Article:

Oxley, W. and Kim, E.-J. (2018) Scalings and fractals in information geometry: Ornstein–Uhlenbeck processes. Journal of Statistical Mechanics: Theory and Experiment , 2018. 113401. ISSN 1742-5468

https://doi.org/10.1088/1742-5468/aae851

Scalings and fractals in information geometry: Ornstein-Uhlenbeck processes

William Oxley and Eun-jin Kim

School of Mathematics and Statistics,

University of Sheffield, Sheffield, S3 7RH, UK

Abstract

We propose a new methodology to understand a stochastic process from the perspective of information geometry by investigating power-law scaling and fractals in the evolution of information. Specifically, we employ the Ornstein-Uhlenbeck process where an initial Probability Density Function (PDF) with a given width ϵ_0 and mean value y_0 relaxes into a stationary PDF with a width ϵ, set by the strength of a stochastic noise. By utilizing the information length $\mathcal{L}$ which quantifies the accumulative information change, we investigate the scaling of $\mathcal{L}$ with ϵ. When $\epsilon = \epsilon_0$, the movement of a PDF leads to a robust power-law scaling with the fractal dimension $D_F = 2$. In general when $\epsilon \neq \epsilon_0$, $D_F = 2$ is possible in the limit of a large time when the movement of a PDF is a main process for information change (e.g. $y_0 \gg \epsilon \gg \epsilon_0$). We discuss the physical meaning of different scalings due to PDF movement, diffusion and entropy change as well as implications of our finding for understanding a main process responsible for the evolution of information.

I. INTRODUCTION

Stochastic processes play a crucial role in different disciplines [1–3]. In particular, a stochasticity in complex systems leads to a random trajectory, which is not smooth in space, possibly leading to fractals and scaling laws that are invariant under the magnification of scale. Furthermore, stochasticity and a fractal trajectory in complex systems suggest a potential utility of considering a resolution-dependent variable (e.g. see [4]). This paper aims to generalize this concept to a statistical space and to investigate scalings and fractals in the evolution of information, as described below.

While the geometry is usually referred to physical spaces, it can be extended to a statistical space by endowing a stochastic process with a metric structure [5]. Formally, the application of a differential geometry to a probability theory is the information geometry, and has attracted a great attention in both classical and quantum systems (see e.g. [6–9] and references therein). In particular, if z^i ($i = 1, 2..., d$) are the parameters of a Probability Density Function (PDF) of a stochastic variable x, a statistical manifold is the d dimensional (parameter) space spanned by z^i in which the Riemannian metric is given by the Fisher information metric g_{ij} [10] with an infinitesimal distance dl

$$\begin{aligned} dl &= g_{ij} dz^i dz^j, \\ g_{ij} &= \int dx \frac{1}{p(x,t)} \frac{\partial p(x,t)}{\partial z^i} \frac{\partial p(x,t)}{\partial z^j}, \end{aligned} \tag{1}$$

where t represents the time. As an example, for a Gaussian PDF

$$p(x,t) = \sqrt{\frac{\beta(t)}{\pi}} e^{-\beta(t)(x - \langle x \rangle)^2}, \tag{2}$$

$d = 2$ and the two parameters are the mean value $\langle x \rangle$ and the inverse temperature β related to the variance as $\beta = \frac{1}{2\langle (x - \langle x \rangle)^2}$. Therefore, for Eq. (2), g_{ij} ($i, j = 1, 2$) is given by [11]

$$g_{ij} = \int dx \frac{1}{p} \frac{\partial p}{\partial z^i} \frac{\partial p}{\partial z^j} = \begin{pmatrix} \frac{1}{2\beta^2} & 0 \\ 0 & 2\beta \end{pmatrix}, \tag{3}$$

where $i, j = 1, 2$, $z^1 = \beta$ and $z^2 = \langle x \rangle$.

A statistical manifold in the parameter space z^i enables us to quantify similarity and disparity between any two PDFs; l is dimensionless and measures the number of statistically different states between two PDFs. For instance, exactly the same two PDFs have exactly the same parameter values and the same coordinates in this statistical manifold, with zero distance in between. The distance between two PDFs increases with the disparity between the two [9, 12–17]. To understand this quantitatively, it is useful to consider a Gaussian PDF where the width of a PDF gives the uncertainty in measuring the peak position. Consequently, the two Gaussian PDFs (i.e. $\langle x \rangle$) which have the same width need to differ in their peak positions by the PDF width for them to be statistically distinguishable. Interestingly, the utility of scaling behaviour in the statistical space was demonstrated through the curvature (the gradient of g_{ij}) for the singularity in phase transitions or critical phenomena (see e.g. [7, 9]), or the information dimension [18] for the rate at which the information contained in a PDF scales with resolution (under coarse-graining).

Far from equilibrium where a PDF continuously changes, we can use time t as a parameter and generalize dl in Eq. (1) above to $d\mathcal{L}$ [11, 19–28] by considering an infinitesimal distance between the two PDFs at time t and $t+dt$ as $d\mathcal{L} = dt/\tau(t)$ where $\tau(t)$ is (time-dependent) correlation time of $p(x,t)$

$$\frac{1}{\tau^2} = \int dx \frac{1}{p(x,t)} \left[\frac{\partial p(x,t)}{\partial t}\right]^2 \left[= g_{ij} \frac{\partial z^i}{\partial t}\frac{\partial z^j}{\partial t}\right]. \tag{4}$$

Here, the last equality can be used only if parameters z^i of a PDF are known. When z^i is unknown, as is often the case where a PDF is constructed from data [20], Eq. (4) can be computed directly from $\partial_t p(x,t)$. For this reason, we use $d\mathcal{L}$ for this general case instead of dl. $d\mathcal{L}$ quantifies the infinitesimal information change along the trajectory or alternatively, the infinitesimal rate at which new information is generated during the evolution of a PDF. The total distance between PDFs at time 0 and t then defines the information length [11, 19–28]

$$\mathcal{L}(t) = \int_0^t \frac{dt_1}{\tau(t_1)} = \int_0^t dt_1 \sqrt{\int dx \frac{1}{p(x,t_1)} \left[\frac{\partial p(x,t_1)}{\partial t_1}\right]^2}. \tag{5}$$

Note that in a statistically stationary state, $\partial_t p(x,t) = 0$ and thus $\mathcal{L} = 0$. $\mathcal{L}$ depends on the time history of $p(x,t)$ and is a Lagrangian distance between PDFs at time 0 and t. Furthermore, $\mathcal{L}$ is invariant under the coordinate transformation of x. Representative of the accumulative change in information, $\mathcal{L}$ quantifies the evolution of information.

The main purpose of this paper is to investigate how $\mathcal{L}$ scales with the strength of a stochastic noise responsible for a random trajectory. To this end, we consider a relaxation problem where a PDF starting from a given initial value evolves in time until it reaches its stationary PDF as $t \to \infty$. Thus, $\tau(t)$ in Eq. (4) depends on the time (e.g. $\tau(t) \to \infty$ as $t \to \infty$) and an initial condition. We will investigate a fractal in the evolution of information by examining scalings of $\mathcal{L}(t)$ during this relaxation problem. Since an initial condition can introduce a scale that affects the scalings for small time, it is likely that the scaling changes with time, the examination of which could give us a better understanding of the dominant process responsible for the evolution of information. To gain a key insight, we employ an analytically solvable model given by the Ornstein-Uhlenbeck (O-U) process, which is a popular model for a noisy relaxation system (e.g. [2]). The remainder of this paper is organized as follows. Section II provides the background in scalings in diffusion versus relaxation problems. Section III introduces scalings in information length. Sections IV-V present detailed scaling analysis. Discussions and Conclusions are found in Section VI.

II. BACKGROUND: DIFFUSION VS RELAXATION PROBLEM

In phase-transition [7] or fully developed fluid turbulence [29], fluctuations appear on a broad range of scales. Thus, the scaling relation under the different coarse graining reveals self-similarity in scale, implying a physical law invariant across scale. Renormalization in quantum field theory utilizes a similar coarse graining in physical space. To elucidate scaling relation in a stochastic process, it is worth reviewing diffusion and relaxation problems. As the simplest example, let us consider a random walk (Brownian motion) driven by a short (delta) correlated stochastic noise ξ as

$$\frac{dx}{dt} = \xi, \tag{6}$$

where x is a variable of interest undergoing the random walk, and

$$\langle \xi \rangle = 0, \quad \langle \xi(t)\xi(t') \rangle = 2D\delta(t-t'). \tag{7}$$

Here, the angular brackets denote the average over ξ; D is the strength of ξ. The solution to Eq. (6) is simply

$$x(t) = x_0 + \int_0^t \xi(t_1)dt_1, \tag{8}$$

where $x_0 = x(t=0)$ is the initial value of x. Thus, Eqs. (7)–(8) give us the mean value and variance of x as

$$\begin{aligned} \langle x \rangle &= \langle x_0 \rangle, & (9) \\ l_{rms}^2 &= \langle (x - \langle x \rangle)^2 \rangle = \langle (x_0 - \langle x_0 \rangle)^2 \rangle + 2Dt, & (10) \end{aligned}$$

where $l_{rms} = \sqrt{\langle (x - \langle x \rangle)^2 \rangle}$ is the Root Mean Square (RMS) displacement. If x_0 is fixed as $\langle x_0 \rangle = x_0$, $\langle (x_0 - \langle x_0 \rangle)^2 \rangle = 0$ and thus $l_{rms}^2 = 2Dt$ in Eq. (10), l_{rms} increases linearly with $D^{1/2}$ and $t^{1/2}$.

To understand how this scaling is affected by an initial condition, let us consider that x_0 is not a single value but is given by the following Gaussian PDF

$$p(x_0) = \sqrt{\frac{\beta_0}{\pi}} e^{-\beta_0 (x_0 - \mu)^2}, \tag{11}$$

where $\mu = \langle x_0 \rangle$ is the mean value of x_0 and $(2\beta_0)^{-1} = \langle (x_0 - \mu)^2 \rangle$ is the variance. Then, the mean value remains the same as $\langle x(t) \rangle = \mu$, and the marginal PDF of x at t is given by the Gaussian PDF

$$p(x,t) = \sqrt{\frac{\beta_*}{\pi}} e^{-\beta_* (x - \langle x \rangle)^2}, \tag{12}$$

where $\langle x \rangle = \mu$ and $\frac{1}{2\beta_*}$ is related to the variance as

$$l_{rms}^2 = \langle (x - \langle x \rangle)^2 \rangle = \frac{1}{2\beta_*} = \frac{1}{2\beta_0} + 2Dt. \tag{13}$$

Obviously, Eq. (13) has the contribution from the two terms, the first, constant initial variance $\langle (x_0 - \mu)^2 \rangle = \frac{1}{2\beta_0}$ and the second term proportional to Dt due to the stochastic noise. For small time, Eq. (13) is dominated by the initial variance, independent of D, while for large time $t \gg \langle (x_0 - \mu)^2 \rangle / 2D$, the second dominates. The critical time determining the transition from D^0 to $D^{1/2}$ scaling of l_{rms} is $t \sim \langle (x_0 - \mu)^2 \rangle / 2D$. This simple exercise elucidates that the length scale set by the initial condition can control the scaling of l_{rms} with D (for small time).

However, at any time, the dependence on D of Eq. (13) can be isolated by considering

$$\frac{\partial l_{rms}}{\partial \ln D} = \frac{Dt}{l_{rms}}. \tag{14}$$

For small time, $\frac{\partial l_{rms}}{\partial \ln D} \propto Dt$ while for large time, $\frac{\partial l_{rms}}{\partial \ln D} \propto \sqrt{Dt} \propto l_{rms}$.

For the diffusion process given by Eq. (6), the variance of PDF in Eq. (12) keeps increasing in time, and there is no stationary PDF. To model a relaxation problem where a system settles into a stationary PDF in the long time limit, we can include a linear frictional term $-\gamma x$ in Eq. (6) so that

$$\frac{dx}{dt} = -\gamma x + \xi, \tag{15}$$

where $\gamma > 0$ is a positive frictional constant. Eq. (15) is the O-U process (e.g. [2]) noted in Section I. The solution to Eq. (15) is

$$x(t) = x_0 e^{-\gamma t} + \int_0^t e^{-\gamma(t-t_1)} \xi(t_1) dt_1, \tag{16}$$

where $x_0 = x(t=0)$ is the initial value of x. Thus, when D is constant and $x(t=0) = x_0$, Eqs. (7) and (16) give us

$$\langle x \rangle = \langle x_0 \rangle e^{-\gamma t} = \mu e^{-\gamma t}, \tag{17}$$

$$l_{rms}^2 = \langle (x - \langle x \rangle)^2 \rangle = \frac{1}{2\beta_*} = \frac{D_0}{\gamma} e^{-2\gamma t} + \frac{D}{\gamma}\left[1 - e^{-2\gamma t}\right], \tag{18}$$

where $\frac{D_0}{\gamma} = \frac{1}{2\beta_0} = \langle (x_0 - \langle x_0 \rangle)^2$ is the variance of the initial PDF given by Eq. (11). Eqs. (17)-(18) completely determine $p(x,t)$ in Eq. (12).

The scaling of l_{rms} in Eq. (18) again depends on β_0, D/γ and t. Specifically, as $t \to \infty$, Eq. (18) approaches a constant variance $l_{rms}^2 \to \frac{D}{\gamma}$, $p(x, t \to \infty)$ settling into its stationary distribution with the PDF width determined by $l_{rms} \propto \sqrt{\frac{D}{\gamma}} \propto D^{\frac{1}{2}}$. This could be obtained from the diffusion case in Eq. (6) by replacing t by γ^{-1} in the long time limit. Considering that the width of a PDF is representative of the uncertainty in measuring the mean position, D/γ determines the smallest scale (resolution) that a physical distance can be measured in the limit of large time. On the other hand, for small time, Eq. (18) is dominated by the initial condition (β_0) or the length scale at the initial time, and is thus independent of D.

The dependence on D can be pulled out again by considering

$$\frac{\partial l_{rms}}{\partial \ln D} = \frac{D}{2\gamma l_{rms}}(1 - e^{-2\gamma t}). \tag{19}$$

For a sufficiently small time $t \ll \frac{D_0}{2D\gamma}$, $\frac{\partial l_{rms}}{\partial \ln D} \propto Dt$, similar to the scaling for the diffusion problem. That is, $\frac{\partial l_{rms}}{\partial \ln D} \propto D$ is due the effect of an initial condition and diffusion for small time. On the other hand, $\frac{\partial l_{rms}}{\partial \ln D} \propto \sqrt{D}$, again similar to the scaling $\sqrt{D}$ in the diffusion problem (apart from the factor of time).

These observations suggest the utility of examining scalings using $\frac{\partial}{\partial \ln D}$. The latter is linked to a dilatation operator, discussed in more detail in Section III. Our goal is to investigate the scaling in the evolution of information by using this dilatation operator.

III. SCALING AND FRACTALS IN INFORMATION LENGTH

A dilatation provides a convenient tool for understanding fractals under the change of the resolution. Specifically, we define a dilatation operator Z

$$Z = \frac{\partial}{\partial \ln \epsilon} = \epsilon \frac{\partial}{\partial \epsilon}, \tag{20}$$

where ϵ is the resolution scale or coarse graining scale. As an example, let us consider the case [4]

$$Z\mathcal{L} = a + b\mathcal{L}, \tag{21}$$

where a and b are constant. If we let $a = \mathcal{L}_0 \tau_F$ and $b = -\tau_F$, we have $a + b\mathcal{L} = -\tau_F(\mathcal{L} - \mathcal{L}_0)$ so that the solution to Eq. (21) can be written as

$$\mathcal{L} = \mathcal{L}_0 + c\epsilon^{-\tau_F} = \mathcal{L}_0 \left[1 + \left(\frac{\lambda}{\epsilon} \right)^{\tau_F} \right], \tag{22}$$

where c is an integral constant, which is expressed in terms of another constant λ as $c = \lambda^{\tau_F} \mathcal{L}_0$. Physically, λ is the critical length scale that determines the existence of a fractal; when ϵ exceeds λ, $\mathcal{L}$ approaches a constant value $\mathcal{L} \to \mathcal{L}_0$ with no scaling region. In comparison, when the resolution ϵ is much smaller than λ, $\mathcal{L} \sim \mathcal{L}_0 \left(\frac{\lambda}{\epsilon}\right)^{\tau_F}$, exhibiting a scaling regime where the decrease in ϵ increases $\mathcal{L}$ as a power-law. This is a manifestation of a fractal, and τ_F is related to the fractal dimension D_F as $\tau_F = D_F - 1$ for one dimensional system where the topological dimension is 1. Despite the constant term in Eq. (22), the power-law scaling of $\mathcal{L}$ is readily obtained for $Z\mathcal{L}$ as

$$Z\mathcal{L} = -\mathcal{L}_0 \lambda^{\tau_F} \tau_F \epsilon^{-\tau_F}. \tag{23}$$

As will be shown later, the entropy change causes $\mathcal{L} \propto \ln \epsilon$, leading to a constant $Z\mathcal{L}$. Alternatively, $Z\mathcal{L} \propto$ constant would imply a log dependence of $\mathcal{L}$ on ϵ.

We now formalize the idea above to understand scalings and fractals in information length for the O-U process. In the O-U process, there are three length scales – the width of the initial PDF, the width of the stationary PDF, and the peak position (mean value) of the initial PDF, given by $\epsilon_0 = \sqrt{\frac{D_0}{\gamma}}$, $\epsilon = \sqrt{\frac{D}{\gamma}}$, and y_0, respectively. As the strength of the stochastic noise sets ϵ and thus the uncertainty in measuring the (peak) position of a PDF for a large time, we are interested how $\mathcal{L}$ is transformed when ϵ is rescaled and how the scaling of $\mathcal{L}$ with ϵ is affected by the relative magnitude of ϵ, ϵ_0 and y_0 and time. For our purpose, it is more convenient to consider

$$Z\mathcal{L} = 2D\frac{\partial \mathcal{L}}{\partial \ln D}. \tag{24}$$

IV. TWO SIMPLE CASES OF THE O-U PROCESS

To elucidate the meaning of $\mathcal{L}$, let us start by examining the two simples cases.

A. $D = D_0$

When $D = D_0$, a time-dependent PDF maintains the same width throughout its entire evolution, and thus the only change in $p(x,t)$ is due to the movement of the peak position (advection) as an initial PDF relaxes to a stationary PDF. $\mathcal{L}(t)$ in this case was shown to be [11, 21]

$$\mathcal{L} = \frac{\sqrt{\gamma} y_0}{\sqrt{D}}[1 - e^{-\gamma t}] = \frac{y_0 - y(t)}{\epsilon}, \tag{25}$$

where $y(t) = y_0 e^{-\gamma t}$. Eq. (25) shows that $\mathcal{L}$ is the total distance that a PDF moves in unit of ϵ so that $\mathcal{L} < 1$ for $\epsilon > y(0) - y(t)$. Furthermore, $\mathcal{L}$ is self-similar. Specifically, Eq. (25) gives us

$$D\frac{\partial \mathcal{L}}{\partial D} = \frac{1}{2} Z\mathcal{L} = -\frac{\sqrt{\gamma} y_0}{2\sqrt{D}}[1 - e^{-\gamma t}] = -\frac{y_0 - y}{2\epsilon}, \tag{26}$$

which is proportional to $D^{-\frac{1}{2}}$ or ϵ^{-1}. Comparing Eq. (25) or (26) with Eq. (22) or (23), we find $\tau_F = 1$ and the fractal dimension $D_F = 2$.

B. Entropy change: $y_0 = 0$

When $y_0 = 0$, we can show [11, 21]

$$\mathcal{L} = \left[\frac{1}{\sqrt{2}} \frac{|r|}{r} \ln\left(\frac{T}{T + r} \right) \right]_{t=0}^{t=t}. \tag{27}$$

Here

$$r = \gamma \left(\frac{D}{D_0} - 1 \right), \; T = \gamma \left[\frac{D}{D_0} (e^{2\gamma t} - 1) + 1 \right]. \tag{28}$$

In this case, $\mathcal{L}$ is caused by the change in the width of a PDF. We now show that $\mathcal{L}$ in Eq. (27) is related to entropy change. To this end, we recall that the differential entropy $S(t)$ is given by

$$S(t) = -\int dx\, p(x,t) \ln p(x,t) = \frac{1}{2} \left[1 - \ln \frac{\beta_*}{\pi} \right], \tag{29}$$

for a Gaussian PDF in Eq. (12). It is important to notice that unlike $\mathcal{L}$, $S(t)$ is independent of the peak position ($\langle x \rangle$) of a PDF. From Eq. (29), it simply follows that $S(t) - S(0) = \frac{1}{2} \ln \frac{\beta_*(0)}{\beta_*(t)}$, again independent of $\langle x \rangle$. Thus, the appearance of the log in $\mathcal{L}$ in Eq. (27) is indicative of the change in $S(t)$.

Now, from Eq. (27), we have

$$\frac{\partial \mathcal{L}}{\partial D} = \left[\frac{|r|}{r\sqrt{2}} \left(\frac{-D_0}{D\left[D(e^{2\gamma t} - 1) + D_0\right]} \right) \right]_{t=0}^{t=t}. \tag{30}$$

Evaluating (30) at the limits $t = t$ and $t = 0$, and multiplying it by D gives us

$$D \frac{\partial \mathcal{L}}{\partial D} = \frac{|r|}{r\sqrt{2}} \left(\frac{-D_0}{\left[D(e^{2\gamma t} - 1) + D_0\right]} \right) + \frac{|r|}{\sqrt{2}r}. \tag{31}$$

The dependence of Eq. (31) on D is a bit more complicated compared with Eq. (26). To understand scalings, it is useful to identify the leading order behaviour of Eq. (31) by looking at the limits of large and small time.

1. Large t

In the limit of a large time where $t \gg \frac{\ln\left(\frac{D_0}{D}\right)}{2\gamma}$ if $D < D_0$ (or for all t if $D > D_0$), Eq. (31) becomes

$$D \frac{\partial \mathcal{L}}{\partial D} \approx -\frac{|r| D_0}{r\sqrt{2} D e^{2\gamma t}} + \frac{|r|}{\sqrt{2}r}, \tag{32}$$

with the first two leading order bahavior $\propto D^0$ and D^{-1}. The leading order term $\propto D^0$ stems from the log dependence of $\mathcal{L}$ on D (due to the entropy change), as noted above.

2. *Small t*

For $t \ll \frac{D_0}{2D\gamma}$, Eq. (31) becomes

$$D\frac{\partial \mathcal{L}}{\partial D} \approx \frac{|r|}{r\sqrt{2}}\left(\frac{-D_0}{(D(2\gamma t)+D_0)}\right)+\frac{|r|}{\sqrt{2}r}. \tag{33}$$

Expanding Eq. (33) out to leading order gives

$$D\frac{\partial \mathcal{L}}{\partial D} \approx \frac{|r|\sqrt{2}\gamma D t}{rD_0}, \tag{34}$$

which is $\propto D^1$. This scaling $Z\mathcal{L} \propto D^1$ is due to the spreading of the PDF width (diffusion) due to the stochastic noise for small time, as seen in Section III. In contrast to a fractal with $D_F = 2$ for $D = D_0$ in Section IV.A, $q = y_0 = 0$ case does not give $D_F > 0$. This shows that the PDF movement could be crucial for $D_F > 0$ and will be confirmed by investigating a more general case in Section V.

V. GENERAL CASES OF THE O-U PROCESS

In general when $D \neq D_0$ and $y_0 \neq 0$, both PDF movement and entropy change contribute to $\mathcal{L}$, making it more difficult to understand the scaling behaviour. In the following, we present the results case by case depending on the signs, relative magnitude of D, D_0 and y_0, and time t. After presenting the detailed analyses, we summarize our results in Table 1 and interpret them. Readers who are interested mainly in final results can directly go to §V.D.

We start by writing down the general expression [11, 21] for $\mathcal{L}$ as follows

$$\mathcal{L} = \left[\frac{1}{\sqrt{2}}\left(\ln\left(\frac{Y-r}{Y+r}\right)\right)+\frac{\sqrt{2}}{r}H\right]_{t=0}^{t=t}, \tag{35}$$

where

$$H = \begin{cases} \sqrt{qr-r^2}\tan^{-1}\left(\dfrac{Y}{\sqrt{qr-r^2}}\right) & \text{if } qr-r^2>0, \quad (36)\\ -\dfrac{\sqrt{r^2-rq}}{2}\ln\left(\dfrac{Y-\sqrt{r^2-rq}}{Y+\sqrt{r^2-rq}}\right) & \text{if } qr-r^2<0, \quad (37)\end{cases}$$

and

$$q = \frac{\gamma^2 y_0^2}{2D_0}, \;\; Y = \sqrt{r^2+qT}. \tag{38}$$

Differentiating Eqs. (28) and (38) with respect to D, we have

$$r' = \frac{\gamma}{D_0}, \quad T' = \frac{\gamma}{D_0}\left(e^{2\gamma t} - 1\right), \quad Y' = \frac{2rr' + qT'}{2Y}, \tag{39}$$

where the prime denotes the derivative with respect to D. From Eqs. (35) and (38)-(39), we have

$$\frac{\partial \mathcal{L}}{\partial D} = \left[\frac{1}{\sqrt{2}}\left(\frac{2}{qT}(rY' - r'Y)\right) + \frac{\sqrt{2}}{r}\left(H' - \frac{H}{D - D_0}\right)\right]_{t=0}^{t=t}, \tag{40}$$

where $H' = \frac{\partial H}{\partial D}$.

In order to calculate H', we need to consider the two cases according to the sign of $qr - r^2$. First, when $qr - r^2 > 0$, Eq. (36) gives us

$$H' = \frac{F'H}{F} + \frac{F(FY' - F'Y)}{q(T + r)}, \tag{41}$$

where $F = \sqrt{qr - r^2}$, and $F' = \frac{qr' - 2rr'}{2F}$.

On the other hand, when $qr - r^2 < 0$, Eq. (37) gives us

$$H' = \frac{G'H}{G} - \frac{G(GY' - G'Y)}{q(T + r)}, \tag{42}$$

where $G = \sqrt{r^2 - rq}$ and $G' = \frac{2rr' - qr'}{2G}$. In the following subsections, we investigate different cases.

A. $q > r > 0$

In this case $qr - r^2 > 0$, so we use Eqs. (40) and (41) to obtain

$$\frac{\partial \mathcal{L}}{\partial D} = \left[\frac{1}{\sqrt{2}}\left(\frac{2}{qT}(rY' - r'Y)\right) + \frac{\sqrt{2}}{r}\left(\frac{F'H}{F} + \frac{F(FY' - F'Y)}{q(T + r)} - \frac{H}{D - D_0}\right)\right]_{t=0}^{t=t}, \tag{43}$$

where H is given by Eq. (36), and $F = \sqrt{qr - r^2}$.

For $q \gg r$, or more precisely $y_0^2 \gg \frac{2(D - D_0)}{\gamma}$, Eq. (43) becomes

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial D} \approx \Bigg[& -\frac{\gamma^2}{\sqrt{2}TYD_0}\left(\frac{D}{D_0}(e^{2\gamma t} - 1) + e^{2\gamma t} + 1\right) \\ & + \frac{D_0\sqrt{2}}{\gamma(D - D_0)}\left(-\frac{H}{2(D - D_0)} + \frac{\gamma^2(D - D_0)^2}{2D_0^2 Y e^{2\gamma t} D} - \frac{q\gamma}{2YD}\right)\Bigg]_{t=0}^{t=t}.\end{aligned} \tag{44}$$

Since $q > r > 0$ implies $D > D_0$, for simplicity, we consider the limit $D \gg D_0$ to recast Eq. (44) to

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial D} \approx &\Bigg[-\frac{\gamma^2}{\sqrt{2}TYD_0}\left(\frac{D}{D_0}(e^{2\gamma t}-1)+2\right) \\ &\quad + \frac{D_0\sqrt{2}}{\gamma D}\left(-\frac{\sqrt{qr}\tan^{-1}\frac{\sqrt{r^2+qT}}{\sqrt{qr}}}{2D} + \frac{\gamma^2 D}{2D_0^2 Y e^{2\gamma t}} - \frac{q\gamma}{2YD}\right)\Bigg]_{t=0}^{t=t}.\end{aligned} \tag{45}$$

Evaluating Eq. (45) at the lower and upper limit of time, we obtain

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial D} \approx & -\frac{\gamma^2}{\sqrt{2}TYD_0}\left(\frac{D}{D_0}(e^{2\gamma t}-1)+2\right) + \frac{\gamma\sqrt{2}}{\sqrt{r^2+q\gamma}D_0} \\ &\quad + \frac{D_0\sqrt{2}}{\gamma D}\left(-\frac{\sqrt{qr}\tan^{-1}\frac{\sqrt{r^2+qT}}{\sqrt{qr}}}{2D} + \frac{\gamma^2 D}{2D_0^2 Y e^{2\gamma t}} - \frac{q\gamma}{2YD}\right). \\ & - \frac{D_0\sqrt{2}}{\gamma D}\left(-\frac{\sqrt{qr}\tan^{-1}\frac{\sqrt{r^2+q\gamma}}{\sqrt{qr}}}{2D} + \frac{\gamma^2 D}{2D_0^2\sqrt{r^2+q\gamma}} - \frac{q\gamma}{2\sqrt{r^2+q\gamma}D}\right).\end{aligned} \tag{46}$$

In order to isolate the leading order behaviour in Eq. (46), we look at the behaviour for both large and small time in the following.

1. Large t

For $t \gg \frac{\ln 2}{2\gamma}$ we have $e^{2\gamma t} \gg 1$, $T \approx \frac{\gamma D e^{2\gamma t}}{D_0}$ and $Y \approx \sqrt{qT}$, reducing Eq. (46) to

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial D} \approx & -\frac{\sqrt{\gamma}}{\sqrt{2DD_0q}e^{\gamma t}} + \frac{\sqrt{2}}{\sqrt{D^2+\frac{\gamma y_0^2 D_0}{2}}} \\ &\quad + \frac{D_0\sqrt{2}}{\gamma D}\left(-\frac{\sqrt{q\gamma}}{2\sqrt{DD_0}}\tan^{-1}(e^{\gamma t}) - \frac{\sqrt{q\gamma D_0}}{2D^{\frac{3}{2}}e^{\gamma t}} + \frac{\gamma^{\frac{3}{2}}\sqrt{D}}{2D_0^{\frac{3}{2}}e^{3\gamma t}\sqrt{q}}\right) \\ & - \frac{D_0\sqrt{2}}{\gamma D}\left(-\frac{\sqrt{q\gamma}}{2\sqrt{DD_0}}\tan^{-1}\left(\frac{\sqrt{r^2+q\gamma}}{\sqrt{qr}}\right) + \frac{\gamma\left(D^2-\frac{D_0\gamma y_0^2}{2}\right)}{2DD_0\sqrt{D^2+\frac{D_0\gamma y_0^2}{2}}}\right).\end{aligned} \tag{47}$$

Using $\tan^- 1(e^{\gamma t}) \approx \frac{\pi}{2}$ for large t and keeping only the three leading terms in Eq. (47), we have (the three leading order terms come from terms 2,3,6 in Eq. (47))

$$\frac{\partial \mathcal{L}}{\partial D} \approx \frac{\sqrt{2}}{\sqrt{D^2 + \frac{\gamma y_0^2 D_0}{2}}} - \frac{\pi\sqrt{2qD_0}}{4\sqrt{\gamma} D^{\frac{3}{2}}} - \frac{\left(D^2 - \frac{D_0\gamma y_0^2}{2}\right)}{\sqrt{2}D^2\sqrt{D^2 + \frac{D_0\gamma y_0^2}{2}}}. \tag{48}$$

Simplifying Eq. (48) further, and multiplying by D, gives

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\pi\sqrt{2qD_0}}{4\sqrt{\gamma}\sqrt{D}} + \frac{\sqrt{D^2 + \frac{D_0\gamma y_0^2}{2}}}{\sqrt{2}D}. \tag{49}$$

It is difficult to simplify much further from here, so we look at the following two cases:
i) When $D^2 \gg \frac{D_0\gamma y_0^2}{2}$, (49) becomes

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\pi\sqrt{2qD_0}}{4\sqrt{\gamma}\sqrt{D}} + \frac{1}{\sqrt{2}}, \tag{50}$$

with the leading order behaviour $\propto D^{-\frac{1}{2}}, D^0$.
ii) When $D^2 \ll \frac{D_0\gamma y_0^2}{2}$, (49) becomes

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\pi\sqrt{2qD_0}}{4\sqrt{\gamma}\sqrt{D}} + \frac{\sqrt{D_0\gamma}y_0}{2D}, \tag{51}$$

with the leading order behaviour $\propto D^{-\frac{1}{2}}, D^{-1}$.

2. *Small t*

For $t \ll \frac{D_0}{2D\gamma}$ we have $e^{2\gamma t} \approx 2\gamma t + 1$ and $T \approx \gamma$. Eq. (46) becomes

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial D} &\approx \frac{\gamma}{\sqrt{2}\sqrt{r^2 + q\gamma}D_0}\left[2 - \left(\frac{D}{D_0}2\gamma t + 2\right)\right] \\ &\quad + \frac{D_0\sqrt{2}}{\gamma D}\left(\frac{\sqrt{qr}\tan^{-1}\frac{\sqrt{r^2+q\gamma}}{\sqrt{qr}}}{2D} - \frac{\sqrt{qr}\tan^{-1}\frac{\sqrt{r^2+qT}}{\sqrt{qr}}}{2D}\right). \\ &\quad + \frac{D_0\sqrt{2}}{\gamma D}\left(\frac{\gamma^2 D}{2D_0^2 Y e^{2\gamma t}} - \frac{q\gamma}{2YD} - \frac{\gamma^2 D}{2D_0^2\sqrt{r^2 + q\gamma}} + \frac{q\gamma}{2\sqrt{r^2+q\gamma}D}\right).\end{aligned} \tag{52}$$

If we also use $tan^{-1}x \approx x$ for small x, and simplify further, Eq. (52) becomes

$$\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\sqrt{2}\gamma D t}{\sqrt{D^2+\frac{y_0^2\gamma D_0}{2}}D_0} - \frac{\gamma^2 y_0^2 t}{2\sqrt{2}D\sqrt{D^2+\frac{y_0^2\gamma D_0}{2}}} - \frac{t\gamma^2 y_0^2\left(D^2-\frac{D_0\gamma y_0^2}{2}\right)}{2\sqrt{2}D\left(D^2+\frac{D_0\gamma y_0^2}{2}\right)^{\frac{3}{2}}} - \frac{\sqrt{2}\gamma t}{\sqrt{D^2+\frac{y_0^2\gamma D_0}{2}}}. \tag{53}$$

Multiplying Eq. (53) by D and simplifying gives

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\sqrt{2}\gamma D^2 t}{\sqrt{D^2+\frac{y_0^2\gamma D_0}{2}}D_0} - \frac{t\gamma^2 y_0^2 D^2}{\sqrt{2}\left(D^2+\frac{D_0\gamma y_0^2}{2}\right)^{\frac{3}{2}}}. \tag{54}$$

By simplifying this further, we arrive at

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\sqrt{2}\gamma D^2 t}{\left(D^2+\frac{D_0\gamma y_0^2}{2}\right)^{\frac{3}{2}}}\left[\frac{D^2}{D_0}+\gamma y_0^2\right]. \tag{55}$$

It is difficult to simplify much further from here, so we look at the following two cases:

i) When $D^2 \gg \frac{D_0\gamma y_0^2}{2}$, (55) becomes

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\sqrt{2}\gamma D t}{D_0}, \tag{56}$$

with the leading order behaviour $\propto D^1$.

ii) When $D^2 \ll \frac{D_0\gamma y_0^2}{2}$, (55) becomes

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{4\sqrt{\gamma}D^2 t}{y_0 D_0^{\frac{3}{2}}}, \tag{57}$$

with the leading order behaviour $\propto D^2$.

B. $r < 0$

In this case $qr - r^2 < 0$, so we use Eqs. (40) and (42) to obtain

$$\frac{\partial \mathcal{L}}{\partial D} = \left[\frac{1}{\sqrt{2}}\left(\frac{2}{qT}(rY' - r'Y)\right) + \frac{\sqrt{2}}{r}\left(\frac{G'H}{G} - \frac{G(GY'-G'Y)}{q(T+r)} - \frac{H}{D-D_0}\right)\right]_{t=0}^{t=t}, \tag{58}$$

where H is given by Eq. (37), and $G = \sqrt{r^2 - rq}$.

For $q \gg |r|$, or more precisely $y_0^2 \gg \frac{2|(D-D_0)|}{\gamma}$, Eq. (58) becomes

$$\frac{\partial \mathcal{L}}{\partial D} \approx \Bigg[-\frac{\gamma^2}{\sqrt{2}TYD_0}\left(\frac{D}{D_0}(e^{2\gamma t}-1)+e^{2\gamma t}+1\right) + \frac{D_0\sqrt{2}}{\gamma(D-D_0)}\left(-\frac{H}{2(D-D_0)}+\frac{\gamma^2(D-D_0)^2}{2De^{2\gamma t}YD_0^2}-\frac{q\gamma}{2DY}\right)\Bigg]_{t=0}^{t=t}. \tag{59}$$

Recalling $r < 0$ implies $D_0 > D$, we now take $D_0 \gg D$ in Eq. (59) to obtain

$$\frac{\partial \mathcal{L}}{\partial D} \approx \left[-\frac{\gamma^2}{\sqrt{2}TYD_0}\left(e^{2\gamma t}+1\right) - \frac{\sqrt{2}}{\gamma}\left(\frac{H}{2D_0} + \frac{\gamma^2}{2De^{2\gamma t}Y} - \frac{q\gamma}{2DY}\right)\right]_{t=0}^{t=t}. \quad (60)$$

In this case, we have $r^2 \ll qT$ for all time. Using this, and substituting our limits for t in (60), we obtain

$$\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\gamma^2}{\sqrt{2}T^{\frac{3}{2}}\sqrt{q}D_0}\left(e^{2\gamma t}+1\right) + \frac{\sqrt{2\gamma}}{\sqrt{q}D_0} + \frac{q-\frac{\gamma}{e^{2\gamma t}}}{\sqrt{2}D\sqrt{qT}} + \frac{\gamma - q}{\sqrt{2}D\sqrt{q\gamma}}$$
$$+ \frac{\sqrt{q}}{2\sqrt{2\gamma}D_0}\left[\ln\left(1 - \frac{2\left(1-\frac{D}{2D_0}\right)}{\sqrt{1+\frac{D}{D_0}(e^{2\gamma t}-1)} + \left(1-\frac{D}{2D_0}\right)}\right) - \ln\left(\frac{D}{4D_0}\right)\right]. \quad (61)$$

We will now look at behaviour for both large and small time.

1. *Large t*

For $t \gg \frac{\ln\left(\frac{D_0}{D}\right)}{2\gamma}$, we have $e^{2\gamma t} \gg 1$. Eq. (61) thus gives us

$$\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\sqrt{\gamma D_0}}{D^{\frac{3}{2}}\sqrt{2q}e^{\gamma t}} + \frac{\sqrt{\gamma}\sqrt{2}}{\sqrt{q}D_0} + \frac{\left(q-\frac{\gamma}{e^{2\gamma t}}\right)\sqrt{D_0}}{\sqrt{2}D^{\frac{3}{2}}\sqrt{q\gamma}e^{\gamma t}} + \frac{\gamma-q}{\sqrt{2}D\sqrt{q\gamma}}$$
$$+ \frac{\sqrt{q}}{2\sqrt{2\gamma}D_0}\left[\ln\left(1 - \frac{2\sqrt{D_0}}{\sqrt{D}e^{\gamma t}}\right) - \ln\left(\frac{D}{4D_0}\right)\right]. \quad (62)$$

Simplifying Eq. (62) further, and multiplying by D, we have (the three leading terms come from terms 2,4,6 in Eq. (62))

$$D\frac{\partial \mathcal{L}}{\partial D} \approx \frac{2D}{\sqrt{\gamma D_0}y_0} + \frac{\sqrt{\gamma}y_0 D}{4D_0^{\frac{3}{2}}}\ln\left(\frac{4D_0}{D}\right) - \frac{\sqrt{\gamma}y_0}{2\sqrt{D_0}}, \quad (63)$$

with the leading order behaviour $\propto D\ln D$, D^1, D^0.

2. *Small t*

For $t \ll \frac{1}{2\gamma}$ we have $e^{2\gamma t} \approx 1 + 2\gamma t$ and $T \approx \gamma$. Thus, we can show that Eq. (61) leads to

$$\frac{\partial \mathcal{L}}{\partial D} \approx \frac{\sqrt{\gamma}}{\sqrt{2q}D_0}(2 - 2\gamma t - 2) + \frac{\sqrt{q}}{2\sqrt{2\gamma}D_0}\left[\ln\left(\frac{D(1+2\gamma t)}{4D_0}\right) - \ln\left(\frac{D}{4D_0}\right)\right]$$
$$+ \frac{1}{2D\sqrt{q\gamma}}\left(q - \frac{\gamma}{1+2\gamma t} + \gamma - q\right). \quad (64)$$

Simplifying Eq. (64) further, and multiplying by D, we arrive at

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{2D\sqrt{\gamma}t}{\sqrt{D_0}y_0} + \frac{2t\sqrt{\gamma D_0}}{y_0} + \frac{D\sqrt{\gamma}y_0}{4D_0^{\frac{3}{2}}}\ln(2\gamma t+1), \tag{65}$$

with the leading order behaviour $\propto D^1$, D^0.

C. $r > q > 0$

In this case $qr - r^2 < 0$, so we use Eqs. (40) and (42) to obtain

$$\frac{\partial \mathcal{L}}{\partial D} = \left[\frac{1}{\sqrt{2}}\left(\frac{2}{qT}(rY' - r'Y)\right) + \frac{\sqrt{2}}{r}\left(\frac{G'H}{G} - \frac{G(GY' - G'Y)}{q(T+r)} - \frac{H}{D-D_0}\right)\right]_{t=0}^{t=t}, \tag{66}$$

where H is given by (36) and $G = \sqrt{r^2 - rq}$.
For $q \ll r$, or more precisely $y_0^2 \ll \frac{2(D-D_0)}{\gamma}$, Eq. (66) becomes

$$\frac{\partial \mathcal{L}}{\partial D} \approx \left[-\frac{\gamma^2 r}{\sqrt{2}TYD_0(T+r)}\left(\frac{D}{D_0}(e^{2\gamma t}-1) + e^{2\gamma t} + 1\right)\right]_{t=0}^{t=t}. \tag{67}$$

Since $r > q > 0$ implies $D > D_0$, we look at $D \gg D_0$ in Eq. (67) to obtain

$$\frac{\partial \mathcal{L}}{\partial D} \approx \left[-\frac{\gamma^2 r}{\sqrt{2}TYD_0(T+r)}\left(\frac{D}{D_0}(e^{2\gamma t}-1) + 2\right)\right]_{t=0}^{t=t}. \tag{68}$$

By evaluating Eq. (68), and using $r^2 \gg q\gamma$, we obtain

$$\frac{\partial \mathcal{L}}{\partial D} \approx -\frac{\gamma^2}{\sqrt{2}TYD_0 e^{2\gamma t}}\left(\frac{D}{D_0}(e^{2\gamma t}-1) + 2\right) + \frac{\sqrt{2}}{D}. \tag{69}$$

We will now look at behaviour for both large and small time.

1. *Large t*

For $t \gg \frac{\ln 2}{2\gamma}$ we have $e^{2\gamma t} \gg 1$ and we rewrite Eq. (69) as

$$\frac{\partial \mathcal{L}}{\partial D} \approx \frac{\sqrt{2}}{D} - \frac{1}{e^{2\gamma t}\sqrt{2D^2 + \gamma y_0^2 D e^{2\gamma t}}}. \tag{70}$$

Multiplying Eq. (70) by D gives

$$D\frac{\partial \mathcal{L}}{\partial D} \approx \sqrt{2} - \frac{D}{e^{2\gamma t}\sqrt{2D^2 + \gamma y_0^2 D e^{2\gamma t}}}. \tag{71}$$

It is difficult to simplify much further from here, so we look at the following two cases.
i) When $D \gg \frac{\gamma y_0^2 e^{2\gamma t}}{2}$, Eq. (71) becomes

$$D\frac{\partial \mathcal{L}}{\partial D} \approx \sqrt{2} - \frac{1}{e^{2\gamma t}\sqrt{2}}, \tag{72}$$

with the leading order behaviour $\propto D^0$.
ii) When $D \ll \frac{\gamma y_0^2 e^{2\gamma t}}{2}$, Eq. (71) becomes

$$D\frac{\partial \mathcal{L}}{\partial D} \approx \sqrt{2} - \frac{\sqrt{D}}{e^{3\gamma t}\sqrt{\gamma} y_0}, \tag{73}$$

which is $\propto D^0, D^{\frac{1}{2}}$ to leading order.

2. *Small t*

If we take $t \ll \frac{D_0}{2D\gamma}$ and use $e^{2\gamma t} \approx 1 + 2\gamma t$, Eq. (69) becomes

$$\frac{\partial \mathcal{L}}{\partial D} \approx -(\sqrt{2}\gamma t)\left(\frac{1}{D_0} - \frac{2}{D}\right). \tag{74}$$

Multiplying Eq. (74) by D, and using $D \gg D_0$ gives

$$D\frac{\partial \mathcal{L}}{\partial D} \approx -(\sqrt{2}\gamma t)\left(\frac{D}{D_0}\right), \tag{75}$$

which is $\propto D^1$ to leading order.

D. A summary of the results

Table 1 summarizes our results of scalings of $Z\mathcal{L}$ for the various limits that were considered in Sections III-IV. The leading order term is marked by an underline.

To understand Table 1, it is useful to translate the limits involving q, r, D and D_0 into three length scales y_0, $\epsilon = \sqrt{\frac{D}{\gamma}}$, $\epsilon_0 = \sqrt{\frac{D_0}{\gamma}}$, introduced in Section III. To this end, we begin by noting that

$$\frac{D}{D_0} = \sqrt{\frac{\epsilon}{\epsilon_0}}, \quad \frac{q}{r} = \frac{y_0^2}{\sqrt{2}\epsilon^2}, \quad \frac{r}{\sqrt{\gamma q}} = \frac{\sqrt{2}\epsilon^2}{y_0 \epsilon_0}. \tag{76}$$

From Eq. (76), we can find that

$$q \gg r \gg \sqrt{q\gamma} \implies \sqrt{2}\frac{\epsilon^2}{\epsilon_0} \gg y_0 \gg \sqrt{2}\epsilon, \tag{77}$$

$$q \gg \sqrt{q\gamma} \gg r \implies y_0 \gg \sqrt{2}\frac{\epsilon^2}{\epsilon_0} \gg \sqrt{2}\epsilon, \tag{78}$$

$$q \ll r \implies y_0 \ll \sqrt{2}\epsilon. \tag{79}$$

Case	Limits taken \ Time	$t \gg \frac{\ln 2}{2\gamma}$	$t \gg \frac{\ln \frac{D_0}{D}}{2\gamma}$	$t \ll \frac{1}{2\gamma}$	$t \ll \frac{D_0}{2D\gamma}$	all time
$q > r > 0$	$q \gg r \gg \sqrt{q\gamma}, D \gg D_0$	$\underline{D^{-\frac{1}{2}}}, D^0$	-	-	$\underline{D^1}$	-
	$q \gg \sqrt{q\gamma} \gg r, D \gg D_0$	$\underline{D^{-\frac{1}{2}}}, D^{-1}$	-	-	$\underline{D^2}$	-
$r < 0$	$q \gg \lvert r \rvert, D_0 \gg D$	-	$D^0, \underline{D^1, Dln(D)}$	$\underline{D^1, D^0}$	-	-
$r > q > 0$	$q \ll r, D \gg D_0$	-	-	-	$\underline{D^1}$	-
	$q \ll r, D \gg D_0, D \gg \frac{\gamma y_0^2 e^{2\gamma t}}{2}$	$\underline{D^0}$	-	-	-	-
	$q \ll r, D \gg D_0, D \ll \frac{\gamma y_0^2 e^{2\gamma t}}{2}$	$D^{\frac{1}{2}}, \underline{D^0}$	-	-	-	-
$q = 0$	none	-	$D^{-1}, \underline{D^0}$	-	$\underline{D^1}$	-
$r = 0$	none	-	-	-	-	$\underline{D^{-\frac{1}{2}}}$

TABLE I: Summary of the scalings of $\mathcal{L}$ in different limits. The leading order term is underlined.

We recall from Section IV, $Z\mathcal{L} \propto D^{-\frac{1}{2}}$ for $r = 0$ for all time, supporting a fractal with $D_F = 2$; the case of $q = y_0 = 0$ does not support a fractal with $D_F > 0$. These cases are also included in Table 1 for completeness.

By examining Table 1 for all other cases and Eqs. (77)-(79), we can conclude that a necessary condition for $D_F = 2$ is (i) $y_0 \gg \epsilon$ (when y_0 exceeds the uncertainty introduced by the stochastic noise); (ii) the limit of large time so that the effect of the movement of a PDF dominates over the entropy change, with no effect of an initial condition. In particular, it is interesting to see that the cases of Eqs. (77) and (78) shown in the first two rows in Table 1 have the same leading order term $\propto D^{-\frac{1}{2}}$ with $D_F = 2$ in the limit of large time. The difference between these two cases appears at the second order. Specifically, for the case of Eq. (77), the second order term $\propto D^0$ is due to the entropy change. In comparison, for the case of Eq. (78), the second order term $\propto D^{-1} \propto \epsilon^{-2}$ suggests a scaling region with $D_F = 3$. This additional fractal scaling arises from the fact that the width of an initial PDF is sufficiently small with a well-separated length scale between y_0, ϵ and ϵ_0; the intermediate length scale $\sqrt{2}\frac{\epsilon^2}{\epsilon_0}$ ($\ll y_0$) is responsible for the appearance of $D_F = 3$ at the second order.

In comparison, when $\epsilon \gg y_0$, the leading order term in Table 1 is $\propto D^0$ in the limit of large time, reflecting that $\mathcal{L}$ is due to the change in entropy. Almost similar behaviour is seen for the case of $r < 0$ ($D_0 > D$) where the width of an initial PDF introduces a scale that is much larger than the uncertainty associated with ϵ due to the stochastic noise. These observations suggest that there are two critical length scales for a fractal scaling region; the upper limit λ in Eq. (22) is y_0 while the lower limit is given by ϵ_0.

Finally, for a sufficiently small time, the leading order behaviour in almost all cases in Table 1 exhibits the scaling $\propto D^1$, which is due to the initial diffusion, as noted in Section IV.B. The exceptional scaling D^2 appears due to the existence of the intermediate length scale $\sqrt{2}\frac{\epsilon^2}{\epsilon_0}$ ($\ll y_0$) for the case of Eq. (78).

VI. DISCUSSIONS AND CONCLUSIONS

We proposed a new methodology to understand a stochastic process from the perspective of information theory by investigating power-law scalings and fractals in the evolution of information. Specifically, we employed the Ornstein-Uhlenbeck process where an initial Probability Density Function (PDF) with a given width ϵ_0 and mean value y_0 relaxes into a stationary PDF with a width ϵ, set by the strength of a stochastic noise D. We utilized the information length $\mathcal{L}$ which quantifies the accumulative information change, ϵ playing the role of the unit of information or resolution in the long time limit, and investigated the scaling of $\mathcal{L}$ with ϵ by calculating $\frac{\partial}{\partial \ln \epsilon}\mathcal{L}$. When $\epsilon = \epsilon_0$, a robust power-law scaling with the fractal dimension $D_F = 2$ results from the movement of a PDF. In general, $\mathcal{L}$ exhibits $D_F = 2$ when the movement of a PDF was a main process. In particular, when $D \neq D_0$, a power-law scaling with $D_F = 2$ is possible when $y_0 \gg \epsilon \gg \epsilon_0$. Physically, $D_F = 2$ represents the increase in information as the uncertainty is reduced, or alternatively, the decrease in information as the uncertainty increased. This is similar to the result in [16] where the Fisher information was shown to decrease under the process of coarse-graining. We also discussed the meaning of different scalings of $Z\mathcal{L}$; $Z\mathcal{L} \propto D$ due to the diffusion while a constant $Z\mathcal{L}$ due to entropy change. It is useful to contrast our proposed fractal dimension of $\mathcal{L}$ with the information dimension $D_1 = \lim_{\epsilon \to 0} \frac{S(t)}{\ln \epsilon}$ (e.g. see [18]), defined based on the entropy $S(t)$ in Eq. (29) where ϵ is a coarse-graining scale. For a Gaussian PDF, $S(t)$ is independent of $\langle x \rangle$, and D_1 thus does not tell us anything about the movement of a PDF. This endorses that our proposed method is new, with a potential to contribute to understanding the evolution of information.

For a nonlinear process, the scaling analysis is much harder due to the unavailability of an exact analytical solution. In [22, 28], we utilized both semi-analytical and numerical approaches to study an nth order nonlinear process governed by $\partial_t x = -\gamma x^n + \xi$ ($n = 3, 5, 7$) and showed that $\mathcal{L}(t \to \infty) \equiv \mathcal{L}_\infty \propto D^{-\frac{n-1}{3n-1}}$ in the limit of $D_0 \gg D$. If we use $p(x, t \to \infty) \propto \exp\left[-\frac{\gamma}{(n+1)D} x^{n+1}\right]$ and let ϵ be the width of the stationary PDF, we obtain $\epsilon \propto D^{\frac{1}{n+1}}$ and thus $\mathcal{L}_\infty \propto \epsilon^{-\frac{n^2-1}{3n-1}}$. This implies that $D_F \sim 1 + \frac{n^2-1}{3n-1}$, which increases with n for $n \gg 1$. This suggests that a nonlinear force generates a more complex fractal structure in the information evolution. How $\mathcal{L}(t)$ scales in general is yet to be studied in future.

We note that there are nonlinear processes where an exact time-dependent non-Gaussian PDF can be found from a Gaussian PDF by a change of variables. Example include the logistic and Gompertz equations with a delta-correlated multiplicative noise in [25] (and also the motion of a quantum particle in [26]). As the information length is invariant under the change of the coordinates, it can be conveniently calculated by using the Gaussian PDF. It will be of interest to investigate scalings and fractals in these cases as well as in the case of more general non-Gaussian PDFs.

Acknowledgement: This research was partially supported by the London Mathematical Society (URB 17-18-10).

[1] Risken, H. The Fokker-Planck equation: methods of solutions and applications (Berlin: Springer, 2013).

[2] Klebaner, F. Introduction to Stochastic Calculus with Applications (Imperial College Press, 2012).

[3] Gardiner, C. Stochastic methods, 4th Ed., (Springer, 2008).

[4] Nottale, L. *Scale relativity and fractal space-time: a new approach to unifying relativity and quantum mechanics*, Imperial College Press, London (2011).

[5] Gibbs, A.L.; Su, F.E. On choosing and bounding probability metrics. *Int. Stat. Rev.* **2002**, *70*, 419–435.

[6] Felice, D.; Cafaro, C.; Mancini, C. Information geometric methods for complexity. *Chaos* **2018**, *28*, 032101.

[7] Kumar, P.; Sarkar, T. Geometric critical exponents in classical and quantum phase transitions. *Phys. Rev. E* **2014**, *90*, 042145.

[8] Maity, R.; Mahapatra, S; Sarkar, T. Information geometry and the renormalization group. *Phys. Rev. E* **2015**, *92*, 052101.

[9] Ruppeiner, G. Thermodynamics: A Riemannian geometric model. *Phys. Rev. A* **1979**, *20*, 1608.

[10] Frieden, B.R. *Science from Fisher Information* (Cambridge Univ. Press, 2000).

[11] Kim, E.; Lee, U.; Heseltine, J.; Hollerbach, R. Geometric structure and geodesic in a solvable model of nonequilibrium process. *Phys. Rev. E* **2016**, *93*, 062127.

[12] Wootters, W.K. Statistical distance and Hilbert space. *Phys. Rev. D* **1981**, *23*, 357.

[13] Schlögl, F. Thermodynamic metric and stochastic measures. *Z. Phys. B Cond. Matt.* **1985**, *59*, 449–454.

[14] Braunstein, S.L.; Caves, C.M. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **1994**, *72*, 3439.

[15] Nulton, J.; Salamon, P.; Andresen, B.; Anmin, Q. Quasistatic processes as step equilibrations. *J. Chem. Phys.* **1985**, *83*, 334–338.

[16] Plastino, A.R.; Casas, M; Plastino, A. Fisher's information, Kullback's measure, and H-theorems. *Phys. Lett. A* **1998**, *246*, 498.

[17] Polettini, M; Esposito, M. Nonconvexity of the relative entropy for Markov dynamics: A Fisher information approach. *Phys. Rev. E.* **2013**, *88*, 012112.

[18] Ott, E. *Chaos in Dynamical Systems*; Cambridge University Press: Cambridge, 2002.

[19] Nicholson, S.B.; Kim, E. Investigation of the statistical distance to reach stationary distributions. *Phys. Lett. A* **2015**, *379*, 83-88.

[20] Nicholson, S.B.; Kim, E. Structures in sound: Analysis of classical music using the information length. *Entropy* **2016**, *18*, 258.

[21] Heseltine, J.; Kim, E. Novel mapping in non-equilibrium stochastic processes. *J. Phys. A* **2016**, *49*, 175002.

[22] Kim, E.; Hollerbach, R. Signature of nonlinear damping in geometric structure of a nonequilibrium process. *Phys. Rev. E* **2017**, *95*, 022137.

[23] Hollerbach, R.; Kim, E. Information geometry of non-equilibrium processes in a bistable system with a cubic damping. *Entropy* **2017**, *19*, 268.

[24] Kim, E.; Tenkès, L.-M.; Hollerbach, R.; Radulescu, O. Far-from-equilibrium time evolution between two gamma distributions. *Entropy* **2017**, *19*, 511.

[25] Tenkès, L.-M.; Hollerbach, R.; Kim, E. Time-dependent probability density functions and information geometry in stochastic logistic and Gompertz models. *J. Stat. Mech. Theo. & Exp.* **2017**, *123201.*

[26] Kim, E.; Lewis, P. Information length in Quantum system. *J. Stat. Mech. Theo. & Exp.* **2018**, *043106.*

[27] Kim, E; Investigating information geometry in classical and quantum systems through information length. *Entropy* **2018**, *20*, 574.

[28] Hollerbach, R.; Dimanche, D.; Kim, E. Information geometry of nonlinear stochastic systems, *Entropy* **2018**, *20*, 550.

[29] Pope, S.B. Turbulent flows (Cambridge Univ. Press, 2000).