地球物理系统中的信息流动

P.J. van Leeuwen

美国科罗拉多州柯林斯堡，科罗拉多州立大学大气科学系

要点：

- 本文推导出一种适用于高维地球物理系统的信息流新框架，只需进行低维计算。
- 信息演化与能量或熵等量的演化截然不同，并且因变量而异。
- 将其应用于高度非线性系统，有助于加深我们对非线性项相互作用及可预测性的理解。

arXiv:2601.01795v1 [cs.IT] 5 Jan 2026

通讯作者：P.J. van Leeuwen，peter.vanleeuwen@colostate.edu

摘要

我们提出了一种用于分析地球物理系统中信息演化的新框架。理解信息及其对应面——不确定性——如何传播，是可预报性研究的核心，对预报不确定性量化和风险管理等应用也有重要意义。此外，这一框架还能帮助我们深入了解系统的底层物理机制。信息传播与因果性密切相关：系统的某一部分如何影响另一部分，以及某些区域如何保持动力学上的隔离。我们将这一框架应用于一维、高度非线性的 Kuramoto–Sivashinsky 模型，以及代表中纬度大气带的浅水方程。值得注意的是，我们观察到信息可以逆着流体流动传播，而且不同模型变量的信息演化模式各不相同。例如，与压力相关的信息和相对涡度的传播方式不同，这反映了重力波与平衡流动力学的影响差异。这一新框架有望成为研究复杂动力系统的一种有价值的诊断工具。

通俗语言总结

我们引入一种新的诊断量——信息，以深化我们对地球物理系统的理解。信息被定义为在给定位置上关于特定变量的可用知识量。我们建立了一个通用框架，用于计算和分析其时间演化，重点关注信息如何在系统中传播。这一视角不仅为可预测性提供了宝贵的洞见，也有助于理解支配系统行为的潜在物理过程。我们通过一个高度非线性的玩具模型和一个简化的大气模型展示了该框架，说明它如何与既有的物理认识相联系，同时揭示新的动力学。

1 引言

尽管人们常用能量、动量或涡度等物理量的演变来理解地球物理系统的动力学，但不确定性——或者反过来说，信息——的演变却远未受到同等程度的关注。以往的研究大多聚焦于可预报性研究（Carnevale & Holloway, 1982; Shukla, 1998; Smith 等, 1999）和数据同化（Van Leeuwen & Evensen, 1996;

Evensen等, 2022），其中研究了初始条件、模型参数或边界条件所导致的不确定性增长，以确定误差加倍时间和预测时域等量。然而，研究信息如何在系统中传播，还能揭示更多内容。信息流可能反映已知的物理过程，甚至揭示此前未发现的机制，其中一些机制可能难以通过传统的预报变量捕捉。

信息可以通过多种方式定义。直观而言，它的反面——不确定性——通常与变量概率密度函数（pdf）的“宽度”相关联，但即使是这种宽度，也可以用不同方式量化。一种应用广泛且通用的测度是香农微分熵：

$$H_\psi = -\int p(\psi)\log p(\psi);d\psi \tag{1}$$

这一定义不假设概率密度函数具有特定形状，适用于光滑分布。其推广形式包括 R´enyi 熵（Renyi，1961），它们能够灵活调节对尾部行为及其他分布特征的敏感度。

Kleeman（2002）首次将基于熵的测度引入可预报性理论，他使用相对熵（库尔贝克–莱布勒散度）来量化相对于气候态的预报技巧。在数据同化领域，Van Leeuwen（2003）运用熵来理解非线性数据同化对高度非线性系统的影响。他发现，与线性数据同化相反，在非线性数据同化中，后验方差可能更大于先验方差。然而，熵仍在减小。

《熵》也一直是因果发现领域的核心。变量集合 $\boldsymbol{\psi}_e$ 对目标变量 ψ 的因果影响通常定义为： ψ（t）中的变异部分，这部分变异可归因于较早时刻 $\boldsymbol{\psi}_e$ 的变化；这一概念可追溯至 Wiener（1956）和 Granger（1969）。尽管最初主要应用于滞后线性回归，但多年来已经发展出基于信息度量的非线性扩展，例如互信息和传递熵；参见 Runge等（2015，2019）的综述。展示所得因果网络的标准方式是使用所谓的有向无环图（DAGs），例如 Pearl（2009）所述的图，其中箭头从一个变量指向其产生影响的其他变量，并通过各种方式对这种影响进行量化，例如，

通过互信息。DAGs 本质上是无环的，因此在设计上排除了循环过程，降低了其在实际应用中的实用性。一旦我们将因果性重新定义为信息从驱动变量流向目标变量，便可以轻松地纳入循环过程，Liang, 2021 已在线性系统中对此进行了展示。此外，DAGs 排除了两个或更多变量对其他变量的联合影响，而这正是非线性系统中典型的因果交互。Van Leeuwen 等（2021）近期提出的因果框架通过纳入交互信息来捕捉变量之间的联合效应，从而解决了仅成对影响的局限性。尽管交互信息通常被认为难以解释，但该研究表明，在动力系统背景下，交互信息提供了一种严谨且具有物理意义的联合因果关系测度。

尽管因果发现主要聚焦于观测数据，但也可以应用于数值模型。虽然在单个时间步内，模型方程的因果结构是明确已知的，但非线性动力学可能会掩盖更长时间范围内的因果路径。一种方法是固定驱动变量，并观察目标变量由此产生的变化。然而，这类干预往往存在问题。例如，试图通过人为固定气压来评估大气压力对风速的影响，会扰乱其他反馈，从而使实验失效。因此，将传统的基于干预的方法应用于复杂的耦合系统中的预报变量时，往往会失败。

一项重大进展来自将动力系统方程与相应概率密度函数（pdf）的演化联系起来，这通过 Fokker–Planck（或 Kolmogorov）方程实现（Fokker, 1914；Planck, 1917；Kolmogorov, 1931）。Liang 和 Kleeman（2005）利用这一形式推导出了变量之间信息传递的表达式，随后将其扩展到多变量系统（Liang 和 Kleeman，2007a，2007b），并基于底层动力学进行了严格推导，同时在（Liang, 2016）中明确建立了与因果性的联系，并在（Liang, 2021）中进一步将其推广到线性系统。Pires等（2024）进一步完善了该理论。这些研究为连续系统和离散系统中的熵流提供了重要洞见。

尽管这些论文取得了长足进展，但该形式体系仍存在一些问题。香农熵在非线性变换下并不具有不变性，这使得信息流取决于所使用的变量。例如，使用云液态水含量或其对数时，将会得到不同的结果

。此外，连续变量的公式——本文关注的正是这类变量——必须处理这样一个问题：熵可能为负，具体取决于概率密度函数的宽度。例如，标准差为 σ 的高斯概率密度函数的香农熵等于 $1/2 \log(2\pi e\sigma^2)$，当 $\sigma^2 < 1/2\pi e$ 时为负值。当 σ 越过这一阈值时，这将导致熵流动方向突然改变。Liang（2018）发现了这些问题，并以相对熵的形式重新制定了 Liang, 2016 的方法论。

Liang–Kleeman 形式体系以及（Liang, 2016, 2018）形式体系的另一个局限在于计算方面：它们需要对整个状态空间进行积分，因此对于包含数百万个变量的高维模型而言，计算将变得不可行。Pires等（2024）提出通过降低系统维度，并在降维后的系统上进行信息传递计算；但这必然会限制可研究现象的范围，而且信息流会如何受到降维影响，并非显而易见。

最后，（Liang（2016））及相关形式体系都基于冻结潜在驱动变量来推断因果关系。如上所述，这可能导致实际计算问题，因为固定某个变量可能会造成不符合物理规律甚至不稳定的模拟。线性因果关系不存在这一问题，因为在这种情况下存在显式解，Liang（2016）中的线性示例已对此进行了说明。

~~最后，正如第 2 节将要说明的，自信息与从其他变量传递而来的信息之间的划分，~~基于 Liang 和 Kleeman（2005）提出的一种启发式论证，缺乏严谨的理论基础。这种划分带有任意性，可能影响对结果的解读。

本文提出了一种新的信息和信息流形式体系，为所有这些问题提供了解决方案。该体系以相对熵为基础，遵循 Kleeman（2002）和 Liang（2018）的研究，而非微分熵；即使在超高维系统中，所有积分仍保持低维，从而揭示物理联系的局部性质。此外，该框架表明，确定信息流乃至因果性并不需要冻结变量。这一推导建立在以往基于熵的方法之上，同时进行了关键改进，使该框架更加

稳健、通用且易于解释。该框架适用于现实中的高维系统，无需对变量变换或概率密度函数形状作出假设。

我们通过两个案例研究展示该框架的实用性：作为典型非线性系统的 Kuramoto–Sivashinsky方程，以及代表中纬度大气动力学的浅水模型。在这两个案例中，该框架都揭示了信息如何在不同变量和尺度间传播与相互作用的新见解。我们还说明了如何利用规模相对较小的系综（~ 100 个成员）可靠地估算所需的积分，从而确保该框架适用于业务和研究应用。

本文结构如下。第 2 节将阐述该框架，并通过一个简单示例说明如何根据系统的底层物理机制分析信息演化。随后，第 3 节介绍新框架在 Kuramoto–Sivashinsky 模型（Kuramoto, 1978; Sivashinsky, 1977, 1980）中的应用，第 4 节则介绍其在大气浅水模型中的应用。第 5 节总结全文并讨论研究结果。三个附录分别详述若干证明、将形式体系扩展至偏微分方程所描述的无限维空间的方法，以及计算所需积分的实用方法。

2 方法论

我们从状态向量 $\boldsymbol{\phi} \in \Re^n$ 所对应系统的演化方程出发，如下：

$$\frac{\partial \boldsymbol{\phi}}{\partial t} = \mathbf{F}(\boldsymbol{\phi}, t) \tag{2}$$

该方程通过数值模型进行模拟，因此，即使状态是无限维的，经过离散化后也会得到一个有限维状态向量。相应的 Fokker-Planck（或 Kolmogorov）方程如下：

$$\frac{\partial p(\boldsymbol{\phi}, t)}{\partial t} = -\frac{\partial}{\partial \boldsymbol{\phi}}\Big(\mathbf{F}(\boldsymbol{\phi}, t)p(\boldsymbol{\phi}, t)\Big) \tag{3}$$

并描述向量 $\boldsymbol{\phi}$的概率密度函数（概率密度函数）随时间和空间的演化。

2.1 信息

我们关注特定网格点处特定变量的信息随时间的演变，记作 $\psi(t)$。信息没有通用的定义，但它与

香农熵有明确的联系。香农熵定义如下：

$$H = -\sum_{i=1}^{N} p_i \log p_i, \tag{4}$$

对于离散变量，其中 p_i 是事件 i 发生的概率。对于地球科学中遇到的连续变量，可以定义微分熵

$$H_\psi = -\int p(\psi) \log p(\psi)\, d\psi, \tag{5}$$

其中 $p(\psi)$ 是连续变量 ψ的概率密度函数（pdf）。然而，正如引言中所述，这个量存在一个问题：对于较窄的概率密度函数，它可能为负。这意味着，将熵通常解释为不确定性的测度会变得棘手。以负微分熵的概率密度函数为简单例子，考虑一个高斯分布概率密度函数，其熵等于 $1/2 \log(2\pi e\sigma^2)$，其中 σ 是该概率密度函数的标准差。对于 $\sigma^2 < 1/(2\pi e)$，对数的真数小于一，因此熵为负。

避免这个问题的一种方法是使用相对熵

$$I_\psi = \int p(\psi) \log \frac{p(\psi)}{q(\psi} \, d\psi, \tag{6}$$

其中 $q(\psi)$ 是参考概率密度函数。这个量也称为 Kullback–Leibler 散度，用来衡量两个概率密度函数之间的差异。严格来说，它并非真正的距离度量，因为它的两个自变量不具对称性。不过，它始终非负，并且当且仅当两个概率密度函数相同时等于零，因为两个概率密度函数之比为一时，对数项就变为零。

要有意义地应用相对熵，必须指定一个参考密度。由于我们关注的是与概率密度函数 $p(\psi)$ 变化相关的信息演化，参考密度应保持时间不变。此外，随着我们的确定性增加，信息也应增加；也就是说，当 $p(\psi)$ 变得更加集中（更窄）时，信息应增加。这意味着，随着其与参考密度的偏离程度增大，相对熵也应增加。

这些考量促使我们采用一种自然且广受认可的参考密度选择：气候学概率密度函数，它由足够长的时间序列获得，从而使分布达到统计平稳。本研究采用这一气候学概率密度函数作为参考密度，以计算相对熵。

2.2 信息流

现在我们将推导信息的演化方程。为此，我们将推导系统中每个变量 ψ 的信息演化方程。第一步是对 $\boldsymbol{\phi}$ 中所有不包含 ψ的变量积分福克-普朗克方程。由此得到

$$\frac{\partial p(\psi,t)}{\partial t} = -\frac{\partial}{\partial \psi}\int p(\boldsymbol{\phi},t)F_\psi(\boldsymbol{\phi},t)\; d\boldsymbol{\phi}\backslash\psi - \int_{\partial\boldsymbol{\phi}\backslash\partial\psi} p(\boldsymbol{\phi},t)\;\mathbf{F}(\boldsymbol{\phi},t)\cdot\mathbf{n}\; dS. \tag{7}$$

其中 $\mathbf{F}\psi(\boldsymbol{\phi},t)$ 表示变量 ψ 的驱动项。第二项是边界项。尽管边界项在许多应用中都很重要，但下文中我们假设它为零，并将其详细处理留待后续论文。

现在我们利用这样一个物理事实：在一个时间步 ψ 内，它只取决于其局部邻居 $\boldsymbol{\psi}_e$。这意味着，对于 $\boldsymbol{\phi}$ 中远离 ψ 的那些变量（即不属于 $\boldsymbol{\psi}_e$ 的变量），$F_\psi(\boldsymbol{\phi},t) = 0$。若用 $\boldsymbol{\phi}_{ne}$ 表示这些变量，我们得到：

$$\begin{aligned}
\frac{\partial p(\psi,t)}{\partial t} &= -\frac{\partial}{\partial \psi}\int p(\boldsymbol{\phi},t)F_\psi(\boldsymbol{\phi},t)\; d\boldsymbol{\phi}\backslash\psi \\
&= -\frac{\partial}{\partial \psi}\int p(\boldsymbol{\phi}_{ne},t|\psi,\boldsymbol{\psi}_e)p(\psi,\boldsymbol{\psi}_e,t)F(\psi,\boldsymbol{\psi}_e,t)\; d\boldsymbol{\phi}_{ne}d\boldsymbol{\psi}_e, \\
&= -\frac{\partial}{\partial \psi}\int p(\psi,\boldsymbol{\psi}_e,t)F(\psi,\boldsymbol{\psi}_e,t)\; d\boldsymbol{\psi}_e
\end{aligned} \tag{8}$$

这里我们省略了 F 上的下标 ψ，因为从第一个自变量可以清楚看出 $F(\psi,\boldsymbol{\psi}_e,t)$ 的含义。

现在我们乘以 $[1+\log(p(\psi,t)/q(\psi))]$，并对 ψ 积分，得到左侧：

$$\begin{aligned}
&\int\left[1+\log\frac{p(\psi,t)}{q(\psi)}\right]\frac{\partial p(\psi,t)}{\partial t}\; d\psi = \\
&\qquad = \frac{\partial}{\partial t}\int\left[1+\log\frac{p(\psi,t)}{q(\psi)}\right]p(\psi,t)\; d\psi - \int p(\psi,t)\frac{\partial \log p(\psi,t)}{\partial t}\; d\psi = \\
&\qquad = \frac{\partial}{\partial t}\int p(\psi,t)\log\frac{p(\psi,t)}{q(\psi)}\; d\psi = \frac{\partial I_\psi}{\partial t},
\end{aligned} \tag{9}$$

这里我们将最后一项识别为变量 ψ的相对熵对时间的导数，并且用到了

$$\int p(\psi,t)\frac{\partial \log p(\psi,t)}{\partial t}\; d\psi = \int \frac{\partial p(\psi,t)}{\partial t}\; d\psi = \frac{\partial}{\partial t}\int p(\psi,t)\; d\psi = 0. \tag{10}$$

因为变量的概率密度函数积分后等于常数 1。

对于右端的第一项，我们得到：

$$-\int\left[1+\log\frac{p(\psi,t)}{q(\psi)}\right]\frac{\partial}{\partial\psi}\Big[F(\psi,\boldsymbol{\psi}_e,t)p(\psi,\boldsymbol{\psi}_e,t)\Big]\,d\boldsymbol{\psi}_e d\psi = \\ =\int F(\psi,\boldsymbol{\psi}_e,t)p(\psi,\boldsymbol{\psi}_e,t)\frac{\partial}{\partial\psi}\left(\log\frac{p(\psi,t)}{q(\psi)}\right)d\boldsymbol{\psi}_e d\psi, \tag{11}$$

其中我们再次假设边界项为零。

综合以上各项，我们得到：

$$\frac{\partial I_\psi}{\partial t}=\int F(\psi,\boldsymbol{\psi}_e,t)p(\psi,\boldsymbol{\psi}_e,t)\frac{\partial}{\partial\psi}\left(\log\frac{p(\psi,t)}{q(\psi)}\right)d\boldsymbol{\psi}_e d\psi, \tag{12}$$

其中，$\boldsymbol{\psi}_e$ 包含离散化系统中计算 $F(\psi,\boldsymbol{\psi}_e,t)$在一个时间步内所需的所有其他变量。

所提出推导的一个关键特点是，其中涉及的所有积分都是低维的。特别是，维数最高的积分来自用于计算变量 ψ演化方程右端的局部模板。对于复杂的数值平流格式，尤其是在非结构化三维网格上使用的格式，这个模板仍可能涉及相对大量的变量。不过，我们的方法有一个重要优势：它无需对积分进行时间迭代。相反，我们可以依靠底层数值模型，在每个时间步提供必要的状态变量。这样，我们便能使用简化的导数近似，而不影响精度。下一节将介绍一个实际示例，说明如何高效计算该框架中的各个项。

2.3 平流-扩散方程示例

如上所述，$F(\psi,\ \boldsymbol{\psi}_e,\ t)$ 表示概率密度函数在 ψ空间中的平流速度。它通常包含物理平流和扩散项，以及与源汇和相变等过程相关的局部项；这些局部项仅取决于同一模型网格点上的变量。

为简化记法，本节省略显式的时间依赖。对于连续系统，某个量 ρ 的平流扩散方程为：

$$\frac{\partial\rho}{\partial t}=-u\frac{\partial\rho}{\partial x}-v\frac{\partial\rho}{\partial y}+D\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)\rho+\lambda f(\rho,\boldsymbol{\psi}_e), \tag{13}$$

我们的量 ψ 已在空间上离散化，因此其空间导数不存在。有两种可行的方法。我们可以先将该方程离散化，然后采用

第 2.2 节中的框架；或者按照附录 B 的详细说明，对无限维空间中测度 $\mu_t(\rho)$ 演化所对应的刘维尔方程进行边缘化。如该附录所示，这种边缘化仍会在空间中的每个位置得到变量 ψ 的概率密度函数的局部描述。如果随后将所得方程离散化，以便进行实际评估，我们会发现，这两种方法得到的结果相似。这里，我们先将 ρ 的方程离散化，然后直接应用第 2.2 节中的框架，因为这对 JAMES 的读者来说会更加熟悉。

我们将沿 x 方向的空间差分算子记为 $L_x(\psi)$，二阶差分算子也类似。于是，对于 $F(\psi,\ \boldsymbol{\psi}_e)$，我们得到：

$$F(\psi,\boldsymbol{\psi}_e) = -uL_x(\psi) - vL_y\psi) + D\big(L_{xx}(\psi) + L_{yy}(\psi)\big) + \lambda f(\psi,\boldsymbol{\psi}_e), \tag{14}$$

其中，$u = u(\psi,\ \boldsymbol{\psi}_e)$、$v = v(\psi,\ \boldsymbol{\psi}_e)$ 和 $f(\psi,\ \boldsymbol{\psi}_e)$ 是仅依赖于局部变量 $\psi,\ \boldsymbol{\psi}_e$ 的函数。下面我们将逐项求值。

对于 u 平流，由方程（12）可得，略去显式的 t 依赖：

$$-\int p(\psi,\boldsymbol{\psi}_e)uL_x(\psi)\frac{\partial}{\partial\psi}\left(\log\frac{p(\psi)}{q(\psi)}\right)\,d\psi d\boldsymbol{\psi}_e. \tag{15}$$

为求出这一项，我们利用概率密度函数仅通过 ψ 依赖于 x 这一点，得到：

$$L_x\left(\log\frac{p(\psi)}{q(\psi)}\right) \approx L_x(\psi)\frac{\partial}{\partial\psi}\left(\log\frac{p(\psi)}{q(\psi)}\right), \tag{16}$$

误差为 Δx 阶或更小，具体取决于所使用的离散化格式，详见附录A。

于是我们得到：

$$\begin{aligned}
&-\int p(\psi,\boldsymbol{\psi}_e)\ u\ L_x(\psi)\frac{\partial}{\partial\psi}\left(\log\frac{p(\psi)}{q(\psi)}\right)\,d\psi d\boldsymbol{\psi}_e = \\
&\quad\approx -\int p(\psi,\boldsymbol{\psi}_e)\ u\ L_x\left(\log\frac{p(\psi)}{q(\psi)}\right)\,d\psi d\boldsymbol{\psi}_e = \\
&\quad\approx -L_x\left(\int p(\psi,\boldsymbol{\psi}_e)u\log\frac{p(\psi)}{q(\psi)}\,d\psi d\boldsymbol{\psi}_e\right) + \int p(\psi,\boldsymbol{\psi}_e)L_x(u)\log\frac{p(\psi)}{q(\psi)}\,d\psi d\boldsymbol{\psi}_e
\end{aligned} \tag{17}$$

与 $p(\psi,\boldsymbol{\psi}_e)$ 的空间导数相关的项不会出现，误差取决于离散化格式，详见附录 B。

同样，我们可以得到经向平流：

$$\begin{aligned}
&-\int p(\psi,\boldsymbol{\psi}_e)\ v\ L_y(\psi)\frac{\partial}{\partial\psi}\left(\log\frac{p(\psi)}{q(\psi)}\right)\,d\psi d\boldsymbol{\psi}_e = \\
&\quad= -L_y\left(\int p(\psi,\boldsymbol{\psi}_e)v\log\frac{p(\psi)}{q(\psi)}\,d\psi d\boldsymbol{\psi}_e\right) + \int p(\psi,\boldsymbol{\psi}_e)L_y(v)\log\frac{p(\psi)}{q(\psi)}\,d\psi d\boldsymbol{\psi}_e
\end{aligned} \tag{18}$$

现在我们将平流信息流定义为

$$\mathbf{I}_{adv} = \begin{pmatrix} \int p(\psi, \boldsymbol{\psi}_e) u \log \frac{p(\psi)}{q(\psi)} \, d\psi d\boldsymbol{\psi}_e \\ \\ \int p(\psi, \boldsymbol{\psi}_e) v \log \frac{p(\psi)}{q(\psi)} \, d\psi d\boldsymbol{\psi}_e \end{pmatrix}, \tag{19}$$

以及一个信息速度通过

$$\mathbf{u}_I = \mathbf{I}_{adv}/I_\psi \qquad \text{such that} \qquad \mathbf{I}_{adv} = \mathbf{u}_I I_\psi \tag{20}$$

信息流使我们能够将平流项写为:

$$\begin{aligned} \left(\frac{\partial I_\psi}{\partial t}\right)_{adv} &= \int p(\psi, \boldsymbol{\psi}_e) F_{adv}(\psi, \psi_e) \frac{\partial}{\partial \psi} \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi \, d\boldsymbol{\psi}_e = \\ &= -\mathbf{L}_\mathbf{x} \cdot \mathbf{I}_{adv} + \int p(\psi, \boldsymbol{\psi}_e) \big(L_x(u) + L_y(v) \big) \log \frac{p(\psi)}{q(\psi)} \, d\psi d\boldsymbol{\psi}_e, \end{aligned} \tag{21}$$

其中我们使用 $\mathbf{L}_\mathbf{x} = (L_x, \ L_y)^T$，其中两个分量都是空间差分算子。

对于扩散项，我们得到如下结果。首先，我们写作 $L_{xx}(\psi) = L_x\big(L_x(\psi)\big)$。

接下来，从纬向扩散项开始，我们有:

$$\begin{aligned} &\int p(\psi, \boldsymbol{\psi}_e) \ D \ L_{xx}(\psi) \frac{\partial}{\partial \psi} \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi d\boldsymbol{\psi}_e = \\ &= DL_x \left(\int p(\psi, \boldsymbol{\psi}_e) L_x \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi d\boldsymbol{\psi}_e \right) \\ &\qquad -D \int p(\psi, \boldsymbol{\psi}_e) L_x(\psi) L_x \left(\frac{\partial}{\partial \psi} \left(\log \frac{p(\psi)}{q(\psi)} \right) \right) d\psi d\boldsymbol{\psi}_e = \\ &= DL_{xx} \left(\int p(\psi) \log \frac{p(\psi)}{q(\psi)} \, d\psi \right) - D \int p(\psi, \boldsymbol{\psi}_e) \big(L_x(\psi) \big)^2 \frac{\partial^2}{\partial \psi^2} \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi d\boldsymbol{\psi}_e \end{aligned} \tag{22}$$

经向扩散项的变换方式类似，因此总扩散项为:

$$\begin{aligned} &\left(\frac{\partial I_\psi}{\partial t}\right)_{diff} = \int p(\psi, \boldsymbol{\psi}_e) D \left(L_{xx}(\psi) + L_{yy}(\psi) \right) \frac{\partial}{\partial \psi} \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi d\boldsymbol{\psi}_e = \\ &= D \left(L_{xx} + L_{yy} \right) \left(\int p(\psi) \log \frac{p(\psi)}{q(\psi)} \, d\psi \right) \\ &\qquad -D \int p(\psi, \boldsymbol{\psi}_e) \left[\big(L_x(\psi) \big)^2 + \big(L_y(\psi) \big)^2 \right] \frac{\partial^2}{\partial \psi^2} \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi d\boldsymbol{\psi}_e = \end{aligned} \tag{23}$$

$$= D \ L_{\mathbf{xx}}(I_\psi) - D \int p(\psi, \boldsymbol{\psi}_e) \left[\big(L_x(\psi) \big)^2 + \big(L_y(\psi) \big)^2 \right] \frac{\partial^2}{\partial \psi^2} \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi d\boldsymbol{\psi}_e. \tag{24}$$

最后，对于无导数项，我们可以直接计算积分。如果 f 仅是 ψ 的函数，我们可以进一步计算为:

$$\begin{aligned} \int p(\psi, \boldsymbol{\psi}_e) \lambda f(\psi) \frac{\partial}{\partial \psi} \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi \, d\boldsymbol{\psi}_e &= \int p(\psi) \lambda f(\psi) \frac{\partial}{\partial \psi} \left(\log \frac{p(\psi)}{q(\psi)} \right) d\psi = \\ &= -\lambda \int p(\psi) \log \frac{p(\psi)}{q(\psi)} \frac{df(\psi)}{d\psi} d\psi. \end{aligned} \tag{25}$$

将所有项合并后，我们得到：

$$\frac{\partial I_\psi}{\partial t} = -\mathbf{L_x} \cdot \mathbf{I}_{adv} + D\ L_{\mathbf{xx}}(I_\psi) + \left(\frac{\partial I_\psi}{\partial t}\right)_{local}, \tag{26}$$

其中

$$\begin{aligned}\left(\frac{\partial I_\psi}{\partial t}\right)_{local} \quad = \quad & \int p(\psi, \boldsymbol{\psi}_e)\,(L_x(u) + L_y(v)) \log\frac{p(\psi)}{q(\psi)}\, d\psi d\boldsymbol{\psi}_e \\ & -D \int p(\psi, \boldsymbol{\psi}_e)\left[(L_x(\psi))^2 + (L_y(\psi))^2\right] \frac{\partial^2}{\partial \psi^2}\left(\log\frac{p(\psi)}{q(\psi)}\right) d\psi d\boldsymbol{\psi}_e \\ & + \int p(\psi, \boldsymbol{\psi}_e)\lambda f(\psi, \boldsymbol{\psi}_e)\frac{\partial}{\partial \psi}\left(\log\frac{p(\psi)}{q(\psi)}\right) d\psi\ d\boldsymbol{\psi}_e, \end{aligned} \tag{27}$$

其中，正如前文所述，若 f 不依赖于 $\boldsymbol{\psi}_e$，则最后一项可以简化。下标 局部 表示这些积分中包含 $\log p(\psi)/q(\psi)$，但无法进一步化简为作用于相对熵的算子。这一局部贡献可解释为以下因素导致的信息变化：(1) 流收敛，该项可以为正，也可以为负；(2) 修正后的耗散项；以及 (3) ψ 空间中局部强迫乘以 $\log p(\psi)/q(\psi)$ 的散度。

我们也可以将这一结果写成

$$\frac{\partial I_\psi}{\partial t} = -\mathbf{L_x} \cdot \mathbf{I} + \left(\frac{\partial I_\psi}{\partial t}\right)_{local}, \tag{28}$$

其中，平流加扩散通量写作

$$\mathbf{I} = \mathbf{I}_{adv} + D\ \mathbf{L_x}(I_\psi). \tag{29}$$

如本文所推导，信息变化的局部贡献并不等同于 Liang 和 Kleeman（2005）以及 Pires等（2024）所用的所谓自信息项。这些研究将自信息定义为：

$$\int p(\psi, \boldsymbol{\psi}_e)\frac{\partial}{\partial \psi} F(\psi, \boldsymbol{\psi}_e)\, d\psi d\boldsymbol{\psi}_e \tag{30}$$

当信息通过微分熵而非相对熵来定义时，便会得到这一表达式；从启发式角度看，它表示在其他所有变量 $\boldsymbol{\psi}_e$保持不变时， ψ 的信息变化。也就是说，当 $\boldsymbol{\psi}_e = \boldsymbol{\psi}_{fixed}$时。然而，尽管这种定义直观易懂，但在实践中却证明其用途较小，甚至会带来问题。

为说明这一点，考虑上文讨论的平流扩散示例，并假设速度分量 u 和 v 独立于 ψ 和 $\boldsymbol{\psi}_e$。在

这种情况下，将自信息项解释为 ψ 与$\boldsymbol{\psi}_e = \boldsymbol{\psi}_{fixed}$ 之间的信息演化，可得：

$$\begin{aligned}\left(\frac{\partial I_\psi}{\partial t}\right)_{local} &= \int p(\psi, \boldsymbol{\psi}_{fixed}) \left(L_x(u) + L_y(v)\right) \log \frac{p(\psi)}{q(\psi)}\, d\psi \\ &\quad + \int p(\psi, \boldsymbol{\psi}_{fixed}) \lambda f(\psi, \boldsymbol{\psi}_{fixed}) \frac{\partial}{\partial \psi} \left(\log \frac{p(\psi)}{q(\psi)}\right) d\psi, \end{aligned} \tag{31}$$

其中省略了扩散项，因为它无法写成局部自信息。

项。通过分部积分并假设边界项为零，可以进一步计算最后一项，得到：

$$\begin{aligned}\left(\frac{\partial I_\psi}{\partial t}\right)_{local} &= \int p(\psi, \boldsymbol{\psi}_{fixed}) \left(L_x(u) + L_y(v)\right) \log \frac{p(\psi)}{q(\psi)}\, d\psi \\ &\quad - \int p(\psi) \lambda \frac{\partial}{\partial \psi} \left(\frac{p(\psi, \boldsymbol{\psi}_{fixed})}{p(\psi)} f(\psi, \boldsymbol{\psi}_{fixed})\right) d\psi \\ &\quad - \int p(\psi, \boldsymbol{\psi}_e) \lambda f(\psi, \boldsymbol{\psi}_{fixed}) \frac{\partial}{\partial \psi} \log q(\psi)\, d\psi \\ &= \int p(\psi, \boldsymbol{\psi}_{fixed}) \left(L_x(u) + L_y(v)\right) \log \frac{p(\psi)}{q(\psi)}\, d\psi \\ &\quad - \int p(\psi) \lambda f(\psi, \boldsymbol{\psi}_{fixed}) \frac{\partial}{\partial \psi} \left(\frac{p(\psi, \boldsymbol{\psi}_{fixed})}{p(\psi)}\right) d\psi \\ &\quad - \int p(\psi, \boldsymbol{\psi}_{fixed}) \lambda \frac{\partial}{\partial \psi} f(\psi, \boldsymbol{\psi}_{fixed})\Bigg) d\psi \\ &\quad - \int p(\psi, \boldsymbol{\psi}_e) \lambda f(\psi, \boldsymbol{\psi}_{fixed}) \frac{\partial}{\partial \psi} \log q(\psi)\, d\psi. \end{aligned} \tag{32}$$

为了与 Liang 和 Kleeman（2005）的表述进行比较，我们现在采用微分熵而非相对熵，得到：

$$\begin{aligned}\left(\frac{\partial I_\psi}{\partial t}\right)_{local} &= \int p(\psi, \boldsymbol{\psi}_{fixed}) \left(L_x(u) + L_y(v)\right) \log p(\psi)\, d\psi \\ &\quad - \int p(\psi) \lambda f(\psi, \boldsymbol{\psi}_{fixed}) \frac{\partial}{\partial \psi} \left(\frac{p(\psi, \boldsymbol{\psi}_{fixed})}{p(\psi)}\right) d\psi \\ &\quad - \int p(\psi, \boldsymbol{\psi}_{fixed}) \lambda \frac{\partial}{\partial \psi} f(\psi, \boldsymbol{\psi}_{fixed})\Bigg) d\psi \end{aligned} \tag{33}$$

相较之下，Liang 和 Kleeman（2005）的表述只能得到：

$$\int p(\psi, \boldsymbol{\psi}_{fixed}) \frac{\partial}{\partial \psi} F(\psi, \boldsymbol{\psi}_{fixed})\, d\psi = \int p(\psi, \boldsymbol{\psi}_{fixed}) \lambda \frac{\partial}{\partial \psi} f(\psi, \boldsymbol{\psi}_{fixed})\, d\psi \tag{34}$$

它只涵盖了与局部信息流相关的部分动力学。

此外，正如引言中所指出的，在复杂的高维系统中，通常难以甚至不可能从概率密度函数 $p(\psi, \boldsymbol{\psi}_{fixed})$ 中获取样本，这进一步限制了该定义在实际应用中的实用性。

该示例说明了如何将信息的演化系统地分解为各种贡献。对系统中的每个变量ψ 重复这一过程，我们便可以在物理空间中——或任何相关的变量空间中——构建有效的信息流场。在这项引言性研究中，我们不讨论与**F**场相关的边界项；在更现实的开放域系统或具有有界变量的系统中，这些边界项可能会变得十分重要。这些问题将留待未来研究。

2.4 计算积分的实用方法

由于概率密度函数 $p(\psi, \boldsymbol{\psi}_e)$ 的形状未知，一种实用的表示方法是采用粒子表示。具体而言，我们写作：

$$p(\psi, \boldsymbol{\psi}_e) \approx \frac{1}{N} \sum_{i=1}^{N} \delta(\psi - \psi_i)\delta(\boldsymbol{\psi}_e - \boldsymbol{\psi}_{ei}), \tag{35}$$

并利用此表达式将积分转换为对粒子的求和。这种基于粒子的表示构成了许多高效数值近似方法的基础，可用于计算信息论量。

为了近似参考密度 $q(\psi)$，可以采用多种方法，包括核密度估计。在本文的示例中，我们将长期模型积分所生成样本的高斯拟合作为参考密度。这种选择不一定是近似——任何任意密度都可以作为有效的参考密度——但结果表明，对于此处考虑的系统，高斯分布具有足够的准确性。

附录 C 提供了准确计算所需求和的详细算法，包括对现有熵估计器的改编，以及新开发的基于核的估计器。这些估计器适用于诸如 $d/d\psi_i \log[p(\psi_i)/q(\psi_i)]$，之类的项，并能在集合规模相对较小时可靠地工作。

3 Kuramoto-Sivashinsky 模型的结果

Kuramoto-Sivashinsky（KS）模型如下：

$$u_t = -uu_x - u_{xx} - u_{xxxx}, \tag{36}$$

其中，下标表示对下标变量求微分。其离散化版本为：

$$u_t = -u\, L_x(u) - L_{xx}(u) - L_{xxxx}(u). \tag{37}$$

该模型包含非线性平流、倾向于增强梯度的反扩散项，以及用于抑制小尺度特征的稳定四阶扩散项。该系统具有混沌特性，其通向混沌的途径是倍周期分岔。

该系统在周期性区域上采用含有 1024 个网格点的规则网格进行离散化，并使用伪谱代码求解。概率密度函数通过由 200 个粒子构成的系综近似，每个粒子均从一个正的余弦平方形隆起开始，如下所示：

$$u_i = A\cos^2\big(2\pi(i-525)/100\big) \tag{38}$$

在位置范围 $[500, 550]$ 内，其他位置为零，振幅为 $A \sim N(1, 10^{-1})$。这一初始条件系综通过 KS模型方程向前推进，并计算每个网格点上对信息演化有贡献的不同项。

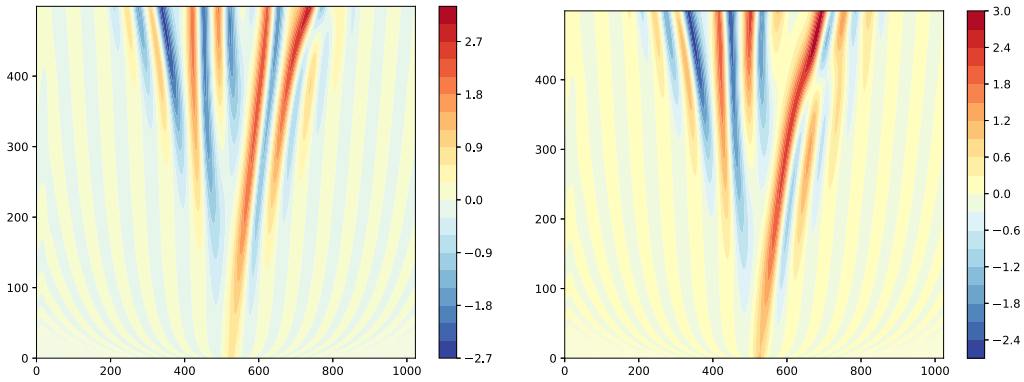

图 1. 两个系综成员时空演化的示例。这两个成员起初相似，但随着时间推移逐渐显著偏离。

图1以时空图的形式展示了两个系综成员的解。起初，这两个成员非常相似；经过约200个时间步后，开始出现明显差异，而在400个时间步后，差异变得很大。这体现了KS方程众所周知的混沌特性。其底层结构是一种波状特征，波长约为50个网格点，与这次短时模拟的初始条件相似。

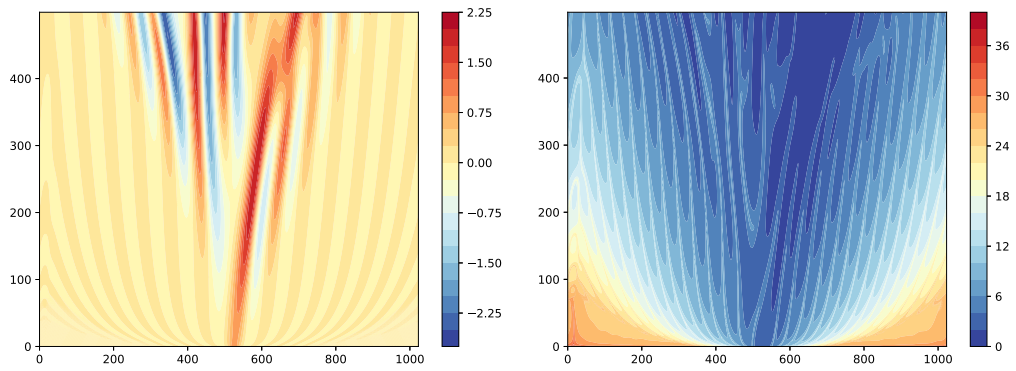

图 2. 系综平均（左）与信息（右）的时空演化。横轴为空间，纵轴为时间。

图2展示了系综平均（左图）和信息含量（右图）的演变。信息含量定义为相对熵，即库尔贝克–莱布勒散度。橙红色表示信息量高，即所有系综成员都相当相似；蓝色表示信息量低。在第350至第500个时间步之间，网格点600至800之间区域的系综平均变得不那么清晰。这表明，随着时间推移，不同系综成员的轨迹逐渐分歧。

右图显示，初始时信息量较高，反映出系综成员聚集紧密，且我们对系统状态很有信心。随着时间推移，信息量逐渐降低，表明我们对状态估计的信心减弱，不确定性增加。

信息场的空间周期性结构颇为有趣，它反映了 KS 解的波动特性。在解 u 的零交叉点附近，信息值往往较高，这表明系综成员对波的相位（或位置）比对其振幅更加一致。因此，高信息量的窄带会沿着零交叉线传播。

最后，在系综平均变得弥散的同一区域，信息降至接近零（约在第 350–500 个时间步、网格点 600–800 处）。这表明，正如模拟后期所预期的那样，系综概率密度函数开始接近参考“气候态”概率密度函数 q。

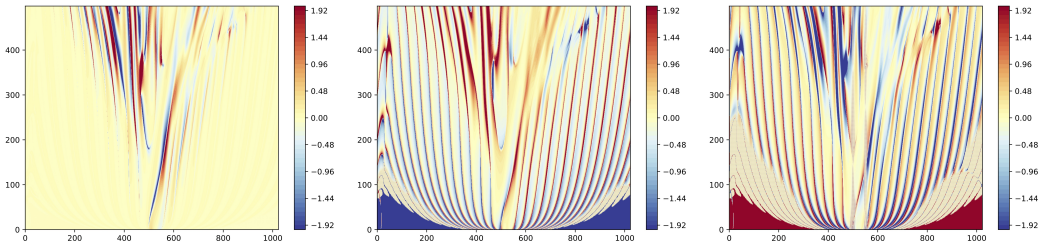

图 3。信息方程中强迫项的时空演化：非线性平流（左）、反扩散（中）和耗散（右）。

为了更好地理解信息的演变，我们分析了信息方程右端的强迫项，如图 3 所示。这些项分别对应非线性平流、反扩散和四阶耗散项，它们定义为以下各项：

$$-\int p(u,\mathbf{u}_e)\Big[u\,L_x(u)+L_{xx}(u)+L_{xxxx}(u)\Big]\frac{\partial}{\partial u}\left(\log\frac{p(u)}{q(u)}\right)\,du d\mathbf{u}_e, \tag{39}$$

其中，$\mathbf{u}_e$ 包含计算一阶、二阶和四阶导数所需的 u 值。其具体内容取决于所采用的空间离散数值格式。由于数值模型采用伪谱法，$\mathbf{u}_e$ 将包含所有网格点。为避免由此产生的高维积分，我们采用有限差分表示，最多使用 4 个网格点$\mathbf{u}_e$。尽管这是一种近似，但仍相当精确，因为我们在每个时间步都基于伪谱解进行此计算，因此离散误差不会随时间累积。为更清楚地显示各项的空间结构，绘图值被截断在范围 $[-2,2]$ 内。

人们可能会认为，非线性平流项和反扩散项会降低信息（因此显示为蓝色），因为它们会导致流动中的不稳定性。此外，耗散项预计会增加信息（因此显示为红色），因为它会耗散小尺度结构。然而，信息变化要微妙得多。

可以识别出三个主要区域：(1) 早期时刻和边缘网格点，此处出现大片红色和蓝色区域；(2) 不断扩大的中心区域，非线性平流项在此占主导；(3) 由两者之间平滑曲线所界定的区域。

在早期时刻以及远离中心轴的位置，也就是区域 (1)，可以观察到反扩散（减少信息）和耗散（增加信息）预期的作用。到了较晚时刻，尤其是在中心区域 (2)，这种行为发生了变化。反扩散项变为正值（增加信息），而耗散项的符号则像平流项一样不断变化。事实上，仔细观察会发现，平流项和反扩散项的综合作用大致抵消了耗散项。

在区域（3）的这两个区域之间，反扩散项和耗散项中可以看到平滑曲线，与图2中系综平均解的波状解以及信息本身类似。仔细观察可见，在反扩散图中，所有曲线在中心轴附近呈红色，远离中心轴则呈蓝色；而耗散图中的颜色恰好相反。如果假设空间中存在正弦解，例如 $\sin\ kx$，那么反扩散项将为 $k^2 \sin kx$，耗散项将为 $-k^4 \sin kx$，因此这两项的符号相反。然而，系综平均解的主导波长只有反扩散项和耗散项波长的一半。

认识到反扩散项和耗散项可以写成信息的反扩散和耗散，再加上局部信息源与汇（如公式（24）所示），便可解开这一悖论。KS方程的系综平均解在正值和负值处的信息都较低，这解释了主导波长的变化。

如果我们追踪该区域 (3) 内单个网格点的信息演变，会发现随着高信息波经过，反扩散项起初会降低信息，而耗散则会增加信息。随后，两者的符号反转。这使信息的净变化与耗散强迫相似，这表明在此状态下，耗散起主导作用。

最后，考虑中心区域（2）。这里非线性平流很强，信息场呈现出高值的窄带，并与波结构的过零点对齐。这里的强迫项也主要集中在这些曲线上。反扩散项倾向于增加信息（红色），而平流项和耗散项则既有正贡献也有负贡献，两者大体上抵消了前者。在这个非线性更强的状态下，现象更难理解，波状解也不复存在。非线性平流项对系综成员之间的微小差异敏感得多，其高值不再局限于这些线，并且会在某些地方增加信

息，在其他地方减少信息。非线性平流项中这种更随机的结构，大体上被耗散项抵消。我们看到，耗散项对信息的作用不同于对状态向量 u 的作用；对于后者，它总是倾向于降低能量。

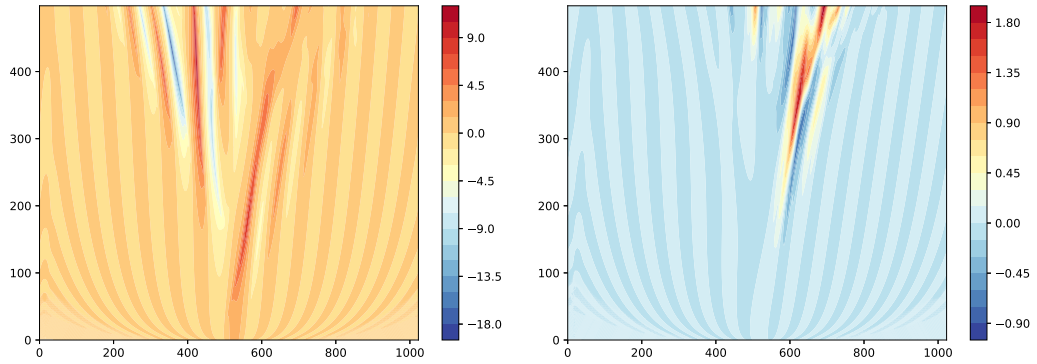

图4。 信息流的时空演化（左）以及信息流与系综平均解乘以信息之差（右）

图4左图显示了方程（19）中定义的平流信息流。正值（橙色/红色）表示流向右侧，负值（黄色/蓝色）表示流向左侧。注意，图中部两侧的流动方向交替，这与KS方程解的波状结构有关。在中央区域，我们看到与图2中系综平均值相似的结构，但振幅更大。这引出了一个问题：信息流是否仅仅是以平均速度平流的信息。

为探究这一点，我们将信息流 $\mathbf{I}_{adv}$ 与信息随流动的平流（定义为 uI_u）进行比较，方法是用前者减去后者。尽管在区域的大部分范围内，信息流与信息的平流相似，但在演化后期会出现红色（正值）和深蓝色（负值），尤其是在信息迅速减少的区域；这与不同系综成员中的不同不稳定性有关。

这个简单的例子表明，信息并非只是随流动一起平流，还会因局部不稳定性而增加或减少。当流动趋于稳定、集合离散度减小时，也会产生信息。该框架并非只是追随集合离散度，而是

上文建立的框架使我们能够理解，哪些项促使信息增加或减少。

4 浅水模型的结果

这个二维浅水模型有三个预报变量：两个风分量 u 和 v，以及流体层的深度 h。模型模拟地球上 20°N 至 70^oN 之间的一条通道，空间分辨率为 100 千米（约一度），时间步为 1 分钟。区域在 x方向上是周期性的，并设有南北两侧的固壁边界，在这些边界上 $v = 0$, 且 h 固定为其初始值。守恒形式（或通量形式）的浅水方程可写为：

$$\begin{aligned} \frac{\partial h}{\partial t} &= -\frac{\partial hu}{\partial x} - \frac{\partial hv}{\partial y} \\ \frac{\partial hu}{\partial t} &= -\frac{\partial}{\partial x}\Big(u^2 h + \frac{1}{2}gh^2\Big) - \frac{\partial}{\partial y}huv + fhv \\ \frac{\partial hv}{\partial t} &= -\frac{\partial}{\partial x}huv - \frac{\partial}{\partial y}\Big(v^2 h + \frac{1}{2}gh^2\Big) - fhu, \end{aligned} \tag{40}$$

其中， f 是科里奥利参数， g 是重力加速度。在这个模型中，科里奥利参数表示为 $f = f_0 + \beta(y - y_0)$，这称为 β 平面近似，其中 y_0 位于区域中部。

我们使用 Lax-Wendroff格式对该模型进行时间积分，该格式具有二阶精度，只需要前一个时间步，并且无需在时间或空间上加入人工扩散即可保持稳定。该格式可视为二维的时间中心、空间中心步进，并采用半时间步。模型的空间离散版本将标准中心差分算子用作 L_x 和 L_y ，此处不予展示。

模型采用2000年7月1日的ECMWF再分析资料中的气压场进行初始化。该气压场实际上来自南半球，但经过翻转后用于北半球的模拟，且不考虑地形。气压场通过 $h = 0.99 * pressure/g$ 转换为浅水方程中的变量高度。初始速度场与该气压场处于地转平衡，即以下内容的离散化版本：

$$\begin{aligned} fv &= g\frac{\partial h}{\partial x} \\ fu &= -g\frac{\partial h}{\partial y}. \end{aligned} \tag{41}$$

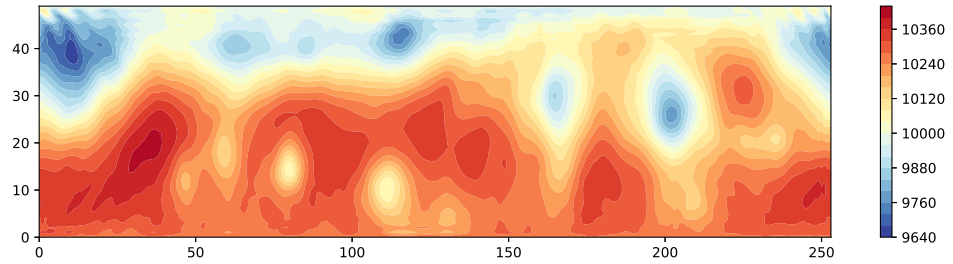

图 5. 模拟期间的高度场示例。

模型在纬向有254个网格点，在经向有50个网格点。图5展示了模拟中高度场的一个示例。空气倾向于沿等高度线流动，主导流动方向为自西向东，并沿区域中纬度蜿蜒前进。高、低气压单体也清晰可见。

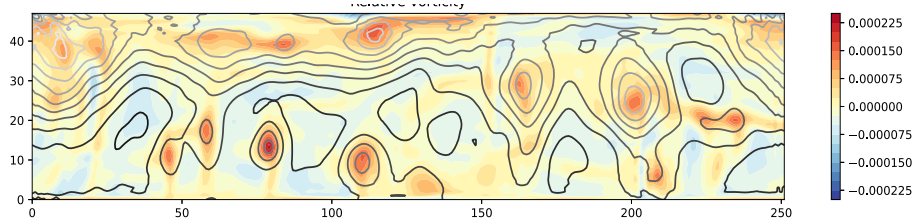

图 6. 相对涡度场，灰色等压线。

图 6 显示了相对涡度，其定义为 $\zeta = v_x - u_y$。曲线为等压线。与图 5 中的高度场相比，我们可以看到，低气压区域对应较高的相对涡度，反之亦然。区域中部的科里奥利参数为 $10^{-4}\ s^{-1}$，而典型的相对涡度值是其两倍，表明该动力学由非线性动力学主导，罗斯贝数约为 2。

通过运行模型并每隔 10 个时间步对场进行采样，生成包含 200 个模型状态的系综。随后，对每个样本运行一天，并计算与信息相关的量。参考概率密度函数 q 通过计算状态在时间、系综和纬向上的均值和方差获得。由此得到均值和方差的经向剖面，并假设它们在纬向上保持不变

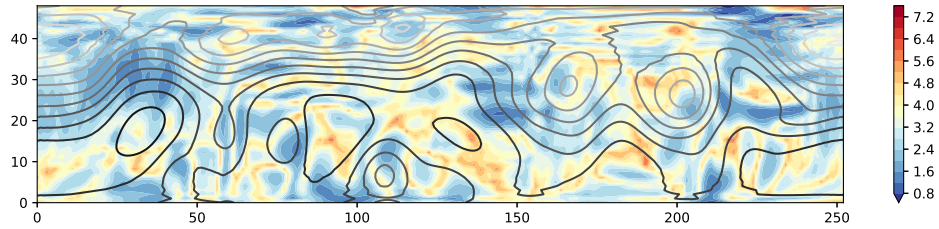

图7. 相对涡度的信息场，并叠加等压线。

方向上。我们选择这一做法，是因为模型在纬向上具有平移不变性。

图7显示了相对涡度场的信息场，并叠加了等高度线。该模型的相对涡度方程可由上述模型方程推导得出：

$$\frac{\partial \zeta}{\partial t} = -u\frac{\partial \zeta}{\partial x} - v\frac{\partial \zeta}{\partial y} - \beta v - (\zeta + f)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right), \tag{42}$$

其中常数为 $\beta = \partial f/\partial y$。该方程包含相对涡度的平流、行星涡度的平流（即 βv 项）以及一个局地拉伸项。其空间离散化版本写作：

$$\frac{\partial \zeta}{\partial t} = -u\, L_x(\zeta) - v\, L_y(\zeta) - \beta v - (\zeta + f)\Big(L_x(u) + L_y(v)\Big), \tag{43}$$

其中 L_x， L_y 为中心差分。

信息场包含小尺度结构；当系综规模增加至原来的2倍或4倍时，这些结构仍然存在，这表明它们是真实的，并且与信息动力学直接相关。总体而言，高信息区域（黄色/红色）往往集中在高压和低压气压单体的中心及其周围，而低信息区域（蓝色）往往集中在急流经向位移较大的区域。这并不意外，因为这些区域表明流动不稳定性，会导致集合离散度增大，从而使信息量降低。

更具体地说，高信息区域往往是在气流较弱（等压线较稀疏）的区域中呈现为狭窄的“条带”，但有时气流也很强，例如（东经 110–130 度、北纬 30–40 度）区域。这表明，某一时刻的信息不仅由该时刻的流动结构决定，更是流动不断演变的结果。事实上，信息无法从传统的

局部流动量中推导出来，而是需要概率密度函数信息，而后者取决于信息随时间的演变。

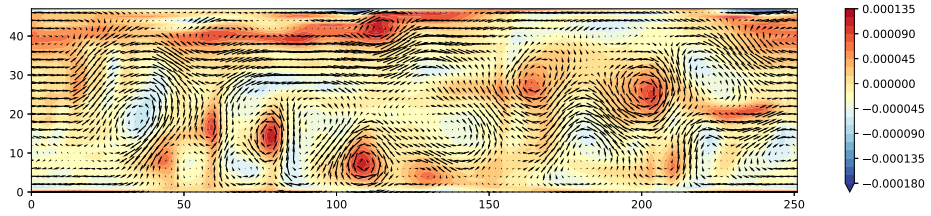

图8。 相对涡度的平流信息流，叠加于集合平均涡度场之上。

图8展示了相对涡度信息流中平流部分的瞬时图像，如方程（19）所示。由于该模型不包含显式扩散，因此可以将其视为完整的信息流；图中还叠加了相对涡度场。我们看到，信息流大体沿着相对涡度相等的等值线方向流动，这符合预期。然而，仔细观察图像，并结合图9中的放大图可以发现，信息流的方向和大小并不与涡度一致。该流动常常发生偏移，有时偏移超过5个网格点，即500千米。这种偏移并非系统性地朝某一方向，而似乎与更大尺度的流动有关。

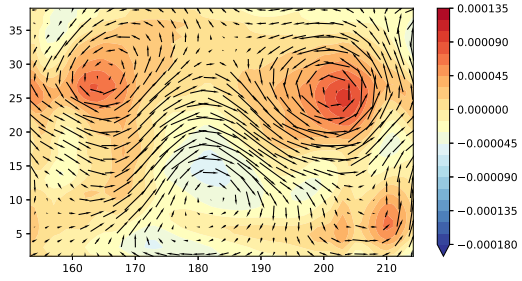

图9。 相对涡度平流信息流的细节，叠加于集合平均涡度场之上。

将信息流偏离相对涡度等值线的程度（或信息流穿过相对涡度色带的位置）与图 7 中信息量的大小进行比较，或许会让人觉得偏离越大，信息量越小（颜色越蓝），但我们并不认为两者之间存在如此直接的联系。

为进一步展示该方法论的强大之处，我们讨论两个动力学特性不同的变量的信息演变。我们以 ECMWF 气压场为初始场构建系综，同时在各处加入振幅为 1 m、空间相关尺度为 5 个网格点的小幅高斯噪声 η_1，并加入一个高斯形状的随机扰动：

$$dh(i,j) = \eta_1(i,j) + A\eta_2 \exp\Big[-\frac{(i-ic)^2 + (j-jc)^2}{2L^2}\Big], \tag{44}$$

其中 $ic = 120$、$jc = 30$、$L = 4$、$A = 100\ m$ 和 $\eta_2 \sim N(0,1)$。初始速度场与 ECMWF 气压场处于地转平衡，但与这一随机扰动不处于地转平衡。

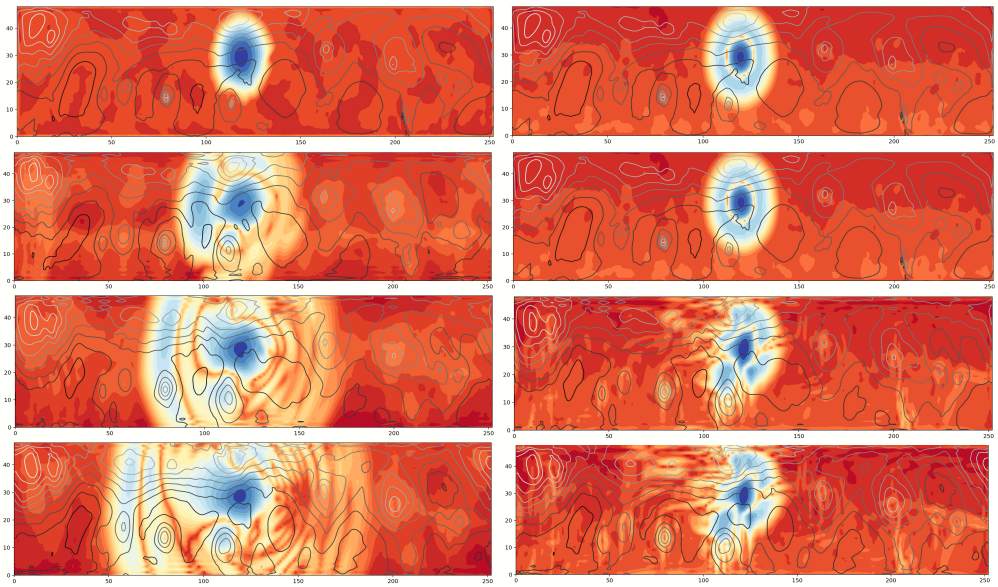

图10。 高度信息场（左）和相对涡度信息场（右）的演变，背景气压等值线以灰色表示。时间间隔为1小时40分钟。

图10展示了左侧各面板中高度场信息以及右侧各面板中相对涡度场信息的演变。面板之间的时间间隔为1小时40分钟。

左右两组面板之间显著的差异，是左侧面板中从源区向外辐射的波状模式。其传播速度约为390 m/s，这是重力波的明显信号；重力波的理论传播速度为$\sqrt{gh} \approx 330\ m/s$。地转调整过程中预计会形成重力波，重力波会被激发出来，以促进地转

—

平衡流的发展。因此，信息正随这些重力波传播。向西的波传播速度大于向东的地转流速度，导致信息传播方向与空气质点的流动方向相反。

此外，尽管主流向东流动，信息在向西流动时仍保持集中，而向东流动时则更加分散。这可能是由于一种类似的效应：当波逆着平均流传播时，波能量会趋于集中；但进一步的研究超出了本文的范围。

源区后方仍留有一个较大的信息“团块”，表明经过地转调整的信息传播与地转流密切一致，并且与右图中的相对涡度场相吻合。相对涡度信息没有向外辐射的原因在于，重力波不具有旋转性，也不携带相对涡度。相反，相对涡度信息开始根据背景速度发生变形，即开始与背景压力等值线对齐。

在高度信息和相对涡度信息的演化过程中，我们都注意到狭窄“条带”的形成，这与图7所示原始实验中的发现相似。高度信息演化表明，重力波可能在其中发挥了更大的作用。相比之下，相对涡度信息演化显示，在低信息“团块”内部形成了一个高信息环的变形，这可能表明图7中的结构更多是由平流作用形成的。

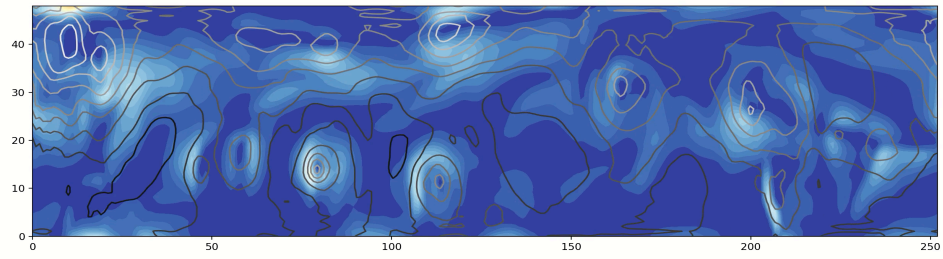

图11. 1小时40分钟后的能量场，灰色等压线。

以上表明，信息并非只是随气流平流而动，这或许意味着信息的演化与能量的演化关系更为密切。在这个正压模式中，能量主要是动能。然而，图 11 所示的模拟开始 1 小时

40 分钟时的动能场完全没有表现出这种行为。这表明，信息演化与任何地球物理量都没有直接关系，而是系统中不确定性的演化。能量的信息图呈现出高度和相对涡度行为的混合特征，此处未予展示。关于信息演化，还有许多可以探讨之处，但鉴于本文具有探索性质，我们将留待后续研究。

5 结论

本文提出了一个新的框架，用于计算和分析复杂高维系统中信息的演化。该框架解决了现有方法的若干局限，包括依赖可能为负的微分熵、处理高维积分的困难，以及自信息定义和使用冻结变量时存在的歧义。

该形式体系建立在有限维离散空间中，以便使用概率密度函数。不过，我们也在附录 B 中讨论了向无限维系统的扩展，并表明即使在这些系统中，所得表达式仍具有局部性，因此，在实际计算中将方程离散化时，所有积分仍为低维积分。

对二维平流扩散方程的初步解析应用展示了如何分析信息力，并将其分解为其他直观的项。我们引入信息流 和信息速度 作为核心概念，以帮助理解信息动力学。

我们将这一形式体系应用于 Kuramoto–Sivashinsky 模型和浅水模型的数值解，并展示它如何从信息论视角帮助我们更深入地理解系统演化。值得注意的是，我们表明信息流不一定沿着物理速度场传播。此外，系统中不同变量的信息动力学可能大不相同，本文通过一个地转调整实验展示了这一点。一般而言，信息并不直接对应于任何单个系统变量，其演化遵循的动力学也不同于能量传播等过程。

这一框架为研究复杂的高维系统提供了一种全新的视角。随着人们日益认识到，信息传播

或不确定性的传播，在许多科学领域与能量输运同样至关重要，基于信息的诊断可能会成为科学工具箱中不可或缺的补充。潜在应用包括研究完整地球系统模型中的信息动力学、高分辨率海气相互作用研究，以及云发展过程的大涡模拟。

更广泛地说，该框架有望助力我们理解系统动力学、可预测性、数据同化和风险量化。这为广泛的社会应用打开了大门，涵盖天气预报、环境监测、风险管理、保险以及数字孪生的开发。本文仅初步展示了该框架广泛的应用范围。

附录A 空间导数近似

在正文中，为了得到最终表达式，我们进行了若干近似处理。本附录将讨论我们采取的三个基本步骤。

首先，我们将 $L_x(f(\psi))$ 改写为 $L_x(\psi)$ 的形式，其中 $f(..)$ 是关于 ψ 的某个光滑函数。根据定义，我们可以写成

$$L_x(f(\psi)) = \sum_i b_i f(\psi_i) \tag{A1}$$

其中 b_i 是定义离散化格式的常数，ψ_i 是 ψ 在相邻网格点 x_i 处的值。b_i 是 $O(1/\Delta x)$。由于 L_x 表示空间导数的离散化版本，而常数的空间导数为零，因此我们发现，如果取 $f(\psi) = 1$，则有 $\sum_i b_i = 0$。据此，利用一阶泰勒展开，我们可以写成：

$$\begin{aligned} L_x(f(\psi)) &= \sum_i b_i f(\psi_i) \\ &= \sum_i b_i(\psi_i - \psi)\frac{\partial f(\psi)}{\partial \psi} + O\Big(\sum_i b_i(\psi_i - \psi)^2\Big) \\ &= \sum_i b_i \psi_i \frac{\partial f(\psi)}{\partial \psi} + O\Big(L_x(\psi^2) - 2\psi L_x(\psi)\Big) \\ &= L_x(\psi)\frac{\partial f(\psi)}{\partial \psi} + L_{xx}(\psi)O\big(\Delta x^n\big) \end{aligned} \tag{A2}$$

其中 $O(\Delta x^n)$ 是数值模型中所用数值离散化的精度。在上述步骤中，我们使用了 $\sum_i b_i f(\psi) = 0$ 和 $\sum_i b_i \psi \partial f(\psi)/\partial \psi = 0$，这是由于 $\sum_i b_i = 0$。

其次，我们将 $f(\psi)\ L_x(g(\psi))$ 改写为 $g(\psi)\ L_x(f(\psi))$ 的形式，其中 $f(..)$ 和 $g(..)$ 是 ψ 的光滑函数。定义 $\Delta f_i = f(\psi_i) - f(\psi)$ 后，我们有：

$$\begin{aligned} L_x\Big(g(\psi)f(\psi)\Big) &= \sum_i b_i f(\psi_i)g(\psi_i) = \sum_i b_i (f(\psi) + \Delta f_i)(g(\psi) + \Delta g_i) \\ &= \sum_i b_i f(\psi)(g(\psi) + \Delta g_i) + \sum_i b_i g(\psi)(f(\psi) + \Delta f_i) + \sum_i b_i \Delta g_i \Delta f_i \\ &= f(\psi)\sum_i b_i g(\psi_i) + g(\psi)\sum_i b_i f(\psi_i) + O\Big(\sum_i b_i(\psi_i - \psi)^2\Big) \\ &= f(\psi)L_x(g(\psi)) + g(\psi)L_x(f(\psi)) + L_{xx}(\psi)\big(\Delta x^n\big) \end{aligned} \tag{A3}$$

or

$$f(\psi)\ L_x(g(\psi)) = L_x\Big(g(\psi)f(\psi)\Big) - g(\psi)\ L_x(f(\psi)) + O(\Delta x^n) \tag{A4}$$

其中，$O(\Delta x^n)$ 是数值离散化格式的精度。

最后，我们考察 L_x 对概率积分的作用。我们假设 L_x 使用 m 个变量 ψ_i。我们可以写成：

$$\begin{aligned} L_x\left(\int p(\psi)g(\psi)\,d\psi\right) &= \sum_i b_i \int p(\psi_i)g(\psi_i)\,d\psi_i \\ &= \sum_i b_i \int p(\psi_1,...,\psi_m)g(\psi_i)\,d\psi_1...d\psi_m \\ &= \int p(\psi_1,...,\psi_m)\sum_i b_i g(\psi_i)\,d\psi_1...s\psi_m \\ &= \int p(\psi)L_x(g(\psi))\,d\psi. \end{aligned} \tag{A5}$$

附录B 无限维情形

从描述无限维空间中偏微分方程（PDE）解的测度演化出发，我们将推导有限维子空间中概率密度函数（pdf）演化的方程。详情可参见 Ambrosio 等（2008）。然后我们将说明，物理系统具有因果性这一事实——即每个物理系统中信息传播的最快速度有限——会使概率密度函数的演化具有局部闭式解。由此，我们将证明，正文中讨论的有限维信息演化可以扩展到无限维区域，即扩展到由偏微分方程描述的系统。

B1 偏微分方程及其相关的刘维尔方程

考虑希尔伯特空间 H上的一个偏微分方程：

$$\frac{\partial u}{\partial t} = F(u) \tag{B1}$$

其中 $F: H \to H$ 是一个表征动力学的算子。

给定初始测度 μ_0 （它表示初始条件的不确定性）在 H 上，测度 μ_t 按照刘维尔方程（也称为连续性方程或输运方程）演化：

$$\frac{\partial \mu_t}{\partial t} + \nabla \cdot (\mu_t F) = 0 \tag{B2}$$

该方程的含义可由涉及测试函数 φ 的弱恒等式 $H \to \mathbb{R}$ 得出：

$$\frac{d}{dt}\int_H \varphi(u)\,d\mu_t = \int_H \langle D\varphi(u), F(u)\rangle\,d\mu_t \tag{B3}$$

其中 $D\varphi(u)$ 是弗雷歇导数或方向导数。对于任意方向 $h \in H$，方向导数定义为：

$$\langle D\varphi(u), h\rangle = \frac{d}{ds}\Big|_{s=0} \varphi(u + sh) \tag{B4}$$

B2 投影到有限维空间与边缘化

我们通过将无限维空间H 投影到有限维空间 H_N，以获得有限维描述，从而可以写成：

$$H = H_N \oplus H_\perp \tag{B5}$$

其中， $H_N = \text{span}\{e_1, \ldots, e_N\}$ 是有限维空间， $H_\perp$ 是其正交补。例如，有限维空间可以是主文中所述的某个网格点处的变量 ψ 。

现在，我们通过以下方式定义一个柱状测试函数

$$\varphi(u) = f(\pi(u)) \tag{B6}$$

其中， $\pi:\ H \to \mathbb{R}^N$ 是到 H_N的投影，因此是其参数相对于基 $\{e_1, \ldots, e_N\}$的坐标向量； $f:\ \mathbb{R}^N \to \mathbb{R}$ 是具有紧支撑的光滑函数。具体而言，

$$\pi(u) = \langle u, e_1\rangle, \ldots, \langle u, e_N\rangle. \tag{B7}$$

方向 $h \in H$ 上的 Fréchet 导数，对于柱状 测试函数为：

$$\langle D\varphi(u), h\rangle = \frac{d}{ds}\Big|_{s=0} \varphi(u + sh) = \frac{d}{ds}\Big|_{s=0} f(\pi(u + sh)) \tag{B8}$$

由于 π 是一个线性投影算子：

$$\pi(u + sh) = \pi(u) + s\pi(h), \tag{B9}$$

我们得到：

$$\langle D\varphi(u), h\rangle = \frac{d}{ds}\Big|_{s=0} f(\pi(u) + s\pi(h)) = \langle D_\pi f(\pi(u)), \pi(h)\rangle \tag{B10}$$

其中， $D_\pi f$ 表示 f 关于其在 $\mathbb{R}^N$中的自变量的梯度。

根据伴随算子的定义，我们可以写出：

$$\langle D_\pi f(\pi(u)), \pi(h)\rangle = \langle \pi^*(D_\pi f(\pi(u))), h\rangle \tag{B11}$$

因此，Fr´echet 导数为：

$$D\varphi(u) = \pi^*(D_\pi f(\pi(u))) \tag{B12}$$

现在回到测度 μ_t 的刘维尔方程，对于柱状测试函数，我们有：

$$\langle D\varphi(u), F(u)\rangle = \langle \pi^*(D_\pi f(\pi(u))), F(u)\rangle \tag{B13}$$

并根据伴随算子的定义：

$$\langle \pi^*(D_\pi f(\pi(u))), F(u)\rangle = \langle D_\pi f(\pi(u)), \pi(F(u))\rangle \tag{B14}$$

定义 $\nu_t = \pi_{\#}\mu_t$ 在 $\mathbb{R}^N$ 上的边缘测度，其中 $\pi_{\#}$ 是从无限维空间上的测度映射到有限维空间上的概率密度函数的映射，由此可得（利用测度分解定理）：

$$\int_{\mathbb{R}^N} f(\xi)\, d\nu_t(\xi) = \int_H f(\pi(u))\, d\mu_t(u) \tag{B15}$$

于是：

$$\frac{d}{dt}\int_{\mathbb{R}^N} f(\xi)\, d\nu_t(\xi) = \int_H \langle D_\pi f(\pi(u)), \pi(F(u))\rangle\, d\mu_t(u) \tag{B16}$$

为了使其在 $\mathbb{R}^N$ 上定义一个柯尔莫哥洛夫方程，我们通过下式确定概率密度函数 ν_t 的漂移项 F_ν：

$$\int_H \langle D_\pi f(\pi(u)), \pi(F(u))\rangle\, d\mu_t(u) = \int_{\mathbb{R}^N} \langle D_\pi f(\xi), F_\nu(\xi)\rangle\, d\nu_t(\xi) \tag{B17}$$

由此将 F_ν 定义为

$$F_\nu(\xi) = \mathbb{E}[\pi(F(u)) \mid \pi(u) = \xi] \tag{B18}$$

概率密度函数的刘维尔方程现在变为：

$$\frac{\partial \nu_t}{\partial \xi} + \nabla_\xi \cdot (\nu_t F_\nu(\xi)) = 0 \tag{B19}$$

B3 闭合问题及其解

如上所推导，有效漂移系数为：

$$F_\nu(\xi) = \mathbb{E}[\pi(F(u)) \mid \pi(u) = \xi] \tag{B20}$$

要计算这个条件期望，我们需要知道 $F(u)$ 如何依赖完整的无限维解 u，而不仅仅是有限维投影 $\xi = \pi(u)$。为了更清楚地说明这一点，我们写作 $u = u_\parallel + u_\perp$，其中 $u_\parallel = \pi(u)$ 和 $u_\perp \in H_\perp$：

$$F_\nu(\xi) = \mathbb{E}[\pi(F(u_\parallel + u_\perp)) \mid u_\parallel = \xi] \tag{B21}$$

这需要条件分布 $\mathbb{P}(u_\perp \mid u_\parallel = \xi)$，它编码了完整的无限维动力学。这就是闭合问题：要在有限维空间上求解概率密度函数的演化方程，我们需要无限维空间上的测度。

然而，物理系统具有因果性，这意味着点（x，t+Δt）处的解在时间 t 仅取决于一个有限维区域 S_x 。这与任何物理系统中信息的有限最大传播速度有关。例如，在正文所述的浅水系统中，最大传播速度为 $\sqrt{gH}$，依赖域的半径将为 $\Delta t\sqrt{gH}$。

—— ——

这意味着条件期望

$$F_\nu(\xi) = \mathbb{E}[\pi(F(u_\parallel + u_\perp)) \mid u_\parallel = \xi] \tag{B22}$$

仅取决于依赖域内 $u_\perp$ 的条件分布，即在时间步 Δt内与已解析模态 ξ 具有因果联系的区域。

如果现在考虑显式有限差分格式，时间步的选择应使依赖域小于网格点之间的距离。这是为了避免数值格式变得不稳定。因此，在一个模型时间步内，我们只需要对 $u_\perp$ 中小于网格间距的那一部分取期望。当然，也可以在模型中使用隐式格式，但物理上对最大传播速度的依赖仍然存在，从而确保依赖域始终与网格间距同量级。

这一论证表明，要解出概率密度函数的演化，我们只需考虑无限维空间中位于有限维空间各网格点依赖域内的那一部分；换言之，闭合仅具有局地性。此外，我们在数值求解偏微分方程方面积累了大量经验，知道借助先进的数值格式和局地参数化，可以控制次网格尺度物理过程对已解析物理过程的影响。这表明，次网格尺度对概率密度函数演化的影响也可以用类似方式加以控制，尤其是考虑到我们可以通过在有限维空间上进行系综积分来近似这种演化。因此，这有力地说明了，基于概率密度函数的信息流计算能够告诉我们无限维空间中的信息流。

具体而言，对于某个网格点，我们可以近似

$$F_\nu(\psi) \approx F(\psi, \boldsymbol{\psi}_e, t) \tag{B23}$$

后者则是像正文中那样，通过对偏微分方程直接离散化得到的。最重要的是，这一因果论证对于理解为何计算信息流各项贡献时所需的计算是局部的、因而是低维的，至关重要。

顺带一提，抛物型偏微分方程具有无限传播速度。任意一点上的扰动都会瞬间扩散到所有其他点，因此依赖域包含整个空间。不过，应当记住，抛物型偏微分方程是有效理论，仅在特定尺度上有效，并且是通过对更小尺度进行平均或粗粒化而产生的。这一点也反映在这样一个事实上：显式有限差分格式虽然具有局部依赖域，却能对抛物型方程给出非常准确的结果（但可能并非最高效）。

最后需要提到的是，如果偏微分方程具有特定的线性形式，或具有点态非线性（例如不包含导数的非线性，如 u^3），则可以精确计算条件期望。假设漂移项分解为微分部分和局部非线性部分，

$$F(u) = Au + g(u), \tag{B24}$$

其中 A 是一个线性算子，而 $g(u)(y) = g(u(y))$ 是逐点的。于是

$$\pi(F(u)) = \pi(Au) + \pi(g(u)), \tag{B25}$$

现在将其用于线性项：

$$\pi(Au)_i = \langle e_i, Au \rangle = \langle A^* e_i, u \rangle \tag{B26}$$

如果对于每个 i，$A^* e_i \in \mathrm{span}\{e_1, \ldots, e_N\}$，那么 $\pi(Au)$ 仍然是到有限维空间的一个投影，并且这一项不存在闭合问题。例如，A 是一个常矩阵时就是如此。对于逐点非线性项，我们发现 $g(u)$ 被直接投影到有限维空间上。

附录 C：积分的实际计算

计算相对熵时，我们采用 Ebrahimi 等人（1994）的方法。我们发现，与最近邻估计量相比，这种估计量更适合我们的用途。假设

我们有 N 个样本 ψ_i，按从低到高的顺序排列，按照 Ebrahimi 等人（1994）的方法，微分熵估计量为：

$$H = \frac{1}{N}\sum_{i=1}^{N}\frac{N}{c_i m}(\psi_{i+m} - \psi_{i-m}), \tag{C1}$$

其中

$$c_i = \begin{cases} 1 + i/m & \text{for} \quad 1 \le i \le m \\ 2 & \text{for} \quad m+1 \le i \le N-m \\ 1 + (m-i-1)/m & \text{for} \quad N-m+1 \le i \le N \end{cases}, \tag{C2}$$

以及 $m = \sqrt{N} + 0.5$。此外，对于 $i \le m$，$\psi_{i+m} - \psi_{i-m} = \psi_{i+m} - \psi_1$，以及 $\psi_{i+m} - \psi_{i-m} = \psi_N - \psi_{i-m}$ 对于 $i \ge N - m + 1$。

由此可得到基于相对熵的信息估计量，假设参考密度为 $q(\psi) = N(m_{ref},\ v_{ref})$，则：

$$I_\psi = -H + \log\left(\sqrt{2\pi v_{ref}}\right) - \frac{1}{N}\sum_{i=1}^{N}\frac{(\psi_i - m_{ref})^2}{2v_{ref}}. \tag{C3}$$

研究发现，在信息估计中使用 q 的解析表达式，比使用集合估计更准确。我们也可以将此算法用于表达式

$$G_\psi = \int g(\psi)p(\psi)\log\frac{p(\psi)}{q(\psi)}\,d\psi, \tag{C4}$$

只需在求和中添加因子 $g(\psi_i)$。

我们还需要估计以下形式的积分

$$\int p(\psi, \boldsymbol{\psi}_e)g(\psi)\frac{\partial}{\partial\psi}\left(\log\frac{p(\psi)}{q(\psi)}\right)d\psi d\boldsymbol{\psi}_e. \tag{C5}$$

为此，我们使用 $d/d\psi \log p(\psi)$ 的集合估计，这有时（容易引起混淆地）称为得分函数。该方法论主要遵循 Pinkse 和 Schurter（2020年），并以核为基础。为了找到核，我们可以利用这样一个性质：对于高斯分布，$d/d\psi \log p(\psi)$ 是一条直线，因此得到如下所述的二次核。

假设样本来自一个未知的概率密度函数 $p(z)$。我们将使用恒等式

$$\frac{d}{dz}p(z) = p(z)\frac{d}{dz}\log p(z), \tag{C6}$$

如下所述。我们将通过涉及核函数 $g(t)$ 的分部积分，得到对对数梯度的局部估计。首先定义变量

$$t = \frac{z - x}{h}, \tag{C7}$$

其中 z 和 x 是标量样本， h 是带宽。核函数必须满足 $g(-1) = g(1) = 0$。接下来注意到

$$\begin{aligned}\int_{x-h}^{x+h} \frac{d}{dt}\left(g(t)\right) p(z)\, dz &= h\int_{x-h}^{x+h} \frac{d}{dz}\left(g\left(\frac{z-x}{h}\right)\right) p(z)\, dz \\ &= hg\left(\frac{z-x}{h}\right) p(z)\Big|_{x-h}^{x+h} - h\int_{x-h}^{x+h} g\left(\frac{z-x}{h}\right)\frac{d}{dz}p(z)\, dz \\ &= -h\int_{x-h}^{x+h} g\left(\frac{z-x}{h}\right) p(z)\frac{d}{dz}\log p(z)\, dz, \end{aligned} \tag{C8}$$

这里我们利用核的条件消去了常数项。现在定义对数先验在该区间上的平均梯度:

$$\int_{x-h}^{x+h} \frac{d}{dt}\left(g(t)\right) p(z)\, dz = -\frac{d}{dx}\log p(x)\int_{x-h}^{x+h} g(t)p(z)h\, dz. \tag{C9}$$

现在我们使用概率密度函数的样本估计，因此写作:

$$p(z) = \frac{1}{N}\sum_{i=1}^{N}\delta(z - z_i), \tag{C10}$$

并将 $i \in a_x$ 定义为位于区间 $[x-h, x+h]$内的那些样本。现在，我们可以定义对数在区间 $[x-h, x+h]$ 上的平均梯度:

$$\frac{d}{dx}\log p(x) = -\frac{1}{h}\frac{\sum_{i\in a}\frac{d}{dt}g(t_i)}{\sum_{i\in a} g(t_i)}. \tag{C11}$$

我们选择核函数

$$g(t) = (t+1)(1-t) = 1 - t^2 \quad \text{on} -1 < t < 1, \tag{C12}$$

并在该区间之外将其置零。这样选择的原因是，它满足 $g(1) = g(-1) = 0$，且其导数为

$$\frac{d}{dt}g(t) = -2t \quad on -1 < t < 1, \tag{C13}$$

它关于 t呈线性。这使得对于高斯分布或指数尾概率密度函数，梯度确实呈线性。使用这个核，我们得到:

$$\frac{d}{dy}\log p(x) = \frac{1}{h}\frac{\sum_{i\in a} 2t_i}{\sum_{i\in a}(1-t_i^2)}, \tag{C14}$$

其中

$$t_i = \frac{z_i - x}{h}. \tag{C15}$$

经过反复试验，带宽选取为

$$h = 2\sigma_z, \tag{C16}$$

其中， σ_z 是基点的标准差。然后检查感兴趣的测试点周围距离 h 以内是否至少有 10 个基点。如果不是，则将带宽扩大 2 倍并再次测试，直到满足该条件为止。

开放研究部分

这些数据可使用 Robin Hogan 最初开发的浅水模型轻松复现，该模型可在以下网址获取：

https://www.met.reading.ac.uk/~swrhgnrj/shallow_water_model/

致谢

本研究最初由欧洲研究委员会 CUNDA 项目（编号：694509）资助开展。作者感谢一位匿名审稿人和梁博士提出的深入意见，这些意见使论文得到了显著改进。

作者声明不存在利益冲突。

参考文献

Ambrosio, L., Gigli, N., & Savar´e, G. (2008). 度量空间与概率测度空间中的梯度流. Birkhauser Vrelag AG. (ISBN: 978-3-7643- 8721-1) Carnevale, G. F., & Holloway, G. (1982). 信息衰减与湍流的可预测性。流体力学杂志，*116*，115-121. Ebrahimi, N., Pflughoeft, K., & Soofi, E. (1994). 两种样本熵测度。《统计与概率快报》，*20*，225-2342. Evensen, G., Vossepoel, F. M., & Van Leeuwen, P. J. (2022). 数据同化基础. doi: 10.1007/978-3-030-96709-3 Fokker, A. D. (1914). 辐射场中旋转电偶极子的平均能量。《物理学年鉴》，*348*，810-820. doi: 10.1002/ andp.19143480507 Granger, C. (1969). 运用计量经济模型和交叉谱方法研究因果关系。《计量经济学》，*37*，424-438. doi: 10.1143/PTPS.64.346 Kleeman, R. (2002). 使用相对

熵衡量动力预测效用。《大气科学杂志》，*59*，2057 - 2072. doi: 10.1175/1520-0469(2002)059⟨2057:MDPUUR⟩2.0.CO;2¨Kolmogorov, A. (1931). 关于概率论中的解析方法。《数学年刊》，*104*，415-458. doi: 10.1007/BF01457949 Kuramoto, Y. (1978). 反应系统中由扩散引发的混沌。理论物理进展补编，*64*，346-367. doi: 10.1143/PTPS.64.346 Liang, X. S. (2016). 信息流与因果性作为严格的先验概念。《物理评论*E*》，*94*。doi: 10.1103/PhysRevE.94.052201 Liang, X. S. (2018). 关于相对熵的因果作用与信息流。《混沌》，*28*。doi: 10.1063/1.5010253 Liang, X. S. (2021). 归一化多元时间序列因果分析与因果图重构。《熵》，*23*。doi: 10.3390/e23060679 Liang, X. S., & Kleeman, R. (2005). 动力系统分量之间的信息传递。《物理评论快报》，*95*，244101. doi: 10.1103/PhysRevLett.95 .244101 Liang, X. S., & Kleeman, R. (2007a). 动力系统分量之间信息传递的严格形式体系。i. 连续流。《物理学*D*》，*227*，173-182. Liang, X. S., & Kleeman, R. (2007b). 动力系统分量之间信息传递的严格形式体系。i. 离散映射。《物理学*D*》，*231*，1-9. Pearl, J. (2009). 因果性。剑桥 纽约。 Pires, C. A., Docquier, D., & Vannitsem, S. (2024). 非线性系统中信息传递估计的一般理论。《物理学*D*：非线性现象》，*458*，133988. doi: 10.1016/j.physd.2023.133988¨Planck, M. (1917). 关于统计动力学中的一个定理及其在量子理论中的推广。柏林普鲁士科学院会议报告，*24*，324-341. Renyi, A. (1961). 关于信息与熵的测度。载于《第四届伯克利数学、统计学与概率论研讨会论文集，*1961*》（第547-561页）。doi: 10.1063/5.0054228 Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., . . . Zscheischler, J. (2019). 从地球系统时间序列推断因果关系

科学。《自然·通讯》，10，2553。Runge, J., Petoukhov, V., Donges, J. F., Hlinka, J., Jajcay, N., Vejmelka, M., . . . Kurths, J. (2015)。在复杂时空系统中识别因果网关和中介变量。《自然·通讯》，6，8502。Shukla, J. (1998)。混沌之中的可预测性：气候预测的科学基础。《科学》，282，728-731。Sivashinsky, G. I. (1977)。层流火焰中流体动力学不稳定性的非线性分析——I. 基本方程的推导。《宇航学报》，4，1177-1206。doi: 10.1016/0094-5765(77)90096-0 Sivashinsky, G. I. (1980)。化学计量比条件下的火焰传播。《SIAM应用数学杂志》，39，67-82。doi: 10.1137/0139007 Smith, L. A., Ziehmann, C., & Fraedrich, K. (1999)。混沌系统中的不确定性动力学与可预测性。《英国皇家气象学会季刊》，125，2855-2886。Van Leeuwen, P. J. (2003)。适用于大尺度应用的方差最小化滤波器。《月度天气评论》，131，2071-2084。doi: 10.1175/1520-0469(2002)059⟨2057:MDPUUR⟩2.0.CO;2 Van Leeuwen, P. J., DeCaria, M., Chakraborty, N., & Pulido, M. (2021)。不可干预系统中因果发现的框架。《混沌：非线性科学跨学科期刊》，31，123128。doi: 10.1063/5.0054228 Van Leeuwen, P. J., & Evensen, G. (1996)。以概率表述阐述数据同化与逆方法。《月度天气评论》，128，2898-2913。Wiener, N. (1956)。预测理论。载于 E. Beckenbach (编)，《工程师现代数学》。麦格劳-希尔，纽约。doi: 10.1143/PTPS.64.346