Preprints.org

Article

Black Holes and Kullback-Leibler Divergence, Decomposing Path-Dependent Processes

Shoshauna Gauvin [*]

Article

Black Holes and Kullback-Leibler Divergence, Decomposing Path-Dependent Processes

Shoshauna Gauvin[ID]

University of Waterloo, Department of Physics; s3gauvin@uwaterloo.ca

Abstract: This work develops an information geometric framework that unifies the principle of stationary action with the minimization of Kullback-Leibler divergence in stochastic systems to probe the shortcomings of General Relativity. By reformulating path probabilities using maximum entropy methods, we decompose complex, path-dependent distributions into symmetric sub-components, thereby isolating error terms that correspond to missing dynamics in black hole mechanics. Analytical and computational analyses reveal that an entropy-maximizing state $\Delta D_{KL} = 0$ signifies an ideal match between theory and observation, while deviations expose hidden mechanisms underlying curvature evolution. Extending the approach with Fisher Information Geometry and Ricci flow, we derive a novel relation, $\gamma(T \cdot ds/dE) = d\theta/d\tau$, which links thermodynamic quantities with spacetime expansion dynamics. This connection offers fresh insights into phenomena such as black hole evaporation and event-horizon viscosity. Our findings underscore the necessity of incorporating path-dependence into gravitational models, providing a possible avenue for reconciling classical and quantum descriptions of spacetime evolution.

Keywords: entropy; Kullback-Leibler divergence; information geometry; path-dependent processes; black holes; Ricci flow

---

1. Introduction

General Relativity achieved remarkable success in predicting a wide range of astronomical observations. However, there remain narrow regimes where Einstein's theory falls short. In these domains, an Information Geometry approach offers new techniques to probe the underlying deficits. In this paper, we demonstrate that the principle of stationary action is equivalent to minimizing the Kullback-Leibler (KL) divergence in stochastic systems. This equivalence allows us to decompose complex, path-dependent distributions into symmetric sub-components. By leveraging the additivity property of divergence, we isolate an "error" distribution that captures the information where General Relativity fails to match observation.

Building on established work that connects maximum entropy with stationary action to recover dynamical laws [17,26,28], our formulation expresses the KL divergence in terms of the path-dependent distribution $P[\Gamma]$, preserving the Euler-Lagrange equations of classical mechanics. Through computational and analytical methods it is shown that $\Delta D_{KL} = 0$, an entropy maximum, corresponds to an ideal state where theory and observation align. Specifically, one can decompose a complex distribution into distributions that have specific properties that are invariant under specific transformations. These invariance are closely linked to conserved quantities and informational symmetries akin to Noether's Theorem. This model-to-observation deficit is leveraged to analyze the "missing" dynamics of General Relativity through deriving a measure for multi-path dependence. Finally, we derive a novel relationship between entropy, energy, and the expansion scalar:

$$\gamma \frac{TdS}{dE} = \frac{d\theta}{d\tau}$$

which offers fresh insight and a more intuitive understanding into black hole relaxation and viscosity.

2. Methods and Established Theory

2.1. Review of Stochastic Action Principle (SAC)

The Stochastic Action Principle (SAP) provides us with a framework whereby the principle of stationary action is equivalent to maximized entropy (Wang, 2007). First hypothesized by Jaynes (1957) [13] Principle of Maximum Entropy, this idea has been built up by various researchers over the years. However, for the benefit of brevity, we will concentrate on formalisms and experimentation provided by Dr. Wang's Stocastic Action Principle derived from Shannon entropy, and Dr. Yahalom's (2022) work starting from the Fisher Information Perspective. To start we will have path information encoded as Shannon entropy in relation to the path probability distribution. Specifically, rather than looking at a singular path we instead look at the average path taken by a large number of particles:

$$S(\Gamma) = -\Sigma_\gamma P(\Gamma) ln P(\Gamma)$$

The Action Value measured via the Action Functional:

$$A[\Gamma] = \int_{t_a}^{t_b} L(x(t), \dot{x}(t), t) dt$$

then determine the path (Γ) by maximizing entropy subject to system knowledge constraints, average action:

$$< A >= \Sigma_\Gamma P(\Gamma) A[\Gamma]$$

The entropy maximization framework is seen via normalization, with fixed <A>:

$$\delta(S - \eta \Sigma_\Gamma P(\Gamma) A[\Gamma] - \alpha \Sigma_\Gamma p[\Gamma]) = 0$$

The Lagrangian for variation, functional J extremized:

$$J = -S + \alpha(\Sigma_\Gamma P(\Gamma) - 1) + \eta(\Sigma_\Gamma P(\Gamma) A[\Gamma] - < A >)$$

Take variation with respect $P(\Gamma)$ for small variation.

$$\frac{\partial J}{\partial P(\Gamma)} = 0 => -(ln P(\Gamma) + 1 + \alpha + \eta A[\Gamma] = 0$$

or

$$ln P(\Gamma) = -(1 + \alpha) - \eta A[\Gamma]$$

and $1 + \alpha$ is constant so exponentiation both sides gets.

$$P(\gamma) = \frac{1}{Q} e^{-\eta A[\Gamma]} \propto e^{-\eta A[\Gamma]}$$

This is the final path probability distribution, more precisely, the probability of a path decays exponentially with its action deviation. In discrete notation for paths k from a to b:

$$p_{ab}(k) = \frac{1}{Q} e^{-\eta A_{ab}(k)}$$

With extremal condition:

$$ln P(\Gamma) + \eta A[\Gamma] = const$$

This shows balance of maximizing entropy and minimizing action cost term i.e., "Maximum Path Entropy Principle". These results were showing in computational experiment by Wang (2007) showing

a Gaussian cloud of dust particles travel through paths which obey this principle. They also find Shannon entropy:

$$S_{ab} = -\int D(r)p_{ab}lnp_{ab}$$

which mimics thermodynamics $dQ = dU + dW$, and $U = \bar{E} = \Sigma_i p_i E_i$. Finally we find:

$$\partial^2 p_{ab} = -\eta p_{ab}\partial^2 A, \eta = \frac{1}{2md} > 0$$

then $\partial p_{ab} = 0$ so $\partial A = 0$ so $\partial^2 A > 0$ or in other words, the least action path is the most probable path [14]. When we have vanishing noise, stochastic dynamics simply tend towards regular dynamics:

$$S_{ab} = lnZ + \eta\bar{A}$$

implying $\partial A = 0$, our stationary action principle, a perspective supported by maximum entropy methods in statistical physics [19].

2.2. Recovering Classical and Quantum Mechanics from SAC and Fisher Information

To recover Classical Mechanics is quite trivial. In the limit where the parameter η is large (small fluctuations or low temperature in statistical mechanics analogy), the weight $e^{-\eta A[\Gamma]}$ becomes extremely small unless $A[\Gamma]$ is at its minimum. This is analogous to the Feynman Path integral construction which for classical systems yields the Euler-Lagrange equations and therefore all of Classical Mechanics.

For Quantum Mechanics this is less straightforward, as the path integral weighs each path by a complex exponential:

$$A[\Gamma] \propto e^{\frac{i}{\hbar}A[\Gamma]}$$

For this we must allow a non-vanishing "spread" of paths, by modifying maximum entropy approach to work with complex probability amplitudes rather than probabilities in an entropic dynamics framework [5]. By incorporating interference between paths, the resulting theory reproduces the quantum dynamics governed by the Schrodinger equation. In effect, classical principle of least action is replaced by a principle where all paths contribute, weighted by $e^{\frac{i}{\hbar}A[\Gamma]}$

An alternative construction has been shown to lead to the quantum Hamilton-Jacobi equation and ultimately the Schrodinger equation reproducing quantum dynamics as shown in [21] and then Dirac Equation [28].

2.3. Extending to Fisher Information Geometry

Now we will briefly review some concepts with Fisher Information Geometry which will be required for our later analysis. These are, that information geometry has a dualistic structure. In particular, two conjugate torsion-free affine connections coupled with the metric. Dual parallel transport that is metric-compatable, 2^k types of k-gons. Compared with Riemannian manifolds which exhibit opposite holonomies (net rotation of vector) i.e., Einstein Field Symmetry/Geodesic congruence. These features are not necessarily true for information manifolds when: duel affine connections exist such as (e-connect, m-connect), connections with torsion, non-metric connections (Weyl), or non-Abelian gauge field or spin structures i.e., transporting a spinor.

This paper argues that these unique properties of an information manifold actually allow us to better model spacetime dynamics, particularly Kullback-Liebler Divergence which exhibits $D_{KL}(P||Q) \neq D_{KL}(Q||P)$ or more precisely, not necessarily equal.

2.4. Kullback-Liebler Divergence

Divergence between two distributions P and Q is:

$$D_{KL}(P||Q) = \Sigma_x P(x) ln \frac{P(x)}{Q(x)}$$

For appropriate distributions, we can show that available work exists P relative to Q:

$$F_{neq} - Feq = kT D_{KL}(P||Q)$$

In other words the available work exists relative to equilibrium [16]. Here $(Q||P)$ asks how inefficient it is to use P to model Q. We can easily see that directional imbalance of distributions occurs when:

$$\Delta D = |D(P||Q) - D(Q||P)| = 0$$

It should be intuitive to see this is a maximal entropy state, there are more microstates available when $D(P||Q) = D(Q||P)$, this is already well established as a consequence of:

$$D(P||u) = \Sigma_{x \epsilon X} P(x) log \frac{P(x)}{u(x)} = log|X| - H(P)$$

This is anagolous to Ficks Normal Diffusion (proportional to negative concentration gradient). For any process that distorts distributions P or Q, where P and Q are dependent distributions then there is a tendency to maximize entropy (i.e., moving from high density to low density states). This emerges from balanced Log Ration Condition:

$$ln \frac{P(x)}{Q(x)} so \int (p(x) - Q(x)) ln \frac{P(x)}{Q(x)} = 0$$

2.5. Computing Divergence in identifiable Distributions

There is a rich body of literature which shows various distributions, have this property. However, for this paper the Author ran a fairly trivial computational experiment to verify these distributions. The resulting distributions which produce $\Delta D = 0$ all have symmetrical properties which can be directly related to invarience (conservation under translation). From a range of distributions analyzed, below are some notable distributions whereby $\Delta D = 0$ (and properties to note):

- Gaussian distributions (same covariance, frame of reference invariance, related to time symmetry),
- location shift families (integrated depends on $\theta_1 - \theta_2$, related to translation symmetry),
- Caunchy distributions of same family (same scale, Reflection, Scale Invariant, some Mobius transformations),
- Antipodal or Permuted Discrete Distributions (mirror image or group rotation, rotational symmetry),
- Von Mises-Fisher (preferred direction spin up/down, symmetric under mean direction but not concentration),
- Bregman Divergence (exponential family distributions are symmetric when "distance" d is essentially Euclidean, implying no extra gauge or affine structure, also implies connection used is usual (Levi-Civita) (see, e.g., Amari and Nagaoka, 2000 [1])).

This should not be too surprising, the distributions on an information manifold can be viewed essentially as field states.

2.6. Decomposing Observed Distributions

The obvious next step is to examine an observed distribution signal, which may appear random or chaotic can actually be represented by a series of decomposable symmetric distributions. Given this,

we look towards methods to linearly separate a mixed distribution into their component distribution parts. This again is fairly straightforward using Fourier analysis and signals analysis techniques. In particular, we can find the marginal likelihood for a given model M_i to be a fit within a mixed solution from:

$$p(y|M_i) = \int p(y|\theta_i, M_i)p(\theta_i|M_i)d\theta_i$$

The posterior probability for model M_i from Bayes Rule:

$$p(M_i|y) = \frac{p(y|M_i)p(M_i)}{p(y)}$$

and from Bayes Factor (see, e.g., Kass and Raftery, 1995 [15]):

$$BF_{ij} = \frac{p(y|M_i)}{p(y|M_j)}$$

Alternative checks for likelihood, for one distribution to be the "correct" decomposition can be found using AIC/BIC or Penalized Likelihood/MAP estimation instead or other techniques. The important thing in this analysis is to find a high degree of confidence that a given signal is decomposed correctly into its sub-signals, and if uncertainty exists between configurations this must be factored into analysis. To better avoid this, we can determine boundary conditions based on physical law. Further confidence is increased when evaluating the time evolution of the distribution as each composable signal has a unique time evolution signature when it is perturbed and drifts back to its idealized state.

2.7. Complex Components (Phase Space Projection)

If we have a distribution that exists in phase space, but its distribution can only be observed in R^3, then we inevitably get complex structure and unpredictability. The result of this is a large variation in the probability "certainty" of the linear separation. In other words, a single smooth distribution may appear in R^3 as having multiple modes. If we are optimizing for R^3 space then $D(Q||P)$ again will exhibit mode-seeking behavior, where under time evolution the divergence of $(Q||P)$ will re-center itself around one mode which acts pseudo "orthogonal" to the leftover modes. This mode-seeking behavior is conceptually very similar to super-position of states. Selecting one distribution automatically causes cascading effects into all remaining distribution patterns, mimicking wave function collapse. Similarly, we can only capture one mode at a time, so the higher the convergence on one mode, the less information and therefore higher uncertainty we have about the remaining, unselected mode. For a complex function

$$f(z) = u(x_\mu) + iv(x_\mu)$$

Projected onto the real axis we define projection:

$$g(x_\mu) = Re(f(x_\mu)), h(x_\mu) = Im(f(x_\mu))$$

Complex function is simply a sum of exponentials, real projection exhibit interference patterns:

$$f(z) = \Sigma_n e^{ik_n z}, Re(f(z)) = \Sigma_n cos(k_n x)$$

The complex Fourier modes is simply:

$$f(z) = \Sigma_n A_n e^{ik_n z}$$

An interesting pattern emerges whereby for any complex distribution, the real distribution must have its most likely decomposition to include multiple sub-distributions. When this pattern is renormalized,

the distribution not captured by mode optimization is free for further signal decomposition. The reader may note this is very similar to the way the Standard Model has essentially been constructed:

$$\mathcal{L}_{SM} = \mathcal{L}_{gauge} + \mathcal{L}_{fermion} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}$$

Where each Lagrangian term represents its own path and action minimization and therefore the fields can be written in terms of a path distribution. Conceptually we can think of this process in terms of entropic maximization tends towards phase alignment

2.8. Path Dependence

Instead of just a simple distribution of point P(x), the probability of being in state x. We can consider the probability of observing a particular path $P(\Gamma)$, where Γ can be either discrete steps in a path or a continuous time evolution. Therefore, path-dependent divergence is:

$$D_{KL}(P||Q) = \int \daleth \Gamma P[\Gamma] ln \frac{P[\Gamma]}{Q[\Gamma]}$$

This tells us how surprised we would be for a path to deviate from assumed dynamics Q, rather than the observed dynamics P. Entropy production relative to the ratio of forwards and backwards path probabilities, total entropy production is given by the KL divergence [22]. Where $[\tilde{\Gamma}]$ is the time-reversed trajectory:

$$\Delta S_{tot}[\Gamma] = ln \frac{P[\Gamma]}{Q[\tilde{\Gamma}]}$$

And the average over all trajectories in a stochastic system is:

$$\Delta S_{tot} = D_{KL}(P[\Gamma]||Q[\tilde{\Gamma}])$$

This average KL Divergence quantifies total entropy production for traveling a path, giving us the irreversibility of a process/dynamic (cf. Seifert, 2012 [23]). This in "reverse" $D(Q[\Gamma]||P[\Gamma])$ quantifies the error in observed model when trying to capture idealized measures. This is akin to saying what is the system inefficiency or signal loss.

2.9. Identifying Multi-Path Dependence

The dual affine structure of Information Manifolds provides the perfect grounding for evaluating processes with multiple separable path-dependent processes. It is simple to find examples, an object placed in front of a fluid flow will experience path-dependency from the geometry of the space (assuming constant curvature), while the displacement of fluid flow due to relative differences in inertia result fluid being forced around the object in turbulence. This is consistent with the thermodynamic metrics approach [24]. In this, not only is the object affected by the flow geometry, but the geometry itself is affected by the presence of the object. If we remove this object, its impacts can still be seen until equilibrium is re-established and normal, unimpeded flow returns to an equilibrium. In this, total divergence can be separated by Leibniz rule or additivity of KL divergence, so we can decompose these two separate but related path-dependent processes into:

$$D_{KL}(P||Q) = D_{KL}(P||P_1) + D_{KL}(P_1||Q) + \Sigma_x[P(x) - P_1(x)]ln\frac{P_1(x)}{Q(x)}$$

if $P_1(x)/Q(x)$ is constant i.e., when $P(x) \neq P_1(x)$, $P_1(x) = 0$ or if $P(x) = P_1(x)$ then this cross term vanishes [7]), which is the case if an observed distribution can be perfectly modeled by two separate path models. As such we have the result where we can relate baseline distribution Q, final

distribution P and an intermediate distribution P_1. As such for a distribution comprised only of symmetric distributions:

$$D_{KL}(P||Q) = D_{KL}(P_1||Q) + D_{KL}(P||P_1) = D_1 + D_2 = 0$$

This is very powerful, without requiring any information about a missing path process, we can say that we can measure the energy gap between a theoretical model, and our observed distribution. From this we can gain insight or even potentially recover the path information which will give rise to the underlying mechanisms that our theoretical models do not currently capture. The reader might pause here because D_{KL} can be related to entropy, and in this time evolution we have $\partial D_1 = -\partial D_2$ this seams unphysical, we should not be seeing a negative path derivative of entropy. However, it is already established that entanglement and coherence of subsystems is invariant under ideal Quantum Reference Frame transformations (Cepollaro et al., 2024) [6]:

$$E^A + C^A = E^B + C^B, \Delta E \Delta C \leq 0$$

Additionally, Vopson & Lepadau (2022) found what they coined "the Second Law of Infodynamics"[25], where under repeated observation and experimentation, across vastly different systems they observed an informational entropy S_{inf} with exactly this property, in particular, that $\partial S_{trad} = -\partial S_{inf}$. These two results one would argue exactly meets the condition set for $\partial D_1 = -\partial D_2$ as an entropic maximum.

3. Results

It should be clear that, $D_{KL}(P||Q) = 0$ occurs when our theoretical model matches observation, i.e., when an ideal physical model is known. $D_{KL}(P_1||Q)$ is how much the ideal theoretical model is mismatched by our current model. $D_{KL}(P||P_1)$ tells us how much the amount of work required to bring a universe following our best current model to match our observed universe. In this idealized case, $D_{KL}(P||Q) = 0$, the following must result $D(P_1||Q) = -D(P||P_1)$. This can be described as a decomposition of two separate, but additive types of path dependence (i.e., geometric curvature and drag), but ultimately, knowing the specific underlying mechanism is irrelevant. This is because the optimized feature is an energy deficit from our expected distribution from our models and the observed universe. Lastly, it is well known that Ricci Flow follows as a Lie derivative with respect to a perturbed state. The Linearized operator governing evolution of h(t) has only smooth decaying modes, consistent with the below work. We attempt to reconcile this observational deficit by modeling it as a function of Ricci flow as a spacetime reparative process ensuring a smooth manifold. **3.1 Ricci Flow as Repair** The geometric evolution for a Riemannian metric, the unnormalized flow is given by:

$$\frac{\partial}{\partial t} g_{ij} = -2Ric_{ij}$$

We can examine the rate of "repair" compared to average flow, this is examined later as a function of temperature and density. Setting aside this for now we know the DeTurck Trick, whereby weakly parabolic (freedom to change coords) adds a diffeomorphism term to flow, leading us to a strickly parabolic system:

$$\frac{\partial}{\partial t} g_{ij} = -2Ric_{ij} + \mathcal{L}_V g_{ij}$$

In this, the Lie derivative with respect to a perturbed field V, the linearized operator governing the evolution of h(t) only has decaying modes. This ensures that any perturbation dies out over time in a smooth and exponential rate relative to neighbouring states [18]. In many cases the linearized analysis shows that each mode of the perturbation decays independently and exponentially as well.

This is to say in the language of this framework that if h(t) can be decomposed into eigenmodes of the linearaized operator, then each mode decays like:

$$e^{-\lambda t}$$

with $\lambda > 0$ as t approaches infinity and the metric returns (up to a diffeomorphism) to g_0. It is important to note however that there may exist upper and lower bounds of decay rate. If this is the case this signature should relate to a rate of information transfer (perhaps quantum speed limit) and be extracted from examining signatures upon decomposition.

3.1. Black Holes

To start, we know that black holes can be described in three separate ways consistent with the four laws of black hole mechanics [2]. First, in terms of their event horizon; Second, in terms of the singularity; Third, in terms of the gradient curvature of spacetime between regions. This paper argues that solely examining a black hole from only one of these frameworks is an insufficient description. As such we will relate spacetime compression of states to the hypothesized upper rate limit of information transfer (quantum speed limit) as a function of an upper boundary of decay. By doing so we have an emergent boundary relational to the Bousso or Bekenstein Bound for a given observer's reference frame. Specifically, for an outside observer, you end up with an energy density which approaches infinity at the singularity, but with maximum entropy boundaries across a finite surface area dependent on the radius of the event horizon. Each of these entropic shells have their own energy distribution as given by Helmholtz Free Energy equations. As the radius of a black hole is directly proportional to its mass, which is then proportional to the number of maximally defined states (atomic states) we can understand the reason why smaller black holes evaporate faster if entropy flow dominates energy dissipation. This is given because if $R_s \sim r$ then higher-order multipoles lead to fields dissipating far faster. However, for $R_s \gg r$ it would appear as a point source and the inverse-square law would dominate.

As such we have two reliable boundary conditions for entropy, particularly as it relates to curvature. Specifically, the maximum entropy state, or curvature state that blocks photon escape (i.e., Bousso, Bekenstein Bounds at event horizon). Assuming a stable system, all free energy approaching the event horizon is drawn into the center of the black hole rapidly until maximum entropy density is reached (finite or infinite), if the slope of the gradient of curvature is continuous between these two points, the rate of black hole decay should directly relate to differences in energy states between curvature steps. This is consistent with our expectation that Ricci Flow is the potential description for geometric repair as exponential decay. For a given system we have:

$$S(n) = \left(\frac{1}{C}\right) S(n+1)$$

which can be expressed by the curve:

$$S(x) = S_0 \, e^{-x \ln C}$$

However, because the bounding values of curvature are predefined (escape curvature of photon and maximum relative curvature between neighbour states), the only varying parameter is the radius of the black hole event horizon, and therefore we can write that the change in field decay is moderated by the scaling factor k (a function itself of r):

$$C' = \frac{C}{k}, \quad C'k = C$$

where C represents the maximum possible field decay. Again, we can use the Von Neumann entropy S of a quantum state ρ given by:

$$S(\rho) = -\text{Tr}(\rho \log \rho)$$

If we consider two states then the amount of entropic work performed is related to the difference in their entropies and respective energies.

$$F(\rho) = E(\rho) - TS(\rho)$$

What is interesting here is the equivalence of $TS(\rho)$ to energy, indicating that given the relatively homogeneous temperature within a stable black hole, the change of temperature relative to energy is $dT/dE \propto 1/E^2$; therefore entropy changes are equivalent to changes in energy at a first-order approximation. This should not be too surprising as a commonly defined macroscopic description of temperature as a comparison of two heat transferring states (V, N) is:

$$T := \left(\frac{dU}{dS}\right)_{V,N}$$

In this, there are some specific consequences particularly as it relates to the thermodynamic limit for considering temperature in an information context. This will be examined in a later paper, so for now we can focus on the Work extracted which is related to the difference in free energies:

$$W = F(\rho_1) - F(\rho_2)$$

This can instead be looked at directly from the perspective that a change in entropy is a change in energy state. Specifically, decreasing the entropy increases free energy, which can then propagate away. Along the event horizon, very high entropy inhibits the release of photons, as this entropy dissipates away. In this case, even if $E(\rho)$ within the event horizon (outside singularity) falls to zero, i.e., in the state where all mass/energy falls towards the center, there still exists a flow energy resulting from the flow of entropy. This should make sense as the entropy flowing away from the singularity is equal to the entropy flowing away from the event horizon. In this, the flow of entropy is what makes black holes evaporate as vacuum entropy density is significantly lower. This requires its own gradient resulting in the fluid-like nature of the surface of a black hole. We can look at the rate of volumetric contraction of the black hole as it relates to the loss rate. The ratio of the rate of entropy loss S as it relates to field decay with the respective volume change defined by the total entropy/energy within the sphere.

$$\frac{\frac{TdS}{dt}}{\frac{dV}{dt}} = \frac{dR}{dV} = \frac{3Cr}{4\pi r^3} = \frac{3C}{4\pi r^2}$$

The Schwarzschild radius is given by Black hole thermodynamics was first formulated by Bekenstein (1973) [3] and quantum evaporation by Hawking (1975) [10]:

$$R_s = \frac{2GM}{c^2}$$

Which we can rearrange to convert back into mass:

$$\frac{1}{R_s} = \frac{c^2}{2GM}$$

This gives us the rate of entropy change as a ratio of total volume change (curvature change):

$$\frac{TdS}{dV} = \frac{3Cc^4}{16G^2M^2}$$

The total entropy of a black hole is given by the Bekenstein-Hawking formula and is proportional to the area of the event horizon:

$$S = \frac{k_B c^3}{\hbar G} \frac{A}{4}$$

For which we find the mass loss rate derived from the power equation:

$$\frac{dM}{dt} \sim -\frac{\hbar c^4}{G^2 M^2}$$

And we can factor out the mass loss:

$$\frac{TdS}{dV} = \frac{dM}{dt}\left(\frac{3C}{16\hbar G}\right)$$

Since the rate of entropy loss is the same as the rate of mass loss (energy/mass equivalence) we now have a cosmological constant for the maximum field decay rate between two states.

$$C = \frac{16\hbar G}{3}$$

where $\mathbb{C}_3 = \frac{V}{r^3} = \frac{4\pi}{3}$ is a generalization for an *n*-dimensional sphere of 3 spatial dimensions. The generalized equation for the coefficient of an *n*-dimensional sphere (n-ball) with radius *r* is:

$$C_{n(\text{dimensions})} = \frac{V_n(r)}{r^n} = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}$$

So, for a spacetime with 3 spatial dimensions, we have:

$$C = 4\mathbb{C}_3 hG$$

This can be factored back into the exponential decay equation to get the mass decay rate due to entropy dilution.

Similarly, we can replace mass conversion here with energy to get:

$$\frac{TdS}{dV} = \frac{dE_t}{dt}$$

bounded by

$$C^u\left(\frac{m^3}{s}\right)$$

where adjusting C to relativistic SI units:

$$C^u = \frac{C}{c^2} = \frac{4\mathbb{C}_3 hG}{c^2}$$

This is expressed in terms of SI units (m^3/s). Specifically, it appears to be a maximum expansion rate on the order of one Planck length per Planck time which is a startling, if not intuitive result. Additionally, in evaluating time as proper time τ we can see the rearrangement of the equation to the form of:

$$\frac{dt}{d\tau}\frac{TdS}{dE_t} = \gamma\frac{TdS}{dE_t} = \frac{dV}{d\tau}$$

expansion scalar change over time is respective to the change of entropy in respect to the energy of the system:

$$\frac{TdS}{dE_t} = \theta$$

with C_e as the maximum energy relaxation relative to the observer i.e.,

$$S(x) = S_0\, e^{-x\ln C_e}$$

Or in systems with little or no free energy:

$$\frac{dt}{d\tau}\left(\frac{TdS}{dE_t}\right) = \frac{d\theta}{d\tau}$$

Formulated relativistically with the Lorentz factor:

$$\gamma\frac{TdS}{dE} = \frac{d\theta}{d\tau}$$

It is important to note however that this Lorentz factor does cancel out, given that the scaling factor θ is measured with its own Lorentz factor, giving us a truly covariant solution. This is interesting because S and E can both be described in terms of state functions which can be expanded upon from the completed work in Section 2.3:

$$\frac{S}{E_t} = \frac{-Tr(\rho log(\rho))}{Tr(\rho\hat{H})}$$

And if ρ is thermal then $\rho = -e^{-\beta H}/Z$: $\lambda_i = e^{-\beta E_i}/Z$ where E_i is the energy of the i-th eigenstate of H and:

$$T = \frac{1}{k_B\beta}$$

Or in other words, the change in symmetry compression is relative to the observer. Specifically, we can see dE as a change to the states' eigenvalues and dS as a change to the states' distribution, particularly as it relates to normalization. This implies that a component of curvature relaxation directly relates to the emergence of symmetries in the state function path. Again, it is obvious how this relates back to signals processing theory and in particular data modulation.

Finally, the slope of the entropy flow determined by the radius of the black hole can be related to fluid dynamics where back pressure is directly related to flow in a classical system and is characterized by the slope of the fluid height profile:

$$P(x) = \rho g(h_0 - Sx)$$

In this sense, it takes more work to change the flow of a system with high back pressure. This relates directly to viscosity along the event horizon. A slower entropy decay due to high back pressure would directly appear as a fluid with high viscosity. This directly matches the observed phenomena of fluid-like black hole boundaries. Further calculation is required to get a purely quantum entropy description for this back pressure.

3.2. Expansion Scalar Examination

The expansion scalar θ is used to denote fluid-like flow for an observer. It is defined as the covariant divergence of the four-velocity field:

$$\theta = \nabla_\mu u^\mu$$

Where ∇_μ is the covariant derivative associated with the spacetime metric $g_{\mu\nu}$ and u^μ is the four-velocity field of observers or fluid elements. This expansion scalar is interpreted as expanding or diverging worldlines when $\theta > 0$, contracting or convergent worldlines when $\theta < 0$, and static when equal to 0. Alternatively, we find:

$$\frac{d\theta}{d\tau} = u^\mu\nabla_\mu\theta = -\frac{1}{3}\theta^2 - u_{\mu\nu}u^{\mu\nu} + \omega_{\mu\nu}\omega^{\mu\nu} - R_{\mu\nu}u^\mu u^\nu$$

This scalar is regularly defined as an infinitesimal volume element dV moving with the flow with observers' proper time (see Raychaudhuri, 1955 [20]; Hawking and Ellis, 1973 [11]) such that:

$$\frac{d\theta}{d\tau} = -\frac{1}{3}\theta^2 + \Lambda$$

In other words θ is the fractional rate of volume change per unit of proper time. For homogeneous and isotropic universes, like as described by the Friedman-Lemaitre-Robertson-Walker (FLRW) metric, the expansion scalar is related to the Hubble parameter:

$$\theta = 3H(t)$$

The behavior of θ is crucial in the Penrose-Hawking singularity theorems, with highly negative values leading to singularities, this highlights this derived result acting in congruence with the above formulation. Further, for examining tidal forces and its impact on expansion, shear and rotation congruence we can examine the evolution of the expansion scalar along the flow as governed by the Raychaudhuri equation:

$$\frac{d\theta}{d\tau} = -\frac{1}{3}\theta^2 - \sigma_{\mu v}\sigma^{\mu v} + w_{\mu v}w^{\mu v} - R_{\mu v}u^\mu u^v + a_\mu a^\mu$$

Implying a deep thermodynamic nature of gravity [12]. This above formulation should give ample evidence of the relationship between entropy, energy, and the expansion scalar. In particular, it should become apparent to the reader how the expansion scalar is directly related to dark energy. In particular how the cosmological constant Λ contributes to the Ricci tensor $R_{\mu v}$ such that:

$$R_{\mu v} = 4\pi G g_{\mu v}(\rho - 3p) + \Lambda g_{\mu v}$$

From the first Friedman Equation (energy density equation) we can also see:

$$H(t)^2 = \frac{8\pi G}{3}\rho(t) - \frac{k}{a(t)^2} + \frac{\Lambda}{3}$$

and the second Friedman Equation (acceleration equation) which the Λ term drives observed accelerated expansion:

$$\frac{\frac{d^2a(t)}{dt^2}}{a(t)} = -\frac{4\pi G}{3}(\rho(t) + 3p(t)) + \frac{\Lambda}{3}$$

Also important to note that dark energy has equation state $w = -1$ which means it has negative pressure, whereby we denote pressure as having three times more effect on dynamics than energy density (hence 3p(t)).

If the overall system's expansion scalar approaches a constant value, this reflects a state of continuous accelerated expansion otherwise seen, or more accurately previously modeled as de-Sitter expansion.

3.3. Special Conditions

This derivation results in some obvious special conditions, primarily when a system has constant entropy and energy i.e., conservation of energy and entropy. In this case, we get:

$$dt = dV \text{ bounded by } C_e$$

This should make intuitive sense, as time increases, the total volume enclosed in the light cone of all interactions observed within it (system evolution) increases as well. Additionally, we have:

$$dE = dS\,dt \text{ bounded by } C_e$$

Where a local change of entropy (perturbation of a field) with discrete time steps. This describes a system of two overlapping light cones that are contained in another larger "older" light cone. As time increases, the entropy of the system increases, as such the change of energy contained within curvature also increases. Spacetime contraction requires some quanta of energy, for photons propagating this would look like a redshift as wavelength increases. It is important to note here that this phenomenon would relate directly to the gradient of neighboring states' entropy, as a photon's energy would be transferred into the vacuum state if the two states' normalization is not congruent (i.e., the scalar deformation) consistent with a unitary interaction. On a universal scale, this should on average appear as a linear relationship with distance traveled, with only local deformations resulting from regions with highly similar entropic states or regions where photons gain energy from vacuum states with higher potential. Unlike other "tired light" mechanisms, as this process is based on the relative energy or entropy of the neighboring states that a particle transits through, rather than the particle's energy itself, it should not be wavelength-dependent. As such, this process should not exhibit blurring which is a categorical problem with other "light scattering" or "tired light" models.

3.4. Temperature Regulation and Dark Matter

Another consequence of this formulation is that temperature evaluated between two points, $T = U(x)/S(x)$ is a macro, statistical description of average energy transmission through spacetime (i.e., average momentum of particles) may mediate the flow of the expansion scalar.

In systems with little or no relative free energy, with Von Neuman entropy, on can determine from the eigenvalue populations in the density matrix where:

$$T = \frac{1}{k_B \beta}, \frac{dt}{d\tau} \frac{TdS}{dE} = \frac{d\theta}{d\tau} = \frac{dt}{d\tau} \frac{Td(-Tr(\rho log(\rho)))}{dTr(\rho \hat{H})}$$

and:

$$\rho = \Sigma_{i=1}^{n} |\lambda_i |v_i\rangle \langle v_i|$$

so if ρ is thermal, $\rho = -e^{-\beta H}/Z$ and $\lambda_i = e^{-\beta E_i}/Z$ where E_i is energy of the i-th eigenstate of H.

In other words, average interaction frequency relates to higher rate of unitary interactions which affect the rate that a systems entropy is updated. Another way of thinking about this is it is the degree that a systems internal movement resists curvature relaxation. Intuitively this makes sense, if we repetitively block flow in a river downstream, the geometry of the flow retains deformation proportional to the frequency of blockage and the rate of flow repair. If we look back to Section 2.2, 2.3 we see that a series of unitary operations occur when path evaluation (curvature evaluation) occurs, in other words, a particle transits through a region. Any transiting particle, such as a photon emitted from a source, would cause interactions resulting in the updating of the states normalization, therefore entropy and the curvature of spacetime. Conversely, in very low-temperature settings, fewer and lower energy particles transit through the medium resulting in less impactful changes to curvature. This may give us insight into why Dark Matter halos are predominantly viewed at the edge of galaxies (see, e.g., Bullock and Boylan-Kolchin, 2017 [4]), as dense regions around star formation tend to have higher temperatures and higher average flux density. As such, We can view the above boundary condition as a measure of the proportion of the free energy and locked entropic energy within a system, in other words:

$$T \propto \frac{dE_f}{dE_l}$$

In other words, the higher the temperature, the more dominant free energy will be within the system compared to locked energy, i.e., entropically defined energy. A direct examination of this ratio will be conducted in a later paper. Until then the reader can ponder if the velocity of the leading edge of gravitational waves, and thus gravitational energy density can be directly affected by temperature as it acts on gravitational wave flux density.

3.5. Formatting of Mathematical Components

This is the example 1 of equation:

$$a = 1, \tag{1}$$

4. Discussion

The relationship between quantized spacetime and entropy should be clear. Entropic work is performed by the observation or translation of any particle. For simple particles without complex spin symmetries, this time evolution is forced into gradient dispersion as a direct consequence of the projection of a particle its prior state location influencing its current state location. That entropy is further increased requiring more work to be performed when the neighboring state is dissimilar. Since curvature increases with increased energy, this provides a mechanism for the experience of smooth spacetime curvature as spacetime contracts or expands to become more homogeneous with its surrounding states. It is important to note here that a relaxing volume necessitates its neighbor volume to contract as a particle with a given energy is essentially annihilated and then created in the new volume because of the field propagation. If contraction and expansion of a volume occur non-linearly (such as in General Relativity), a macro-observer in a stable system would observe a change of relative volumes dramatically smaller than the lower bound of the non-zero minimum length. This would make observing this mechanism only possible under extreme circumstances where the gravitational field is violently disturbed. Also note that under this formulation gravitational fields are the combinatory field of all energies in all existing fields, I find no reason to assume that this entropic process prefers energy propagation in any particular mode.

This framework gives us a time evolution description for curvature and a relativistic energy density which relates directly to entropy and increasing system symmetries. This description translates directly to expected results for black holes and dark energy. In particular, we expect dark matter density within the halo to fall off proportional to the cube of radius at large distances. Additionally, we can view the expansion of our observable universe to be an expansion of our light cone whereby low entropy states are being added to our relativistic worldview. By using theory which shows that entanglement information is inversely related to coherence information, also seen in information dynamics (Vopson, 2022) you can derive a relationship between the gradient of entropy. Specifically, one can find an entropic relationship between entanglement and coherence that is in covariant and contravariant forms of a new metric tensor. Additionally, we can find that curvature can be described here as worldlines expressing harmonized modes of information where the real and complex components of the eigenvalues are. These scalar multiples relate directly to modes of force that correspond to the 2D plane delta curvature (as compared to Riemann Flat space). This geometric analysis reveals interesting results given the entropically defined time evolution boundary conditions.

Limitations

Several key limitations must be considered within this paper. First, in deriving the relationship between the change of entropy and energy we assumed stable, non-changing temperature (hence from product rule $SdT = 0$. This cannot be assumed for the vast majority of applications, however, for black holes $dT/dE \propto -1/E^2$ and as such should not play a dominant role, particularly on a first-order approximation. Further work should be conducted in fully incorporating dT as simplifications are highly likely. Secondly, we constructed this formulation at extreme boundaries near a singularity, where mass is pulled towards the center. As such we assumed that the entropic component of total energy was dominant for the region in between the singularity and the event horizon. In systems with elevated levels of free energy, this component must be considered for its impact on curvature deformation according to the stress-energy tensor.

More significantly, is the potential for an alternative decomposition which may better "fit" the distribution divergences. While this author attempted to ground the decomposition in sound logic and

fundamental principles (least action, geometric interpretation) one cannot ignore the fact that much of these ideas are derived from approximation and assumption from best fit.

5. Conclusions

In conclusion, our investigation challenges the traditional path-independent assumptions of tensor contraction in spacetime, revealing that significant information is lost in this simplification. The evidence points to the necessity of path-dependent analyses to account for the entropic dynamics inherent in spacetime, particularly at quantum scales. This realization opens new avenues for research, particularly in the examination of the Riemann Tensor with an emphasis on preserving information and embracing the entropic nature of spacetime evolution. Future work should explore these entropic path-dependent processes in greater detail, aiming to integrate them with existing theoretical frameworks to enhance our understanding of spacetime dynamics. This approach not only promises to refine our grasp of General Relativity also bridges the gap between classical and quantum physics by acknowledging the fundamental role of information entropy in shaping the universe.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Amari, S., & Nagaoka, H. (2000). *Methods of information geometry*. Oxford University Press. https://doi.org/10.1090/mmono/191

2. Bardeen, J. M., Carter, B., & Hawking, S. W. (1973). The four laws of black hole mechanics. *Communications in Mathematical Physics, 31*(2), 161–170. https://doi.org/10.1007/BF01645742

3. Bekenstein, J. D. (1973). Black holes and entropy. *Physical Review D, 7*(8), 2333–2346. https://doi.org/10.1103/PhysRevD.7.2333

4. Bullock, J. S., & Boylan-Kolchin, M. (2017). Small-scale challenges to the ΛCDM paradigm. *Annual Review of Astronomy and Astrophysics, 55*(1), 343–387. https://doi.org/10.1146/annurev-astro-091916-055313

5. Caticha, A. (2011). Entropic dynamics, time, and quantum theory. *Journal of Physics A: Mathematical and Theoretical, 44*(22), 225303. https://doi.org/10.1088/1751-8113/44/22/225303

6. Cepollaro, C., Akil, A., Cieśliński, P., de la Hamette, A.-C., & Brukner, Č. (2024). The sum of entanglement and subsystem coherence is invariant under quantum reference frame transformations. *arXiv preprint arXiv:2406.19448*. https://doi.org/10.48550/arXiv.2406.19448

7. Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory* (2nd ed.). Wiley-Interscience. https://doi.org/10.1002/047174882X

8. Friedan, D. (1985). Nonlinear models in $2 + \varepsilon$ dimensions. *Annals of Physics, 163*(2), 318–419. https://doi.org/10.1016/0003-4916(85)90384-7

9. Hatano, T., & Sasa, S.-i. (2001). Steady-state thermodynamics of Langevin systems. *Physical Review Letters, 86*(16), 3463–3466. https://doi.org/10.1103/PhysRevLett.86.3463

10. Hawking, S. W. (1975). Particle creation by black holes. *Communications in Mathematical Physics, 43*(3), 199–220. https://doi.org/10.1007/BF02345020

11. Hawking, S. W., & Ellis, G. F. R. (1973). *The large scale structure of space-time.* Cambridge University Press. https://doi.org/10.1017/CBO9780511524646

12. Jacobson, T. (1995). Thermodynamics of spacetime: The Einstein equation of state. *Physical Review Letters, 75*(7), 1260–1263. https://doi.org/10.1103/PhysRevLett.75.1260

13. Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review, 106*(4), 620–630. https://doi.org/10.1103/PhysRev.106.620

14. Kaila, V. R. I., & Annila, A. (2008). Natural selection for least action. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 464*(2099), 3055–3070. https://doi.org/10.1098/rspa.2008.0178

15. Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

16. Kawai, R., Parrondo, J. M. R., & Van den Broeck, C. (2007). Dissipation: The phase-space perspective. *Physical Review Letters, 98*(8), 080602. https://doi.org/10.1103/PhysRevLett.98.080602

17. Matsueda, H. (2013). Emergent general relativity from Fisher information metric. *arXiv preprint arXiv:1310.1831*. https://doi.org/10.48550/arXiv.1310.1831

18. Perelman, G. (2002). The entropy formula for the Ricci flow and its geometric applications. *arXiv preprint arXiv:math/0211159*. https://doi.org/10.48550/arXiv.math/0211159

19. Pressé, S., Ghosh, K., Lee, J., & Dill, K. A. (2013). Principles of maximum entropy and maximum caliber in statistical physics. *Reviews of Modern Physics, 85*(3), 1115–1141. https://doi.org/10.1103/RevModPhys.85.1115

20. Raychaudhuri, A. K. (1955). Relativistic cosmology. I. *Physical Review, 98*(4), 1123–1126. https://doi.org/10.1103/PhysRev.98.1123

21. Reginatto, M. (1998). Derivation of the equations of nonrelativistic quantum mechanics using the principle of minimum Fisher information. *Physical Review A, 58*(3), 1775–1778. https://doi.org/10.1103/PhysRevA.58.1775

22. Roldán, É., & Parrondo, J. M. R. (2012). Entropy production and Kullback–Leibler divergence between stationary trajectories of discrete systems. *Physical Review E, 85*(3), 031129. https://doi.org/10.1103/PhysRevE.85.031129

23. Seifert, U. (2012). Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics, 75*(12), 126001. https://doi.org/10.1088/0034-4885/75/12/126001

24. Sivak, D. A., & Crooks, G. E. (2012). Thermodynamic metrics and optimal paths. *Physical Review Letters, 108*(19), 190602. https://doi.org/10.1103/PhysRevLett.108.190602

25. Vopson, M. M., & Lepadatu, S. (2022). Second law of information dynamics. *AIP Advances, 12*(7), 075310. https://doi.org/10.1063/5.0100358

26. Wang, Q. A., Tsobnang, F., Bangoup, S., Dzangue, F., Jeatsa, A., & Le Méhauté, A. (2009). Stochastic action principle and maximum entropy. *Chaos, Solitons & Fractals, 40*, 2550–2556. https://doi.org/10.1016/j.chaos.2007.10.047

27. Yahalom, A. (2022). Fisher information perspective of Pauli's electron. *Entropy, 24*(12), 1721. https://doi.org/10.3390/e24121721

28. Yahalom, A. (2024). Dirac equation and Fisher information. *Entropy, 26*(11), 971. https://doi.org/10.3390/e26110971.