, 2015.

Larson, V. E. and Golaz, J.: Using probability density functions to derive consistent closure relationships among higher-order Moments, Mon. Weather Rev., 133, 1023–1042, https://doi.org/10.1175/MWR2902.1, 2005.

Le Treut, H. and Li, Z. X.: Sensitivity of an atmospheric general circulation model to prescribed SST changes: Feedback effects associated with the simulation of cloud optical properties, Clim. Dynam., 5, 175–187, 1991.

Lorenz, Z. N.: Deterministic nonperiodic flow, J. Atmos. Sci., 20, 130–141, 1963.

Mellor, G. L.: Analytic prediction of the properties of stratified planetary surface layers, J. Atmos. Sci., 30, 1061–1069, 1973.

Mellor, G. L.: Gaussian cloud model relations. J. Atmos. Sci., 34, 356–358, 1977.

Mellor, G. L. and Yamada, T.: A hierarchy of turbulence closure models for planetary boundary layers, J. Atmos. Sci., 31, 1791–1806, 1974.

Poterjoy, J.: A localized particle filter for high-dimensional nonlinear systems, Mon. Weather Rev., 144, 59–76, 2016.

Richard, J. L. and Royer, J. F.: A statistical cloud scheme for ruse in an AGCM, Ann. Geophys., 11, 1095–1115, 1993.

Snyder, C., Bengtsson, T., Bickel, P., and Anderson, J.: Obstacles high-dimensional particle filtering, Mon. Weather Rev., 136, 4629–4640, 2006.

Sommeria, G. and Deadorff, J. W.: Subgrid-scale condensation in models of nonprecipitating clouds, J. Atmos. Sci., 34, 344–355, 1977.

Tompkins, A. M.: A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover, J. Atmos. Sci., 59, 1917–1942, 2002.

van Leeuwen, P. J.: A variance -minimizing filter for large-scale applications, Mon. Weather Rev., 131, 2071–2084, 2003.

van Leeuwen, P. J.: Particle filtering in geophysical systems. J. Climate, 137, 4089–4114, 2009.

Yano, J.-I. and Ouchtar, E.: Convective Initiation Uncertainties Without Trigger or Stochasticity: Probabilistic Description by the Liouville Equation and Bayes' Theorem, Q. J. Roy. Meteor. Soc., 143, 2015–2035, https://doi.org/10.1002/qj.3064, 2017.

Yano, J.-I. and Phillips, V. T. J.: Explosive ice multiplication induced by multiplicative-noise fluctuation of mechanical break-up in ice-ice collisions, J. Atmos Sci., 73, 4685–4697, 2016.

Yano, J.-I. and Plant, R. S.: Finite Departure from Convective Quasi-Equilibrium: Periodic Cycle and Discharge-Recharge Mechanism, Q. J. Roy. Meteor. Soc., 138, 626–637, 2012, https://doi.org/10.1002/qj.957.

Yano, J.-I., Ziemiański, M., Cullen, M., Termonia, P., Onvlee, J., Bengtsson, L., Carrassi, A., Davy, R., Deluca, A., Gray, S. L., Homar, V., Köhler, M., Krichak, S., Michaelides, S., Phillips, V. T. J., Soares, P. M. M., and Wyszogrodzki, A.: Scientific Chal-

lenges of Convective-Scale Numerical Weather Prediction, B. Am. Meteorol. Soc., 99, 699–710, https://doi.org/10.1175/bams-d-17-0125.1, 2018.

Yano, J.-I., Larson, V., and Phillips, V. T. J.: Technical Note: General Formulation For the Distribution Problem: Prognostic Assumed PDF Approach Based on The Maximum–Entropy Principle and The Liouville Equation, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2023-2278, 2024.