Bit threads in higher-curvature gravity

Jonathan Harper,[a] **Matthew Headrick**[a,b] **and Andrew Rolph**[a]

[a]*Martin Fisher School of Physics, Brandeis University,*
Waltham MA 02453, U.S.A.

[b]*Center for Theoretical Physics, Massachusetts Institute of Technology,*
Cambridge, MA 02139, U.S.A.

E-mail: jharper@brandeis.edu, headrick@brandeis.edu,
andrewrolph@brandeis.edu

ABSTRACT: We generalize holographic bit threads to bulk theories with a gravitational action containing higher-curvature terms. Bit threads are a reformulation of holographic entanglement entropy, where the entropy is given by the maximum number of threads emanating from a boundary region into the bulk. We show that the addition of higher-curvature terms adds corrections to the bit thread thickness that depend on the local geometry and thread orientation. Two different methods are given: determination of the density bound by requiring the maximum number of threads through a given surface to reproduce the entanglement entropy functional on that surface, and application of Lagrange dualization. The results of the two methods are applied to Gauss-Bonnet gravity as the simplest non-trivial example.

Contents

1   Introduction

Holographic entanglement entropy (HEE) has had a profound impact on our understanding of quantum gravity, by directly connecting quantum information and geometry. According to the Ryu-Takayanagi (RT) formula [1], the entropy of a boundary region A is given by the area of the minimal-area bulk surface homologous to A:

$$S(A) = \frac{1}{4G_{\mathrm{N}}} \min_{m \sim A} \mathrm{area}(m). \tag{1.1}$$

Bit threads [2] are a reformulation of the RT formula introduced to address some of the conceptual issues arising from the RT formula. To understand them, one first defines a *flow*, a vector field v satisfying $\nabla_\mu v^\mu = 0$, $|v| \leq 1$ everywhere. A consequence of the divergencelessness is that the flux of v through A equals the flux through any surface homologous to A. Maximizing the flux picks out the minimal surface as the "bottleneck", on which the norm bound is saturated, giving a flux equal to the minimal surface area. A

simple analogy is determining the size of a water pipe's bottleneck by maximizing the flow through one end. Mathematically, this is expressed in the max flow-min cut theorem,

$$\min_{m \sim A} \text{area}(m) = \max_v \int_A \sqrt{h}\, n_\mu v^\mu \quad \text{where} \quad \nabla_\mu v^\mu = 0\,, \quad |v| \le 1\,. \tag{1.2}$$

Based on this theorem, we can rewrite the RT formula as

$$S(A) = \frac{1}{4G_N} \max_v \int_A \sqrt{h}\, n_\mu v^\mu\,. \tag{1.3}$$

A bit thread is an integral curve of v. The threads are chosen as a subset of the integral curves with transverse density equal to $|v|/4G_N$; effectively, each thread has cross-sectional area $4G_N$. The fact that threads have thickness and cannot intersect then gives an interpretation to the vector field norm bound: there is a limit to how closely the threads can be packed together. For configurations that maximize the number of threads out of A, the threads are maximally packed on and directionally normal to the minimal surface, which acts as a bottleneck. There is a redundancy in which discrete members in the continuous family of integral curves are given a cross-sectional area and called bit threads. At AdS scales the bit thread number density is of the order N^2, so the discrete family of bit threads is indistinguishable from the continuous family of integral curves, and the terms "bit thread" and "flow" can be used interchangeably. The maximum number of threads that can be placed on A gives the entanglement entropy, as if each bit thread connected an EPR pair of qubits between A and its complement region A^c; bit threads are thus a layer of quantum information theoretic interpretation on top of the vector field v. Unlike the minimal surface, the bit thread configuration changes continuously under continuous deformations of the region A. Furthermore, they allow for very natural expressions for important information-theoretic quantities such as the conditional entropy, mutual information, and conditional mutual information. They also provide proofs of important properties like subadditivity and strong subadditivity that correspond directly to the information-theoretic meaning of these properties.

The RT formula (1.1), and therefore the bit-thread formula (1.3), requires the bulk theory and state to obey a particular set of conditions, including being in the classical limit (large-N limit of the field theory), governed by Einstein gravity (strong-coupling limit of the field theory), and in a state possessing a time-reflection symmetry. It is by now more or less understood how to relax each of these conditions on the RT formula. If the bit threads indeed have a fundamental physical significance — as opposed to being just a mathematical artifact of the simplicity of the RT formula — then it should be possible to relax these conditions for the bit threads as well.

In this paper, we will take up this challenge by exploring how to relax one of the above conditions, namely the assumption that the bulk is described by Einstein gravity. In other words, we will show how to formulate bit threads when the bulk gravitational action includes higher-curvature terms. This is equivalent to moving away from the strong-coupling limit in the boundary theory. For work on black-hole entropy and HEE in higher-curvature gravity theories see [3–13]. The original bit thread paper [2] discussed one possible way

to generalize bit threads to higher-curvature gravity, by employing a conjectural generalization of the Weyl law for the spectrum of the scalar Laplacian to vector fields. In this paper, we will take a more direct approach.

In higher-curvature theories, the right-hand side of the RT formula is corrected by the minimum of a local geometrical functional of the surface:

$$S(A) = \frac{1}{4G_N} \min_{m \sim A} a_\lambda(m), \qquad a_\lambda(m) := \int_m \sqrt{g}\left(1 + \lambda \tilde{\mathcal{A}}\right), \tag{1.4}$$

where λ is a small parameter and $\tilde{A}$ is a function of the intrinsic and extrinsic geometry and the ambient metric [11]. For example, the simplest example of a higher curvature correction is for Gauss-Bonnet gravity, where λ is the coefficient of the Gauss-Bonnet term in the gravitational action and $\tilde{\mathcal{A}}$ is the scalar curvature of the induced metric on m [14, 15].

In order to generalize the bit threads to higher curvature gravity, we need to find a generalization of the max flow-min cut theorem (1.2) in which the area functional is replaced by a_λ and the right-hand side is corrected in an appropriate way. The right-hand side involves three ingredients: the objective to be maximized $\int_A \sqrt{h} n_\mu v^\mu$, the divergencelessness condition $\nabla_\mu v^\mu = 0$, and the norm bound $|v| \leq 1$. We will find that the $\lambda \tilde{\mathcal{A}}$ correction can be accounted for by correcting just the norm bound, replacing $|v| \leq 1$ with

$$|v| \leq F_\lambda[v], \tag{1.5}$$

leaving both the divergencelessness condition and the objective untouched. Here F_λ can depend on both v and its derivatives.

The constant λ will always be considered to be perturbatively small, allowing both the flow v^μ and the norm bound to be expressed as perturbative expansions. Thus we will write

$$F_\lambda[v] = 1 + \sum_{n=1}^{\infty} \lambda^n f_n[v]. \tag{1.6}$$

We will compute the functions f_1 and f_2. Although the entropy functional (1.4) has no order-λ^2 term, it turns out that such a term is necessary in the norm bound. The reason is that maximizing the flux subject to a norm bound with a first-order correction leads to a second-order term in the flux, which must then be cancelled with an explicit second-order correction to the norm bound. Our general result can be found in subsection 2.5, and its application to Gauss-Bonnet gravity in eq. (2.92).

We derive our results by two different methods. The first, which we call the bottleneck method, is described in section 2. It asks what the norm bound needs to be such that the maximum flux of a vector field defined on a given surface m equals $a_\lambda(m)$. This implies that, for any $m \sim A$, $a_\lambda(m)$ bounds the flux through A of any divergenceless vector field. The tightest bound is $\int_A v \leq \min_{m \sim A} a_\lambda(m)$. We then ask whether this inequality can be saturated, in other words whether a vector field defined on the minimizing surface and satisfying the norm bound can be extended to one that is defined everywhere in the space while respecting the divergenceless condition and the norm bound. This issue is non-trivial, and discussed in detail. The bottleneck method is general enough that it can in

principle be used to write any higher-curvature HEE prescription in an equivalent bit-thread formulation. We apply the bottleneck method to the specific case of Gauss-Bonnet (GB) gravity and discuss how to incorporate the accompanying Gibbons-Hawking-York (GHY) boundary term. The second method, described in section 3, applies the machinery of Lagrange dualization of convex programs to the corrected min cut problem. This method is more straightforward, but it requires the corrected min cut problem to define a convex program, which is only true under restrictive assumptions, either when the minimal area surface has a high degree of symmetry or $\tilde{\mathcal{A}}$ has a particularly simple form. The bottleneck method on the other hand is always valid. For GB HEE both methods are valid when considering boundaryless bulk entangling surfaces without extrinsic curvature. Despite the limited applicability of Lagrange dualization to our problem, we include it because it is a non-trivial self-consistency check in results. In section 4, we consider cases where the norm bound is a function purely of flow direction from the viewpoint of bit threads.

2 Bottleneck method

In this section we explain the bottleneck method and use it to derive the norm bound corrections.

2.1 Method

We will retain, from the max flow-min cut theorem, the objective $\int_A v$ as well as the divergenceless constraint $\nabla_\mu v^\mu = 0$. As a result, we still have, for any surface m homologous to A,

$$\int_m v = \int_A v \,, \tag{2.1}$$

so we can measure the flux through any m. Let m_λ^* be the surface that minimizes the functional a_λ whose minimum gives $S(A)$ in the corrected RT formula (1.4). We wish to find a norm bound

$$|v| \le F_\lambda \tag{2.2}$$

such that the maximum flux equals

$$\max_v \int_m v = a_\lambda(m_\lambda^*) = \int_{m_\lambda^*} \sqrt{g} \left(1 + \lambda \tilde{\mathcal{A}} \right) . \tag{2.3}$$

with $\lambda \tilde{\mathcal{A}}$ the perturbative correction to the HEE area functional. This is not a trivial task because F_λ is not allowed to depend explicitly on m_λ^*, only on v and the local geometry since the norm bound (2.2) will be imposed everywhere without reference to any particular surface.

Let us fix a surface m homologous to A, and further restrict m to be perturbatively close to the $\lambda = 0$ RT surface m_0^*. To see that this restriction does not exclude any potential bottlenecks, note that corrections to the flow must be perturbatively small, so the position where the bit threads are maximally packed can only move a perturbatively small distance, assuming no flat directions in the position of the RT surface. A useful property of minimal

area surfaces such as m_0^* is that K the trace of the extrinsic curvature vanishes. This implies, since m is perturbatively close to m_0^*,

$$K = \mathcal{O}(\lambda) \text{ on } m. \qquad (2.4)$$

We will do a local analysis on m, thus in this subsection the vector field v is defined only on m and its neighborhood. Only when it is established that the norm bound we derive has flow solutions that do not violate it in the neighborhood of m, will v be extended off the surface onto the entire time slice. We will derive a norm bound of the form (2.2) such that

$$\max_v \int_m v = a_\lambda(m), \quad |v| \le F_\lambda[v] \text{ on } m. \qquad (2.5)$$

Separating v into its norm $|v|$ and direction $\hat{v}$, the flux is

$$\int_m v = \int_m \sqrt{\tilde{g}}\, u_\mu v^\mu = \int_m \sqrt{\tilde{g}}\, u_\mu \hat{v}^\mu\, |v| \qquad (2.6)$$

with u the unit normal on m. To maximize the flux for a given direction field $\hat{v}$ (assuming $u_\mu \hat{v}^\mu \ge 0$, as it will be) corresponding to a fixed orientation of the bit threads on m, the threads should clearly be maximally packed, saturating the norm bound such that

$$|v| = F_\lambda[v] \quad \text{on } m, \qquad (2.7)$$

so what we want to maximize and match is

$$\max_v \int_m \sqrt{\tilde{g}}\, u_\mu \hat{v}^\mu F_\lambda[v] = \int_m \sqrt{\tilde{g}}(1 + \lambda\tilde{\mathcal{A}}). \qquad (2.8)$$

In subsection 2.2, we will find the conditions on the maximizing vector field and the surface m such that the vector field can be extended off of that surface while respecting both the divergencelessness constraint and the norm bound. In subsection 2.3 we derive some useful perturbative expansions in λ, and in 2.4 we will find a suitable F_λ by first considering the vector field v on m, and requiring that (2.3) hold. This establishes

$$\max_v \int_A v = a_\lambda(m_\lambda^*) \qquad (2.9)$$

as the maximum flux from A is bounded by the maximum flux through the bottleneck, which when (2.3) holds is also the surface that minimizes $a_\lambda(m)$.

In this section Gaussian normal coordinates (GNC) will sometimes be used, with the notation

$$ds^2 = dz^2 + \tilde{g}_{ij}dx^i dx^j. \qquad (2.10)$$

Thus v^z is the component of v normal to hypersurfaces of constant z, and v^i the tangential components. We take m to be the surface with $z = 0$.

2.2 Obstruction equations

The flux $\int_A v$ through the boundary subregion A is equal to the flux $\int_m v$ through any surface m homologous to A, from the divergenceless condition on v and Stoke's theorem. This is assuming there is no obstruction to the flow, that is, given some v on m, it is possible to extend v from m to the boundary $(A \cup A^c)$ without anywhere violating the norm bound. For $\lambda = 0$, this is true by the MFMC theorem. We would like to know whether for $\lambda \neq 0$ this holds true, or whether the maximally packed threads will inevitably collide.

Away from the bottlenecks the bit threads are far from maximally packed; stopping them from running into each other and violating the norm bound is simple, as there is a lot of space for maneuvering. On the minimal surface however, the threads are maximally packed and their directionality fixed; there is no choice on how to orient the threads in order to stop them colliding, so just off the surface is the most likely place for the norm bound to be violated.

Let us calculate the necessary condition for the threads not to collide off a surface m, given that the norm bound on v is saturated on that surface

$$|v| - F_\lambda[v] = 0. \tag{2.11}$$

If the difference between the norm bound and $F_\lambda[v]$ is anywhere positive, on or off the surface, then the norm bound has been violated. Assuming the flow to be smooth, the norm bound must be saturated to linear order in distance from m

$$\partial_z(|v| - F_\lambda[v])|_m = 0, \tag{2.12}$$

which may be written as

$$v_\mu \nabla_z v^\mu = F_\lambda[v] \nabla_z F_\lambda[v]. \tag{2.13}$$

To quadratic order in distance from m the non-violation of the norm bound is expressed as a bound on the second normal derivative

$$\nabla_z^2(|v| - F_\lambda[v])|_m \leq 0. \tag{2.14}$$

which may be written as

$$v_\mu \nabla_z^2 v^\mu + \nabla_z v_\mu \nabla_z v^\mu - F_\lambda[v] \nabla_z^2 F_\lambda[v] - (\nabla_z F_\lambda[v])^2 \leq 0. \tag{2.15}$$

2.3 Perturbative expansions of the flow

Here we introduce our notation for keeping track of λ dependence, and derive useful perturbative expansions of the flow. The bulk geometry is a solution to the Einstein's field equations with higher curvature corrections, so has dependence on λ. The flow solution which maximizes the flux also depends on λ. In perturbative expansions care must be taken to keep track of the perturbative order of every term. The norm bound is expanded

$$F_\lambda[v] = 1 + \sum_{n=1}^\infty \lambda^n f_n[v] \tag{2.16}$$

and the flow

$$v = v_0 + \sum_{n=1}^{\infty} \lambda^n v_n. \tag{2.17}$$

While v_n has no λ dependence, $f_n[v]$ does. For other quantities, perturbative expansions in λ use (n) to denote the nth order in the expansion. In our notation, by definition a quantity with a (n) superscript never has λ dependence. For example, the metric we expand as

$$g_{\mu\nu} = \sum_{n=0}^{\infty} \lambda^n g_{\mu\nu}^{(n)}. \tag{2.18}$$

In GNC the zz component of the metric is exactly 1 by definition,

$$g_{zz}^{(0)} = 1, \qquad g_{zz}^{(n>0)} = 0 \tag{2.19}$$

while the tangential components do have λ dependence with

$$g_{ij}^{(n)} = \tilde{g}_{ij}^{(n)}. \tag{2.20}$$

$\tilde{g}$ is the induced metric on m. In our notation the only quantities without a (n) superscript which do *not* have λ dependence are v_n and quantities which we have explicitly shown and stated to have no λ dependence.

A useful simplification is made using a result from the zeroth-order RT bit threads, that the maximizing flow on the minimal area surface equals the unit normal, $v_0|_{m_0^*} = u$. As m_0^* and m are perturbatively close, v will still be normal to m at zeroth-order and so

$$v_0|_m = u, \tag{2.21}$$

which in turn implies that $|v_0|$ has no λ dependence as

$$\begin{aligned} |v_0| &= (g_{\mu\nu}^{(0)} v_0^\mu v_0^\nu + \lambda g_{\mu\nu}^{(1)} v_0^\mu v_0^\nu + \ldots)^{1/2} \\ &= (g_{zz}^{(0)} + \lambda g_{zz}^{(1)} + \ldots)^{1/2} \\ &= 1 \end{aligned} \tag{2.22}$$

so that on m

$$|v_0| = |v_0|^{(0)} = 1, \qquad |v_0|^{(n>0)} = 0. \tag{2.23}$$

v_0 and $|v_0|$ do not have λ dependence, so neither does $\hat{v}_0$ or the projection tensor

$$\begin{aligned} P^\mu{}_\nu[\hat{v}_0] &:= \delta^\mu{}_\nu - \hat{v}_0^\mu \hat{v}_{0\nu} \\ &= \delta^\mu{}_\nu - g_{\nu\rho} \hat{v}_0^\mu \hat{v}_0^\rho \\ &= \delta^\mu{}_\nu - \delta^z{}_\nu \hat{v}_0^\mu. \end{aligned} \tag{2.24}$$

Taking the norm of v and perturbatively expanding gives

$$\begin{aligned} |v| &= |v_0| + \lambda \hat{v}_{0\mu} v_1^\mu + \frac{\lambda^2}{|v_0|} \left(v_{0\mu} v_2^\mu + \frac{1}{2} v_{1\mu} v_1^\mu - \frac{1}{2}(v_{0\mu} v_1^\mu)^2 \right) + \mathcal{O}(\lambda^3) \\ &= 1 + \lambda v_1^z + \lambda^2 \left(v_2^z + \frac{1}{2} v_{1i} v_1^i \right) + \mathcal{O}(\lambda^3) \end{aligned} \tag{2.25}$$

When $|v|$ saturates its norm bound, we can use the results derived so far to perturbatively expand both sides of (2.7), finding

$$v_1^z = (f_1[v])^{(0)} \tag{2.26}$$

and

$$v_2^z + \frac{1}{2}\tilde{g}_{ij}^{(0)} v_1^i v_1^j = (f_1[v])^{(1)} + (f_2[v])^{(0)}. \tag{2.27}$$

The direction of v can be expanded in v_n to

$$\hat{v}^\mu = \hat{v}_0^\mu + \lambda \frac{P^\mu{}_\nu[\hat{v}_0]v_1^\nu}{|v_0|} - \lambda^2 \frac{(\hat{v}_0^\mu P^\rho{}_\nu[\hat{v}_0] + 2\hat{v}_0^\rho P^\mu{}_\nu[\hat{v}_0])v_1^\nu v_{1\rho} - 2|v_0|P^\mu{}_\nu[\hat{v}_0]v_2^\nu}{2|v_0|^2} + \mathcal{O}(\lambda^3) \tag{2.28}$$

We only need the normal component of $\hat{v}$ for our procedure; see equation (2.8). On m this is

$$\hat{v}^z = 1 - \lambda^2 \frac{v_{1i}v_1^i}{2} + \mathcal{O}(\lambda^3). \tag{2.29}$$

Before proceeding with the maximization of the flux order by order in λ, note that while $\sqrt{\tilde{g}}$ has a λ expansion it is common to both the area functional and flux sides of (2.8) and so is a spectator; while not left out, it will be ignored.

2.4 Maximization of flux

In this subsection we maximize the flux order by order in λ, making use of the perturbative expansions (2.25) and (2.29), and evaluate the obstruction equations (2.13) and (2.15), in order to determine $F_\lambda[v]$. At each order in λ, there are three pieces of information that can be used to constrain v on and off m,

1. The norm bound $|v|$ is saturated on m.

2. The direction $\hat{v}$ is such that the flux through m is maximized.

3. The norm bound cannot be violated anywhere off m.

Information that can be found about f_n or v_n at a given order in λ can be used at higher-order.

Zeroth order. *Norm bound:* the norm bound to zeroth-order is

$$|v|^{(0)} = (\tilde{g}_{ij}^{(0)} v_0^i v_0^j + (v_0^z)^2)^{1/2} = 1. \tag{2.30}$$

Flux: the flux to zeroth order in λ is

$$\int_m \sqrt{\tilde{g}}\hat{v}_0^z \tag{2.31}$$

which given the zeroth order norm bound (2.30) is maximized when

$$v_0^z = 1, \qquad v_0^i = 0 \tag{2.32}$$

As stated earlier this is a known result from RT bit threads, derived here using the novel bottleneck method.

That $v_0 = u$ on m allows us to replace v_0 with u in functionals which do not contain derivatives perpendicular to the surface, for example $k_{ij}[v_0] = K_{ij}$. Those normal derivatives are thus far unconstrained, for example v_0^i on m is known, while $\partial_z v_0^i$ is not.

Linear obstruction equation: the zeroth-order of the linear obstruction equation is

$$(v_\mu \nabla_z v^\mu)^{(0)} = \nabla_z v_0^z = \partial_z v_0^z = 0. \tag{2.33}$$

From this we can show that the trace of the extrinsic curvature on m must vanish to zeroth-order, using

$$\partial_z v_0^z = (\nabla_z v^z)^{(0)} = (\nabla_\mu v^\mu - \nabla_i v^i)^{(0)} = -(\Gamma_{iz}^i)^{(0)} = -K^{(0)}. \tag{2.34}$$

which implies

$$K^{(0)} = 0. \tag{2.35}$$

That the zeroth-order component in the trace of the extrinsic curvature vanishes is thus a no-obstruction constraint on m. This is consistent with the restriction made earlier that m be perturbatively close to a minimal area RT surface. For RT bit threads this is a known result, that it is not possible to extend bit threads off a surface on which the threads are maximally packed without violating the norm bound, unless it is a RT surface, with $K = 0$.

Quadratic obstruction equation: the zeroth order in the quadratic obstruction equation (2.15) is

$$\begin{aligned}
(v_{0\mu} \nabla_z^2 v_0^\mu + \nabla_z v_{0\mu} \nabla_z v_0^\mu)^{(0)} &= (\nabla_z^2 v_0^z + \tilde{g}_{ij} \nabla_z v_0^i \nabla_z v_0^j)^{(0)} \\
&= (R_{zz} - \nabla_i \nabla_z v_0^i + \partial_z v_{0i} \partial_z v_0^i)^{(0)} \le 0.
\end{aligned} \tag{2.36}$$

where the last line follows from

$$\nabla_z^2 v^z = \nabla_z(\nabla_\mu v^\mu - \nabla_i v^i) = -\nabla_z \nabla_i v^i = R_{\mu z} v^\mu - \nabla_i \nabla_z v^i. \tag{2.37}$$

Note that in expressions of the form $(\ldots)^{(0)}$, v can be replaced with v_0 and vice versa. The bound (2.36) is a constraint on how v_0 changes off the surface, and the max-flow-min-cut theorem states that for all minimal area surfaces with $K^{(0)} = 0$ there is always an obstructionless flow v_0, and hence the above constraint inequality places no further condition on m at this order. If at higher orders we need to maximize flux over $\nabla_z v_0^i$ then this inequality will be important, but we will see at second order how the dependence of flux on $\nabla_z v_0^i$ can be removed with a suitable choice for $f_2[v]$.

First order. *Flux:* the flux to first order in λ is

$$\left(\int_m v\right)^{(1)} = \int_m \sqrt{\tilde{g}}((\hat{v}^z)^{(0)}(F_\lambda[v])^{(1)} + (\hat{v}^z)^{(1)}(F_\lambda[v])^{(0)} = \int_m \sqrt{\tilde{g}}(f_1[v_0])^{(0)} \tag{2.38}$$

Comparing this to the first-order term in $a_\lambda(m)$ implies

$$(f_1[v_0])^{(0)} = \tilde{\mathcal{A}}^{(0)}. \tag{2.39}$$

Norm bound: the norm bound saturation (2.26) on m to first order is

$$v_1^z = (f_1[v_0])^{(0)} = \tilde{\mathcal{A}}^{(0)}, \tag{2.40}$$

however the tangential components of v_1 are undetermined at this order.

Linear obstruction equation: the first-order in the linear obstruction equation (2.13) is

$$\begin{aligned}
(v_\mu \nabla_z v^\mu - F_\lambda[v] \nabla_z F_\lambda[v])^{(1)} &= (\nabla_z v_1^z + v_{1i} \partial_z v_0^i - \partial_z f_1[v])^{(0)} \\
&= 0.
\end{aligned} \tag{2.41}$$

where we have used $\partial_z v_0^z = 0$ on m.

Quadratic obstruction equation: the first order of the quadratic obstruction equation does not have any impact on our flux maximization, more details are given in appendix A.

Second order. *Flux:* the flux to second order in λ is

$$\begin{aligned}
\left(\int_m v \right)^{(2)} &= \int_m \sqrt{\tilde{g}} (\hat{v}_z^{(0)} (F_\lambda[v])^{(2)} + \hat{v}_z^{(1)} (F_\lambda[v])^{(1)} + \hat{v}_z^{(2)} (F_\lambda[v])^{(0)}) \\
&= \int_m \sqrt{\tilde{g}} \left(f_1[v]^{(1)} + f_2[v]^{(0)} - \frac{\tilde{g}_{ij}^{(0)} v_1^i v_1^j}{2} \right) \\
&= \int_m \sqrt{\tilde{g}} \left((f_1[v_0])^{(1)} + (f_1[v] - f_1[v_0])^{(1)} + f_2[v]^{(0)} - \frac{\tilde{g}_{ij}^{(0)} v_1^i v_1^j}{2} \right)
\end{aligned} \tag{2.42}$$

We will maximize this contribution to the flux with respect to v_1, so it is important to know the v_1 dependence of each term, to this end in the last line we separated $(f_1[v])^{(1)}$ into terms containing only v_0, and those exactly linear in v_1.

Suppose we took the functional $\tilde{\mathcal{A}}[u]$ and replaced u with $\hat{v}$. Let us call that functional $\tilde{a}[\hat{v}]$. As $\hat{v}_0 = u$ on m, and $\tilde{\mathcal{A}}$ contains only derivatives projected tangentially to the surface we have

$$\tilde{a}[\hat{v}_0] = \tilde{\mathcal{A}}, \tag{2.43}$$

If we choose f_1 to equal $\tilde{a}$ then we have

$$f_1[v_0] = \tilde{\mathcal{A}} \tag{2.44}$$

and the correction to the HEE surface functional is captured to all orders in λ. There is still work to do however as there are terms left over in the second order flux, which as everything in the HEE functional has been accounted for must equal zero. These additional terms come from the flow being perfectly normal to m, only to zeroth order in λ, the higher order corrections to the norm bound (f_2 and above) exist to cancel overcorrections to the flux.

We would like to keep f_1 as general as possible, so note that we can add

$$f_{1b}[v] = \sum_{n=1}^{\infty} p_n[v] (v^\mu \partial_\mu |v|)^n \tag{2.45}$$

to $f_1[v]$ without changing $f_1[v_0]$ on m, as $v_0^\mu \partial_\mu |v_0| = \partial_z v_0^z = 0$ on the surface and so (2.44) is still satisfied. $f_{1b}[v_0]$ does not therefore affect the flux, but f_{1b} is important for the flow

to be obstructionless. p_n are unfixed functions. $v^\mu \partial_\mu |v|$ measures change in bit thread number density tangential to the flow. The first order correction to the norm bound thus has two components,

$$f_1[v] = \tilde{a}[\hat{v}] + f_{1b}[v], \tag{2.46}$$

one which captures the surface functional correction $\tilde{A}$, and the other which ensures flow is obstructionless.

Let us return to the second order flux and calculate the contribution from f_{1b},

$$
\begin{aligned}
\int_m \sqrt{\tilde{g}}(f_{1b}[v] - f_{1b}[v_0])^{(1)} &= \sum_{n=1}^\infty \int_m \sqrt{\tilde{g}}(p_n[v](v^\mu \partial_\mu |v|)^n - p_n[v_0](v_0^\mu \partial_\mu |v_0|)^n)^{(1)} \\
&= \int_m \sqrt{\tilde{g}}(p_1[v])^{(0)}(\partial_z |v|)^{(1)} \\
&= \int_m \sqrt{\tilde{g}}(p_1[v]\partial_z f_1[v])^{(0)}.
\end{aligned}
\tag{2.47}
$$

To reach the last line we have used the first order linear obstruction constraint

$$(\partial_z(|v| - (1 + \lambda f_1[v])))^{(1)} = 0. \tag{2.48}$$

Now let us calculate the contribution to the second order flux from $(\tilde{a}[v] - \tilde{a}[v_0])^{(1)}$. As we will be performing a functional variation around $v = v_0$, we need to understand what derivative terms of v can appear. As derivatives of u in $\tilde{A}$ must be projected tangential to m, all derivative terms of $\hat{v}$ in $\tilde{a}[\hat{v}]$ must be projected onto the normal subspace of v. A consequence of this is that terms involving the normal derivatives of v vanish at zeroth order, for example, suppose that $\tilde{a}[\hat{v}]$ is the trace of the extrinsic curvature

$$k[\hat{v}] = P^{\mu\nu}[\hat{v}]\nabla_\mu \hat{v}_\nu, \tag{2.49}$$

then terms such as

$$\left(\frac{\partial \tilde{k}[\hat{v}]}{\partial(\nabla_z \hat{v}_\mu)}\right)^{(0)} = (P^{z\nu}[\hat{v}])^{(0)} = g^{z\nu} - u^z u^\nu = 0. \tag{2.50}$$

This leaves just derivatives tangential to m, which can be integrated by parts to strip off all the derivatives acting on v_1, in a fashion similar to the derivation of the Euler-Lagrange equation. The contribution from $\tilde{a}$ is

$$
\begin{aligned}
\int_m \sqrt{\tilde{g}}(\tilde{a}[\hat{v}] - \tilde{a}[\hat{v}_0])^{(1)} &= \int_m \sqrt{\tilde{g}}\left(\tilde{a}\left[\hat{v}_0^\mu + \lambda \frac{P^\mu{}_\nu[\hat{v}_0]v_1^\nu}{|v_0|}\right] - \tilde{a}[\hat{v}_0]\right)^{(1)} \\
&= \int_m \sqrt{\tilde{g}}\left(\frac{\partial \tilde{a}[\hat{v}]}{\partial \hat{v}^\mu}\frac{P^\mu{}_\nu[\hat{v}_0]v_1^\nu}{|v_0|} + \frac{\partial \tilde{a}[\hat{v}]}{\partial(\nabla_\rho \hat{v}^\mu)}\nabla_\rho\left(\frac{P^\mu{}_\nu[\hat{v}_0]v_1^\nu}{|v_0|}\right) + \ldots\right)^{(0)} \\
&= \int_m \sqrt{\tilde{g}}\left(\frac{\partial \tilde{a}[\hat{v}]}{\partial \hat{v}^i}v_1^i + \frac{\partial \tilde{a}[\hat{v}]}{\partial(\nabla_j \hat{v}^\mu)}\nabla_j\left(\frac{P^\mu{}_\nu[\hat{v}_0]v_1^\nu}{|v_0|}\right) + \ldots\right)^{(0)} \\
&= \int_m \sqrt{\tilde{g}}\left(\left(\frac{\partial \tilde{a}[\hat{v}]}{\partial \hat{v}^i} - \nabla_j\frac{\partial \tilde{a}[\hat{v}]}{\partial(\nabla_j \hat{v}^i)} + \ldots\right)v_1^i\right)^{(0)} + \text{boundary terms} \\
&= \int_m \sqrt{\tilde{g}}\left(\zeta_i[v]v_1^i\right)^{(0)} + \text{boundary terms}
\end{aligned}
\tag{2.51}
$$

with the definition

$$\zeta_i[v] := \left(\frac{\partial \tilde{a}[\hat{v}]}{\partial \hat{v}^i} - \nabla_j \frac{\partial \tilde{a}[\hat{v}]}{\partial (\nabla_j \hat{v}^i)} + \nabla_k \nabla_j \frac{\partial \tilde{a}[\hat{v}]}{\partial (\nabla_j \nabla_k \hat{v}^i)} - \dots \right) \tag{2.52}$$

Let us assume that the boundary terms vanish. We will explicitly show they do for GB HEE.

As $v_0 = \hat{v}_0$ on m, we could have replaced any of the $\hat{v}_0$ terms in $\tilde{a}[\hat{v}_0]$ with v_0. Then there would be additional terms in the second order flux involving v_0 and $v_1^z = \tilde{\mathcal{A}}^{(0)}$. As functions purely of v_0 such as these are easily removed with a suitable choice for f_2, as we will see, we are in effect only shuffling terms between f_1 and f_2 and nothing is lost by taking $\tilde{a}$ to be purely a function of the direction field $\hat{v}$. Substituting both the contributions from $\tilde{a}$ and f_{1b} into the second order flux gives

$$\left(\int_m v \right)^{(2)} = \int_m \sqrt{\tilde{g}} \left(\tilde{\mathcal{A}}^{(1)} + \zeta_i[v] v_1^i + p_1[v] \partial_z f_1[v] + f_2[v] - \frac{v_1^i v_{1i}}{2} \right)^{(0)} \tag{2.53}$$

For now we will assume there is no constraint on v_1^i from the obstruction equations, maximize the second order flux with respect to v_1^i, then find a $p_1[v]$ such that the obstruction equation is satisfied for this maximizing value of v_1^i. The maximizing value of v_1^i is

$$v_{1i} = (\zeta_i[v])^{(0)} \tag{2.54}$$

for which the second order flux is

$$\int_m \sqrt{\tilde{g}} \left(\tilde{\mathcal{A}}^{(1)} + p_1[v] \partial_z f_1[v] + f_2[v] + \frac{\zeta^i[v] \zeta_i[v]}{2} \right)^{(0)}. \tag{2.55}$$

To determine f_2 we equate this with the second order of $a_\lambda(m)$,

$$(a_\lambda(m))^{(2)} = \int_m \sqrt{\tilde{g}} \tilde{\mathcal{A}}^{(1)} \tag{2.56}$$

which implies

$$\begin{aligned}
(f_2[v])^{(0)} &= \left(-\frac{1}{2} \zeta_i[v] \zeta^i[v] - p_1[v] \partial_z f_1[v] \right)^{(0)} \\
&= \left(-\frac{1}{2} \zeta_i[v] \zeta^i[v] - p_1[v] (\partial_z \tilde{a}[\hat{v}] + p_1[v] \partial_z^2 |v|) \right)^{(0)}.
\end{aligned} \tag{2.57}$$

While this does not fully determine f_2, since we will not continue to third-order in λ any choice satisfying the above constraint is adequate for our purposes, so let us take the simplest, defined over the extended flow domain

$$f_2[v] = -\frac{1}{2} P_{\mu\nu}[\hat{v}] \zeta^\mu[v] \zeta^\nu[v] - p_1[v] v^\mu \partial_\mu \tilde{a}[\hat{v}] - (p_1[v])^2 v^\mu v^\nu \nabla_\mu \nabla_\nu |v|. \tag{2.58}$$

To determine p_1, we return to the first order linear obstruction equation (2.41),

$$\begin{aligned}
(\nabla_z v_1^z + v_{1i} \partial_z v_0^i - \partial_z f_1[v])^{(0)} &= (-\nabla_i v_1^i - K^{(1)} + v_{1i} \partial_z v_0^i - \partial_z \tilde{a}[\hat{v}] - p_1[v] \partial_z^2 |v|)^{(0)} \\
&= 0.
\end{aligned} \tag{2.59}$$

where we have used the divergencelessness of v to relate $\nabla_z v_1^z$ to derivatives tangential to m,

$$\begin{aligned}
(\nabla_\mu v^\mu)^{(1)} &= (\nabla_\mu v_0^\mu + \lambda \nabla_\mu v_1^\mu)^{(1)} \\
&= (\partial_z v_0^z + \partial_i v_0^i + \Gamma_{i\mu}^i v_0^\mu)^{(1)} + (\nabla_\mu v_1^\mu)^{(0)} \\
&= K^{(1)} + (\nabla_z v_1^z + \nabla_i v_1^i)^{(0)} \\
&= 0.
\end{aligned} \tag{2.60}$$

Before we maximized the second order flux with respect to v_1^i, this obstruction equation (2.59) contained two unconstrained components of the flow v_1^i and $\partial_z v_0^i$, plus any further unconstrained derivatives of v_0 from one's choise of p_1. If we simply solve the obstruction equation (2.59) for p_1

$$\begin{aligned}
(p_1[v])^{(0)} &= \frac{(\partial_z |v|)^{(1)} - (\partial_z \tilde{a}[\hat{v}])^{(0)}}{(\partial_z^2 |v|)^{(0)}} \\
&= \left(\frac{1}{\partial_z^2 |v|} (\zeta_i[v]\partial_z v_0^i - \nabla_i \zeta^i[v] - \partial_z \tilde{a}[v] - K^{(1)}) \right)^{(0)}.
\end{aligned} \tag{2.61}$$

and define $p_1[v]$ such that it evaluates to this on m, then no matter what values the unconstrained components of the flow take there is no obstruction to the flow, at first order in λ and first order in distance from m. We were free to choose any function for p_1, however besides the special choice given above, (2.59) gives a constraint on v_1^i in terms of $\partial_z v_{0i}$ and whatever other unconstrained derivatives of v_0 appear in the choice for $(p_1[v])^{(0)}$, and this constraint needs to be imposed when maximizing the second order flux with respect to v_1^i. The above choice for p_1 is merely the most convenient.

The choice for p_1 given by (2.61) is singular whenever $\partial_z^2 |v_0| = 0$ and the numerator is non-zero. Away from m_λ^* this is not an issue as regions where the flow capacity is infinite do not affect the bottleneck position. As the flow always seeks to maximize flux, we only need to assume the existence of any v_0 for which $\partial_z^2 |v_0| \neq 0$ everywhere (corresponding to threads always moving apart), or even if no such v_0 exists, that there is not a new bottleneck created. We also need to argue that the flow can not take advantage of this choice of p_1 in order to increase the capacity of the bottleneck. The correction to the norm bound on m from p_1 is $\lambda(p_1[v]\partial_z|v|)^{(0)}$. Now $(\partial_z|v|)^{(0)}$ always equals 0 on m, however the flow can still try to increase capacity by choosing $\partial_z^2 |v_0| = 0$ in which case we need to apply L'Hôpital's rule to evaluate the ratio $\partial_z |v_0| / \partial_z^2 |v_0|$. $\partial_z^3 |v_0|$ must be zero on m for there to be no obstruction, so we consider $\partial_z^3 |v_0| / \partial_z^4 |v_0|$. Again the flow can take $\partial_z^4 |v_0| = 0$, and so on, for the ratio $\partial_z^{2n-1} |v_0| / \partial_z^{2n} |v_0|$, the limit of which is where $|v_0| = 1$ everywhere.

Now that we have found a condition such that there is no obstruction to the flow, we may extend v off the surface such that it is defined throughout the time slice. Let us choose a function for p_1 whose domain is over this extended flow, which when evaluated on m satisfies the constraint (2.61),

$$p_1[v] = \frac{1}{v^\mu v^\nu \nabla_\mu \nabla_\nu |v|} (P^\mu{}_\nu[v](\zeta_\mu[v]v^\rho \nabla_\rho v^\nu - \nabla_\mu \zeta^\nu[v]) - v^\rho \nabla_\rho \tilde{a}[v] - k_1[v]). \tag{2.62}$$

$k_1[v]$ is the function formed by replacing u in $K^{(1)}$ with $\hat{v}$. This definition for $p_1[v]$ makes no reference to any particular surface, $K^{(1)}$ and hence $k_1[v]$ can be derived directly from the HEE surface functional by taking a variational derivative, only the equation of motion of the minimizing surface is needed, not its solution.

There are terms in $f_{1b}[v]$ that are higher order in $v^\mu \partial_\mu |v|$ that are still unfixed at this order in λ, however as we will not proceed to the next order we are free to set them to zero, $p_{(n \geq 2)}[v] = 0$, giving

$$f_1[v] = \tilde{a}[\hat{v}] + p_1[v] v^\mu \partial_\mu |v|. \tag{2.63}$$

Norm bound: we can determine the value of v_2^z on m from the saturation of the norm bound to second-order. Using (2.25) and (2.29),

$$\begin{aligned} \int_m (|v| \hat{v}^z)^{(2)} &= \int_m (|v|^{(2)} + (\hat{v}^z)^{(2)}) \\ &= \int_m v_2^z \\ &= \int_m \tilde{\mathcal{A}}^{(1)}. \end{aligned} \tag{2.64}$$

This gives the component of v_2 normal to m,

$$v_2^z = \tilde{\mathcal{A}}^{(1)} \tag{2.65}$$

Obstruction equations: the second-order of the obstruction equations gives conditions on v_2^i which would only be relevant if we continued to maximizing flux at third-order.

2.5 Combined results

Combining all results from zeroth to second-order, the maximizing value of v on m is

$$v^\mu|_m = (1, \vec{0}) + \lambda(\tilde{\mathcal{A}}^{(0)}, \zeta^i[v_0]^{(0)}) + \lambda^2 (\tilde{\mathcal{A}}^{(1)}, v_2^i) + \mathcal{O}(\lambda^2). \tag{2.66}$$

with v_2^i unknown at second order in λ. The norm bound extended off of m, without reference to any surface, is

$$\begin{aligned} |v| \leq &1 + \lambda(\tilde{a}[\hat{v}] + p_1[v] v^\mu \partial_\mu |v|) \\ &- \lambda^2 \left(\frac{1}{2} P_{\mu\nu}[\hat{v}] \zeta^\mu[v] \zeta^\nu[v] + p_1[v] v^\mu (\partial_\mu \tilde{a}[v] + p_1[v] v^\nu \partial_\mu \partial_\nu |v|) \right) + \mathcal{O}(\lambda^3) \end{aligned} \tag{2.67}$$

with p_1 defined as

$$p_1[v] := \frac{1}{v^\mu v^\nu \nabla_\mu \partial_\nu |v|} (P^\mu{}_\nu[v](\zeta_\mu[v] v^\rho \nabla_\rho v^\nu - \nabla_\mu \zeta^\nu[v]) - v^\rho \partial_\rho \tilde{a}[\hat{v}] - k_1[v]). \tag{2.68}$$

and ζ defined as

$$\zeta_\mu[v] := \left(\frac{\partial \tilde{a}[\hat{v}]}{\partial \hat{v}^\mu} - P^{\nu\rho}[\hat{v}] \nabla_\nu \frac{\partial \tilde{a}[\hat{v}]}{\partial (\nabla^\rho \hat{v}^\mu)} + P^{\nu\rho}[\hat{v}] P^{\sigma\omega}[\hat{v}] \nabla_\nu \nabla_\sigma \frac{\partial \tilde{a}[\hat{v}]}{\partial (\nabla^\omega \nabla^\rho \hat{v}^\mu)} - \dots \right). \tag{2.69}$$

This is as high in orders of λ as we will go. In principle one could continue the procedure of maximizing the flux and equating it to the HEE functional to even higher order, and this would continue to give corrections to the value of the flow on m and the norm bound. At each order in λ a new degree of freedom v_n^μ is added over which the flux is maximized, and corrections to the norm bound are added to correct for over/undershooting.

2.6 Application to Gauss-Bonnet gravity

Let us apply our results to Gauss-Bonnet (GB) gravity, where the correction to the surface functional is

$$\tilde{\mathcal{A}} = \tilde{R}. \tag{2.70}$$

with $\tilde{R}$ the induced scalar curvature of the surface

$$\tilde{R} = R - 2R_{\mu\nu}u^\mu u^\nu + (K_\mu{}^\mu)^2 - K_{\mu\nu}K^{\mu\nu} \tag{2.71}$$

and $K_{\mu\nu}$ the extrinsic curvature tensor

$$K_{\mu\nu} = P_\mu{}^\rho[u]\nabla_\rho u_\nu. \tag{2.72}$$

Gauss-Bonnet gravity is the simplest extension to Einstein gravity that is a Lovelock theory. The Lagrangian in a Lovelock theory is a sum of Euler densities,

$$\mathcal{L}_{2p} \coloneqq \frac{1}{2^p}\delta^{\nu_1...\nu_{2p}}_{\mu_1...\mu_{2p}}R^{\mu_1\mu_2}_{\nu_1\nu_2}\ldots R^{\mu_{2p-1}\mu_{2p}}_{\nu_{2p-1}\nu_{2p}}, \tag{2.73}$$

quantities whose integrals are topological invariants in $2p$ dimensions. The equations of motion of such a theory contain only second derivatives of the metric, meaning that they require the same initial data as Einstein gravity. GB gravity includes, in addition to the usual cosmological constant ($p = 0$) and Einstein-Hilbert ($p = 1$) terms, the $p = 2$ term:

$$I = \frac{1}{16\pi G_{\mathrm{N}}}\int\sqrt{g}\left(-2\Lambda + R + \lambda(R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\lambda\sigma}R^{\mu\nu\lambda\sigma})\right) + \text{boundary terms}, \tag{2.74}$$

where λ is a parameter with dimensions of length-squared. For HEE in GB gravity, the entropy is given by minimizing a functional which includes the area plus the integrated induced Ricci scalar [13–15]. The GB HEE functional is

$$a_\lambda(m) \coloneqq \int_m\sqrt{\tilde{g}}(1 + \lambda\tilde{R}) + 2\lambda\int_{\partial m}\sqrt{\tilde{h}}\tilde{K}, \tag{2.75}$$

and where we use tildes to denote quantites defined with respect to the induced metric $\tilde{g}_{ij}$ on the surface m, $\tilde{K}$ is the trace of the extrinsic curvature not of m but ∂m. The Gibbons-Hawking-York (GHY) boundary term in (2.75) is necessary to give a well-posed variational problem. We should again emphasize that we could have chosen any higher-curvature correction to Einstein gravity to illustrate our method, as long as the entropy is given by minimizing a local functional on surfaces in the homology class of A. We will not be using any special properties of Lovelock theories, Gauss-Bonnet gravity is merely a simple extension to consider.

Before proceeding, we note an important caveat regarding the GB HEE formula. Naively, it gives $-\infty$ for the entropy of any region. This can be easily seen in $3+1$ bulk dimensions, where the surface m is 2-dimensional and the λ terms in (2.75) are proportional to its Euler character $\chi(m)$:

$$\int_m\sqrt{\tilde{g}}\tilde{R} + 2\int_{\partial m}\sqrt{\tilde{h}}\tilde{K} = 4\pi\chi(m). \tag{2.76}$$

By adding small handles or spheres to the surface m, its Euler character can be made arbitrarily negative or positive without significantly changing the total area. Hence, for either sign of λ, the GB HEE formula, taken at face value, tells us the entropy will always be $-\infty$! However, one should remember that (2.74) should be treated as an effective action, with λ treated as a perturbative parameter, rather than assigned a finite value. Correspondingly, λ should be treated as a perturbative parameter. In other words, the embedding coordinates of the surface m should be written as a power series in λ, and then the surface functional minimized order by order in λ. In turn, all calculations demonstrating our methods on GB gravity will be done perturbatively in the Gauss-Bonnet parameter λ.

By varying $a_\lambda(m)$ we find the equation of motion for m_λ^*,

$$(1 + \lambda\tilde{R})K - 2\lambda\tilde{R}_{ij}K^{ij} = 0, \tag{2.77}$$

which implies

$$K^{(0)} = 0, \qquad K^{(1)} = 2(\tilde{R}_{ij}K^{ij})^{(0)}. \tag{2.78}$$

The correction $\lambda\tilde{\mathcal{A}}$ cannot affect the zeroth order result, so we start at first-order. With $\tilde{\mathcal{A}} = \tilde{R}$, we have $\tilde{a} = r$ where

$$r[\hat{v}] := R - 2R_{\mu\nu}\hat{v}^\mu\hat{v}^\nu + (k_\mu{}^\mu[\hat{v}])^2 - k_{\mu\nu}[\hat{v}]k^{\mu\nu}[\hat{v}]$$
$$k_{\mu\nu}[\hat{v}] := P_\mu{}^\rho[\hat{v}]\nabla_\rho\hat{v}_\nu, \quad P_\mu{}^\nu[\hat{v}] := \delta_\mu{}^\nu - \hat{v}_\mu\hat{v}^\nu. \tag{2.79}$$

The leading order correction to the thread thickness is $\lambda r[v_0] = \lambda\tilde{R}$, so loosely speaking the more curved the surfaces which are perpendicular to the flow are, the more the thread thickness is affected, thicker or thinner depending on the sign of λ.

Given $\tilde{a}$, we next calculate the terms in ζ_i,

$$\frac{\partial r[\hat{v}]}{\partial\hat{v}^i} = -4R_{iz} + 2k_i{}^j[\hat{v}]\nabla_z\hat{v}_j, \qquad \nabla_j\frac{\partial r}{\partial(\nabla_j\hat{v}^i)} = -2\nabla_j k^j{}_i[\hat{v}] \tag{2.80}$$

giving

$$(\zeta_i[v])^{(0)} = (-4R_{iz} + 2k_i{}^j[\hat{v}]\nabla_z\hat{v}_j + 2\nabla_j k^j{}_i[\hat{v}])^{(0)}$$
$$= 2(K_i{}^j\nabla_z\hat{v}_j - R_{iz})^{(0)} \tag{2.81}$$

which uses the identity

$$R_{iz} = \nabla_j K^j{}_i - \partial_i K$$
$$\implies (R_{iz})^{(0)} = (\nabla_j k^j{}_i[\hat{v}])^{(0)}. \tag{2.82}$$

In the derivation for the general case, we neglected the boundary terms arising from the integration by parts. For Gauss-Bonnet these are

$$\int_{\partial m}\sqrt{\tilde{h}}(k_{ij}[\hat{v}_0]\tilde{n}^i v_1^j))^{(0)} \tag{2.83}$$

In an asymptotically AdS spacetime, with spatial metric $ds^2 \sim z^{-2}(dx^\mu)^2$ and cutoff $z = z_0$, m_0^* has extrinsic curvature components K_{ij} which remain finite on the boundary, while $\tilde{n}^i$ goes like z_0, so $(K_{ij}\tilde{n}^i)^2$ goes like z_0^4, and therefore vanishes as $z_0 \to 0$.

With ζ we can calculate f_1. Recall that

$$f_1[v] = \tilde{a}[\hat{v}] + p_1[v]v^\mu \partial_\mu|v| \tag{2.84}$$

with $\tilde{a}[\hat{v}] = r[\hat{v}]$ for GB HEE, and that on m p_1 is

$$\left(p_1[v] = \frac{1}{\partial_z^2|v|}(\zeta_i[v]\partial_z v_0^i - \nabla_i\zeta^i[v] - \partial_z\tilde{a}[v] - K^{(1)})\right)^{(0)} \tag{2.85}$$

Let us calculate each of the terms on the right individually,

$$
\begin{aligned}
(\zeta^i[v]\partial_z v_i)^{(0)} &= 2((k^i{}_j[\hat{v}]\partial_z\hat{v}^j - R^i{}_z)\partial_z v_i)^{(0)} \\
(-\nabla_i\zeta^i[\hat{v}])^{(0)} &= 2(-\nabla_i k^i{}_j[\hat{v}]\partial_z\hat{v}^j - k^i{}_j[\hat{v}]\nabla_i\nabla_z\hat{v}^j + \nabla_i R^i{}_z)^{(0)} \\
&= 2(-R^i_z\partial_z\hat{v}_i - k^i{}_j[\hat{v}]\nabla_i\nabla_z\hat{v}^j + \nabla_i R^i{}_z)^{(0)} \\
-K^{(1)} &= -2(K^{ij}\tilde{R}_{ij})^{(0)} \\
(-\partial_z r[v])^{(0)} &= (-\partial_z R + 2\partial_z R_{zz} + 4R_{zz}\partial_z v^z + 4R^i{}_z\partial_z v_i \\
&\qquad + 2K^{ij}\nabla_z k_{ij}[v] - 2K^i{}_i\nabla_z k^j{}_j[\hat{v}])^{(0)} \\
&= (-\partial_z R + 2\partial_z R_{zz} + 4R^i{}_z\partial_z v_i + 2K^{ij}\nabla_z k_{ij}[\hat{v}])^{(0)} \\
&= (-2\nabla_i R^i{}_z + 4R^i{}_z\partial_z v_i + 2K^{ij}\nabla_z k_{ij}[\hat{v}])^{(0)} \\
&= (-2\nabla_i R^i{}_z + 4R^i{}_z\partial_z v_i + 2K^{ij}\nabla_z((\delta^\rho_i - v_i v^\rho)\nabla_\rho v_j))^{(0)} \\
&= (-2\nabla_i R^i{}_z + 4R^i{}_z\partial_z v_i + 2K^{ij}(\nabla_z\nabla_i v_j - \partial_z v_i\partial_z v_j))^{(0)} \\
&= (-2\nabla_i R^i{}_z + 4R^i{}_z\partial_z v_i + 2K^{ij}(\nabla_i\nabla_z v_j - \partial_z v_i\partial_z v_j + R_{jzzi}))^{(0)}.
\end{aligned}
\tag{2.86}
$$

For $\partial_z r$, in the first line the third and last terms vanish using $\partial_z v_0^z = 0$ and $K^{(0)} = 0$, the third line makes use of the contracted Bianchi identity

$$\partial_\rho R = 2\nabla_\mu R^\mu{}_\rho, \tag{2.87}$$

and the final line uses the relation between the commutator of covariant derivatives and the Riemann tensor

$$[\nabla_\mu, \nabla_\nu]V_\rho = R_{\rho\sigma\mu\nu}V^\sigma. \tag{2.88}$$

Combining these contributions many terms cancel giving

$$
\begin{aligned}
(p_1[v])^{(0)} &= \left(\frac{1}{\partial_z^2|v|}2K^{ij}(\tilde{R}_{ij} - R_{jzzi})\right)^{(0)} \\
&= \left(\frac{1}{\partial_z^2|v|}2K^{ij}(R_{ij} - K_i{}^m K_{mj})\right)^{(0)}
\end{aligned}
\tag{2.89}
$$

on m, using the GNC identities

$$\tilde{R}_{ij} = R_{ij} + \partial_z K_{ij} + K K_{ij} - 2K_i{}^m K_{mj} \tag{2.90}$$

and

$$R_{izzj} = \partial_z K_{ij} - K_i{}^m K_{mj}. \tag{2.91}$$

Thus for GB HEE, the norm bound defined without reference to any surface is

$$|v| \leq 1 + \lambda(r[\hat{v}] + p_1[v]v^\mu \partial_\mu |v|)$$
$$- \lambda^2 \left(\frac{1}{2} P_{\mu\nu}[\hat{v}] \zeta^\mu[v] \zeta^\nu[v] + p_1[v]v^\mu (\partial_\mu r[v] + p_1[v]v^\nu \nabla_\mu \partial_\nu |v|) \right) + \mathcal{O}(\lambda^3) \qquad (2.92)$$

with

$$\zeta^\mu[v] = 2P^\mu{}_\nu[\hat{v}] v^\sigma (K^\nu{}_\rho[\hat{v}] \nabla_\sigma \hat{v}^\rho - R^\nu{}_\sigma) \qquad (2.93)$$

and

$$p_1[v] = \frac{2k^{\mu\nu}[\hat{v}](R_{\mu\nu} - k_\mu{}^\rho[\hat{v}] k_{\rho\nu}[\hat{v}])}{v^\mu v^\nu \nabla_\mu \partial_\nu |v|}. \qquad (2.94)$$

The bit thread formulation of GB HEE simplifies when the RT surface m_0^* has no extrinsic curvature, such that $(K^{ij})^{(0)} = 0$ and so $(p_1[v])^{(0)}$ vanishes on m, and we can choose $p_1 = 0$ to simplify the norm bound. The tangential component of v_1, $(\zeta^i[v])^{(0)}$ also vanishes on m using the identity that relates the iz component of the Ricci tensor to the vanishing extrinsic curvature,

$$R_{iz} = \nabla_j K^j{}_i - \partial_i K. \qquad (2.95)$$

Furthermore, a simpler form for r can be used,

$$r[\hat{v}] = R - 2R_{\mu\nu} \hat{v}^\mu \hat{v}^\nu \qquad (2.96)$$

which still satisfies $(r[\hat{v}])^{(0)} = \tilde{R}^{(0)}$, $(r[\hat{v}])^{(1)} = \tilde{R}^{(1)}$ on m, so is adequate in giving the correct flux up to second order in λ. This gives us the norm bound

$$|v| \leq 1 + \lambda(R - 2R_{\mu\nu} \hat{v}^\mu \hat{v}^\nu) + \mathcal{O}(\lambda^3) \qquad (2.97)$$

for cases where m_0^* has no extrinsic curvature. We will compare this norm bound with the result derived using Lagrange dualization and find agreement.

2.6.1 Gibbons-Hawking-York term

The GHY term in $a_\lambda(m)$ has so far been neglected. We present two ways to incorporate it: adding a term to the norm bound with delta-function support on the boundary, and a doubling trick, taking ∂M to be the boundary both of the original Riemannian manifold M and an identical copy, with bit threads flowing out into both.

The GHY term contains $\tilde{K}[\tilde{n}]$, the divergence of the surface's boundary normal $\tilde{n}$. By allowing an additional flux through on ∂m we capture the GHY term, however the difficulty is doing so without making reference to any surface. While we do not a priori know where the bottleneck will be, we do know what v will be on it from which we can extract $\tilde{n}$ and thus $\tilde{K}[\tilde{n}]$.

Straightforwardly the unit normal $\tilde{n}$ can be written as the normalized projection of the time slice's boundary unit normal n onto the tangent space of m, which is

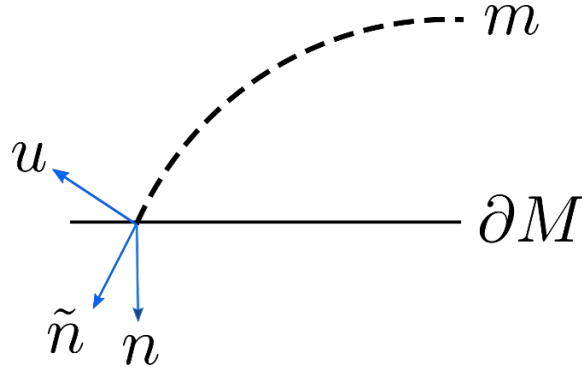$$\tilde{n}^\mu = \frac{P^\mu{}_\nu[u]n^\nu}{|P[u]n|}, \qquad (2.98)$$

Figure 1. Illustration showing u the unit normal to m, $\tilde{n}$ the unit normal to ∂m, and n the unit normal to the bulk time slice boundary ∂M at ∂m.

see figure 1. We would like to adapt this formula for $\tilde{n}$ to use v instead of u, therefore not making reference to any particular surface. Note also that on the boundary of m, as argued earlier the extrinsic curvature always vanishes, hence so too does v_1^i,

$$
\begin{aligned}
v_1^i &= (\zeta^i[v])^{(0)} \\
&= 2(k_i{}^j[\hat{v}]\nabla_z \hat{v}_j - R_{iz})^{(0)} \\
&= 2(K_i{}^j \nabla_z \hat{v}_j - \nabla_j K^j{}_i)^{(0)} \\
&= 0
\end{aligned}
\tag{2.99}
$$

and therefore v is normal to m to at least second order,

$$
v|_{\partial m} = u + \mathcal{O}(\lambda^2).
\tag{2.100}
$$

Let us define a function on the spacetime boundary ∂M

$$
\tilde{N}^\mu[v] := \frac{P^\mu{}_\nu[v]n^\nu}{|P[v]n|}
\tag{2.101}
$$

then at ∂A (2.100) holds and so

$$
\tilde{N}[v] = \tilde{n} + \mathcal{O}(\lambda^2)
\tag{2.102}
$$

on ∂A. To account for the GHY term we add to the norm bound a function with $\delta-$function support on ∂A, which allows additional finite flux through ∂A equal to the GHY term

$$
|v| \leq 1 + \lambda(f_1[v] + 2k^\nu{}_\nu[\tilde{N}[v]]\delta_{\partial A}) + \lambda^2 f_2[v] + \mathcal{O}(\lambda^3)
\tag{2.103}
$$

Note that the norm bound (2.103) is defined purely in terms of boundary geometric data and an unconstrained v, such that bit thread thickness is only a function of local geometry and thread orientation. Using the formula for $\tilde{N}[v]$, our method generalises to any higher curvature HEE prescriptions whose boundary term is a functional of the surface boundary normal $\tilde{n}$.
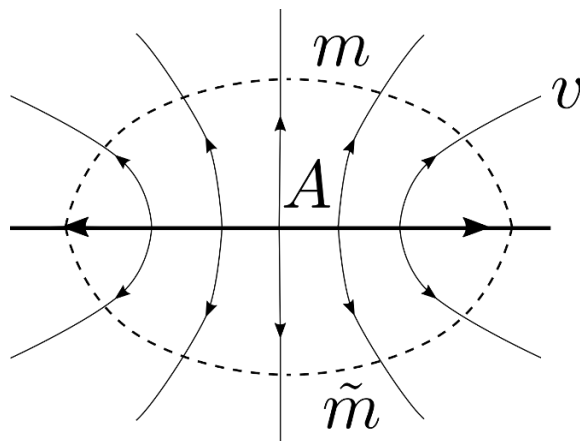
Figure 2. By gluing two copies of the time slice Σ along the boundary $\partial\Sigma$, we create a boundaryless surface $m \cup \tilde{m}$. The GHY term is accounted for by the integral of $\tilde{R}$ over where m and $\tilde{m}$ join.

An alternative way of including the contribution of the GHY term is to employ a doubling trick. Taking m to be a surface homologous to boundary subregion A, and adding the mirror image $\tilde{m}$ of m across the boundary, creates a boundaryless surface $m + \tilde{m}$ for which

$$\int_m \tilde{R} + 2\int_{\partial m} K = \frac{1}{2}\int_{m\cup\tilde{m}} \tilde{R}. \tag{2.104}$$

After gluing the surface m together with its double, there may be a kink in the surface at ∂A, giving a singular induced scalar curvature. The GHY terms can then be understood as accounting for possible delta-function singularities in $\tilde{R}$ where we join m with its mirror image. From the norm bound (2.103), this implies infinite bit thread density at ∂A, though the flux is still finite.

In this doubling trick picture, the entanglement entropy is given by half the maximum flux out of boundary region A,

$$4G_N S(A) = \frac{1}{2}\max_v \int_A v \tag{2.105}$$

where v can flow out into two copies of Σ glued along $\partial\Sigma$, subject to divergenceless of v and the norm bound, see figure 2.

3 Lagrange dualization method

The max flow-min cut theorem (1.2) is proven as a consequence of strong Lagrange duality between two convex optimization problems, namely max flow and a relaxed form of the min cut problem. (A review of these concepts aimed at physicists can be found in [16].) In this section, we will apply these ideas to the Gauss-Bonnet holographic entanglement entropy formula. Unfortunately, as we will see, the λ term in the funcational $a_\lambda(m)$ in general ruins the convexity of the relaxed min cut functional. Therefore, the technique will only work in certain special cases, namely when the minimal surface has no boundary and vanishing extrinsic curvature $K_{ij} = 0$, such as when calculating the entanglement entropy

of one side in the high-temperature thermofield-double state. This will allow us to replace the non-convex optimization problem $\min_{m \sim A} a_\lambda(m)$ with an equivalent convex optimization problem. This is important because non-convex problems generally have a duality gap between the primal and dual problem.[1] Specializing to the situation where the problem is convex will then allow us to use Lagrange dualization to derive the flow reformulation.

3.1 Convex optimization and Lagrange dualization

We present here a brief review of the mathematics of Lagrange dualization and its application to HEE, however the authors strongly suggest that readers unfamiliar with these to read the more detailed expositions in sections 2 and 3 of [16] before trying to follow their extension to the higher-curvature case in subsection 3.2 of this paper.

3.1.1 Review of Lagrange dualization

Lagrange duality is a technique often employed in the fields of linear programming and network theory. For a well defined class of minimization problems (the primal) there exists a description where the problem has been transformed into a maximization problem (the dual). Strong duality is the nontrivial assertion that these two descriptions are in fact the same, that the maximum of one equals the minimum of the other.

Let L_p and $\{f_a\}$ be a set of convex functions, and $\{h_b\}$ a set of affine functions on a vector space parametrized by x. The primal program is given by the constrained optimization program

$$\min_x L_p(x) \quad \text{s.t.} \quad f_a(x) \le 0 \quad h_b(x) = 0 \,. \tag{3.1}$$

We may rewrite L_p by imposing Lagrange multipliers for the constraints

$$L(x, \{\phi_a\}, \{\gamma_b\}) \equiv L_p(x) + \phi_a f_a(x) + \gamma_b h_b(x), \quad \phi_a \ge 0. \tag{3.2}$$

The primal problem L_p may be recovered from L by maximizing with respect to the Lagrange multipliers

$$L_p(x) = \max_{\phi_a \ge 0, \gamma_b} L(x, \{\phi_a\}, \{\gamma_b\}). \tag{3.3}$$

To obtain the dual program however, we instead minimize L with respect to x

$$L_d(\{\phi_a\}, \{\gamma_b\}) \equiv \min_x L(x, \{\phi_a\}, \{\gamma_b\}) \,. \tag{3.4}$$

The difference between the solution of the primal program and the dual program is called the duality gap

$$d_g \equiv \min_x L_p(x_i) - \max_{\phi_a \ge 0, \gamma_b} L_d(\{\phi_a\}, \{\gamma_b\}). \tag{3.5}$$

When the duality gap is zero then strong duality is said to hold. A sufficient, but not necessary condition for strong duality to hold is for $L_p(x_i)$ to be a convex function, and

---

[1]A general procedure exists called convex relaxation which allows one to embed a nonconvex problem in a larger solution space which is convex. When such a relaxation can be done it is possible to find a dual with zero duality gap. So far we have not been able to find such a relaxation which would allow the Gauss-Bonnet holographic entanglement entropy to be calculated in the general case. We leave this for future work.

there to exist an x_i in the relative interior of its domain for which the constraints are satisfied; this is Slater's condition.

When the dual program has a unique optimal configuration (ϕ_a^*, γ_b^*), then the values of the Lagrange multipliers tell us how sensitive the optimal value is to small changes in the constraints. In other words, if we replace the constraint $f_a(x) \leq 0$ by $f_a(x) + \lambda \leq 0$, then to first order in λ the optimal value changes by $\lambda \phi_a^*$. This is derived for example in [16]. In fact, by a slight generalization of that argument, the result holds even when we perturb the constraint by a function $\delta f_a(x)$: if the dual optimal configuration is unique and if $\delta f_a(x^*)$ has the same value for all primal optimal points x^*, then replacing $f_a(x) \leq 0$ by $f_a(x) + \lambda \delta f_a(x) \leq 0$ changes the optimal value by $\lambda \phi_a^* \delta f_a(x^*) + \mathcal{O}(\lambda^2)$. The same result also holds for concave programs: replacing $f_a(x) \geq 0$ by $f_a(x) + \lambda \delta f_a(x) \geq 0$ changes the optimal value by $\lambda \phi_a^* \delta f_a(x^*) + \mathcal{O}(\lambda^2)$. We will use this fact in subsection 3.1.3.

3.1.2 Example: max flow-min cut

As an example to further familiarize readers, and to set the stage for the derivation with GB gravity, we will show how to apply Lagrange dualization to the RT formula with EH gravity, proving the Riemannian MFMC theorem. Let M be a Riemannian manifold with boundary, in this case a constant time slice of a bulk spacetime in a static state of a holographic theory. Given a region $A \subset \partial M$ the HEE is

$$S(A) = \frac{1}{4G_{\mathrm{N}}} \min_{m \sim A} \int_m \sqrt{\tilde{g}} \,. \tag{3.6}$$

To define the problem as a well posed convex program we perform a convex relaxation of the program by adding a scalar field degree of freedom ψ, which is subject to the boundary condition $\psi|_{\partial M} = \chi_A$ with $\chi_A = 1$ on A and 0 on the complement A^c. This has the effect of smearing the surface to form level sets of constant ψ in the bulk. The optimal solution involves stacking these level sets all on the true minimal surface. The space of surfaces m is a subspace of possible $\psi(x)$, when $\psi(x)$ is binary valued, equal to 1 is a bulk region (not necessarily connected) and 0 in the complement, then surface m can be understood as the boundary of these regions. The resulting optimization problem is convex in ψ:

$$\min_{m \sim A} \int_m \sqrt{\tilde{g}} = \min_{\psi} \int_M \sqrt{g} |\partial_\mu \psi|, \quad \psi|_{\partial M} = \chi_A \,. \tag{3.7}$$

In order to proceed with the dualization we introduce Lagrange multiplier term v^μ enforcing the replacement of $\partial_\mu \psi$ with a new vector degree of freedom w^μ, and a boundary term that is minimized when $\psi|_{\partial M} = \chi_A$

$$S(A) = \frac{1}{4G_N} \min_{\psi, v^\mu} \left[\int_M \sqrt{g} \left[|w| + v^\mu(w_\mu - \partial_\mu \psi) \right] + \int_{\partial M} \sqrt{h} |\psi - \chi_A| \right]. \tag{3.8}$$

Optimizing first with respect to the Lagrange multipliers imposes the constraints and returns us to the primal program so let us instead optimize over the fields w_μ and ψ, giving the dual program

$$S(A) = \frac{1}{4G_N} \max_{v^\mu} \int_A \sqrt{h} \, n_\mu v^\mu, \quad |v| \leq 1, \quad \nabla_\mu v^\mu = 0. \tag{3.9}$$

As the primal problem was convex and obeys Slater's condition, strong duality holds, and thus the Riemannian MFMC theorem is proven.

3.1.3 Perturbing the convex program

The flow formulation of RT HEE given by (3.9) is a well-posed convex program (more precisely, concave program, since it involves maximizing a concave functional). Dualizing it returns us to the relaxed min cut program (3.7). (See [16] for the details of this derivation.) By viewing the max flow program as the primal, we can use the relation between perturbations of the primal constraints and changes in the optimal value, described at the end of subsection 3.1.1 above, to figure out how to change the norm bound in the max flow program in order to reproduce the λ term in the GB HEE functional. This gives a very straightforward way to find the first-order correction to the norm bound.

In the dualization of the max flow program, there is a Lagrange multiplier ψ for the divergencelessness constraint and another one ϕ for the norm bound. As long as the minimal surface m_0^* is unique, the dual optimal configuration is also unique; in particular, ϕ^* is a delta function on m_0^*.

In this subsection we will only work to first order in λ. If we perturb the norm bound to make it

$$1 - |v| + \lambda F \geq 0, \tag{3.10}$$

where F is some function on M, then the maximum flux will change by

$$\lambda \int \sqrt{g}\, F \phi^* = \lambda \int_{m_0^*} \sqrt{\tilde{g}} F. \tag{3.11}$$

In particular, if we choose F to be any function which equals $\tilde{\mathcal{A}}$ on m_0^*, then the maximum flux will equal

$$\int_{m_0^*} \sqrt{\tilde{g}}\,(1 + \lambda \tilde{\mathcal{A}}) = a_\lambda(m_0^*) = a_\lambda(m_\lambda^*) + \mathcal{O}(\lambda^2), \tag{3.12}$$

where we used the fact that m_0^* extremizes the area, so $\mathrm{area}(m_\lambda^*) = \mathrm{area}(m_0^*) + \mathcal{O}(\lambda^2)$. In order for F to equal $\tilde{\mathcal{A}}$ on m_0^* for any max flow, we set it equal to $\tilde{a}[\hat{v}]$. The norm bound is thus

$$|v| \leq 1 + \lambda \tilde{a}[\hat{v}]. \tag{3.13}$$

3.2 Lagrange dualization of higher curvature holographic entanglement entropy

In this section, Lagrange dualization is applied to optimization problems of the form

$$\min_{m \sim A} \int_m \sqrt{\tilde{g}}(1 + \lambda \tilde{\mathcal{A}}), \tag{3.14}$$

corresponding to a perturbative correction of the RT HEE prescription.

We now carry out the same convex relaxation as in section 3.1.2, such that the normal vector field with δ-function support becomes a one-form $\partial_\mu \psi$ supported over the bulk time slice M with $\psi \in \mathbb{R}$,

$$u^\mu \to \frac{\partial^\mu \psi}{|\partial \psi|}. \tag{3.15}$$

Heuristically, this convex relaxation smears the surface over the manifold forming a foliation of hypersurfaces with $\partial_\mu \psi / |\partial \psi|$ the unit normal on a component surface. This gives

$$\min_{\psi} \left[\int_M \sqrt{g} \left(1 + \lambda \tilde{a} \left[\frac{\partial^\mu \psi}{|\partial \psi|} \right] \right) |\partial \psi| + \int_{\partial M} \sqrt{h} \left(1 + \lambda \tilde{\mathcal{A}} \right) |\chi_A - \psi| \right] \quad (3.16)$$

where $\chi_A = 1$ in A and 0 in A^c. We restrict ourselves to the case where $\tilde{\mathcal{A}}$ depends on the surface unit normal u, but not derivatives of u as they generally cause the problem to be non-convex. For example, suppose $\tilde{\mathcal{A}}$ contains terms involving the trace of the extrinsic curvature. M is foliated by hypersurfaces of constant ψ, so smooth changes to $\psi(x)$ can lead to discontinous changes in hypersurface foliation, with very different extrinsic curvatures. A consequence of this is that the convexity condition

$$pK[\psi_a] + (1-p)K[\psi_b] \geq K[p\psi_a + (1-p)\psi_b], \quad 0 \leq p \leq 1 \quad (3.17)$$

can be violated to an arbitrary degree, making it a non-convex optimization problem.

We next add a Lagrange multiplier term $v_\mu(w^\mu - \partial^\mu \psi)$ to replace derivatives of ψ with w^μ, arriving at the following Lagrangian:

$$L[\psi, w, v] = \int_M \sqrt{g} \left[(1 + \lambda \tilde{a}[\hat{w}])|w| + v_\mu(w^\mu - \partial^\mu \psi) \right]$$
$$+ \int_{\partial M} \sqrt{h}(1 + \lambda \tilde{\mathcal{A}})|\psi - \chi_A| \quad (3.18)$$

where $\hat{w}^\mu \equiv w^\mu/|w|$. We now minimize over the variables ψ and w^μ on M and ∂M. Integrating the $v_\mu \partial^\mu \psi$ term in (3.18) by parts strips all derivatives off ψ, allowing us to do a pointwise minimization. The terms involving ψ are

$$\int_M \sqrt{g} \, \psi \nabla_\mu v^\mu + \int_{\partial M} \sqrt{h} \left((1 + \lambda \tilde{R})|\psi - \chi_A| + \psi n_{\partial M}^\mu v_\mu \right). \quad (3.19)$$

The bulk integrand is unbounded unless

$$\nabla_\mu v^\mu = 0, \quad (3.20)$$

in which case it vanishes. On the boundary, in order to have a bounded minimum in ψ, we require

$$|n^\mu v_\mu| \leq 1 + \lambda \tilde{\mathcal{A}}; \quad (3.21)$$

this inequality allows for the possibility that A itself is the flow bottleneck. The minimum is at $\psi = \chi_A$, leaving us with

$$\min_{\psi} L[\psi, w, v] = \int_M \sqrt{g} \left[(1 + \lambda \tilde{a}[\hat{w}])|w| + v_\mu w^\mu \right] + \int_A \sqrt{h} v_\mu n_{\partial M}^\mu \quad (3.22)$$

Let us minimise the bulk integrand

$$(1 + v_\mu \hat{w}^\mu + \lambda \tilde{a}[\hat{w}]) |w|. \quad (3.23)$$

with respect to w. If the prefactor for $|w|$ is negative for any value of its direction $\hat{w}$, then the minimum is unbounded by sending its magnitude $|w| \to \infty$. Thus we require, for all values of $\hat{w}$,

$$1 + v_\mu \hat{w}^\mu + \lambda \tilde{\mathcal{A}}(\hat{w}) \geq 0 \,, \tag{3.24}$$

and then the minimum is zero at $|w| = 0$. To see whether (3.24) holds for any $\hat{w}$, we minimize the left-hand side of the inequality with respect to $\hat{w}$, subject of course to the constraint $\hat{w}^\mu \hat{w}_\mu = 1$, finding the minimizing value for $\hat{w}$

$$\begin{aligned}
\hat{w}^\mu &= -\frac{v^\mu + \lambda \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\mu}}{|v + \lambda \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}}|} \\
&= -\hat{v}^\mu - \frac{\lambda}{|v|} P^\mu{}_\nu[\hat{v}] \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\nu} + \lambda^2 \frac{(\hat{v}^\mu P^\rho{}_\nu[\hat{v}] + 2\hat{v}^\rho P^\mu{}_\nu[\hat{v}])}{2|v|^2} \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\nu} \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}^\rho} + \mathcal{O}(\lambda^3)
\end{aligned} \tag{3.25}$$

using (2.28). Contracting with v_μ gives

$$v_\mu \hat{w}^\mu = -|v| + \frac{\lambda^2}{2|v|} P_{\mu\nu}[\hat{v}] \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\mu} \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\nu}. \tag{3.26}$$

Taking this minimizing value of $\hat{w}$ and Taylor expanding the $\lambda \tilde{a}[\hat{w}]$ term in (3.24) about $-\hat{v}$ gives

$$\lambda \tilde{a}[\hat{w}] = \lambda \tilde{a}[\hat{w}]|_{\hat{w}=-\hat{v}} - \frac{\lambda^2}{|v|} P_{\mu\nu}[\hat{v}] \left(\frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\mu} \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\nu} \right) \bigg|_{\hat{w}=-\hat{v}} \tag{3.27}$$

Substituting (3.26) and (3.27) into the inequality (3.24) gives a constraint on v, which is the norm bound

$$|v| \leq 1 + \lambda \tilde{a}[\hat{w}]|_{\hat{w}=-\hat{v}} - \frac{\lambda^2}{2} P_{\mu\nu}[\hat{v}] \left(\frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\mu} \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\nu} \right) \bigg|_{\hat{w}=-\hat{v}} + \mathcal{O}(\lambda^3). \tag{3.28}$$

Bringing the constraints we have found together, we arrive at the dual problem

$$\begin{aligned}
\max_v \int_A \sqrt{h} v_\mu n^\mu \text{ over } \bigg\{ v^\mu : |v| &< 1 + \lambda \tilde{a}[\hat{w}]|_{\hat{w}=-\hat{v}} \\
&- \frac{\lambda^2}{2} P_{\mu\nu}[\hat{v}] \left(\frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\mu} \frac{\partial \tilde{a}[\hat{w}]}{\partial \hat{w}_\nu} \right) \bigg|_{\hat{w}=-\hat{v}} + \mathcal{O}(\lambda^3), \\
\nabla^\mu v_\mu = 0, \quad |n^\mu v_\mu| &\leq (1 + \lambda \tilde{\mathcal{A}}) \bigg\}
\end{aligned} \tag{3.29}$$

Starting with the dual problem (3.29), one can reverse the process and recover the HEE formula (3.31). The details of this calculation are non-essential to the conclusions of this paper, but it is worthwhile to note that from a convex maximal flow problem, one can find a dual minimal cut problem.

3.2.1 Application to Gauss-Bonnet gravity

Here we will apply the results of the previous section to GB HEE, in the special case where the minimal surface has vanishing extrinsic curvature and no boundary, for which the minimization of surfaces becomes a convex problem. There are some non-convex optimization problems whose Lagrange dual obtains strong duality, GB HEE is not one of them.

The GHY term contains the trace of the extrinsic curvature and is not convex, so we consider only surfaces without boundaries, for which the GB HEE formula is

$$S(A) = \frac{1}{4G_N} \int_{m_\lambda^*} \sqrt{\tilde{g}} \left(1 + \lambda (R - 2R_{\mu\nu} u^\mu u^\nu + K^2 - K_{\mu\nu} K^{\mu\nu}) \right). \tag{3.30}$$

As before, m_λ^* is the codimension-2 surface homologous to A that minimizes the surface functional, and $\tilde{g}_{\mu\nu}$ and $\tilde{R}$ are the induced metric and curvature scalar on m_λ^*. The extrinsic curvature terms in (3.30) are problematic to obtaining strong duality as they make the problem non-convex.

We will restrict ourselves to m_0^* having no extrinsic curvature, then the $\lambda K^{\mu\nu} K_{\mu\nu}$ term in the GB HEE functional will be third order on m_λ^* and can be dropped as we are only working to second order. The extrinsic curvature tensor appears only quadratically in GB HEE, so under the assumption that m_0^* has no extrinsic curvature these terms can be removed without affecting the local minimum of (3.30). In cases where m_0^* has vanishing curvature due to Killing symmetries, such as on bifurcation surfaces of Killing horizons, then m_λ^* may also have vanishing extrinsic curvature. This is the case for all known static black hole event horizons in Lovelock gravity [6].

Thus we can take

$$S(A) = \frac{1}{4G_N} \min_{m \sim A} \int_m \sqrt{\tilde{g}} \left(1 + \lambda (R - 2R_{\mu\nu} u^\mu u^\nu) \right) \tag{3.31}$$

as the primal program to dualize. We identify

$$\tilde{\mathcal{A}} = R - 2R_{\mu\nu} u^\mu u^\nu \tag{3.32}$$

as the perturbation to the RT area functional, for cases where the optimum surface m_λ^* has no boundary or extrinsic curvature. Following the procedure given in the previous section, after convex relaxation and substitution of ψ with $\hat{w}$, this becomes

$$\tilde{\mathcal{A}}(\hat{w}) = R - 2R_{\mu\nu} \hat{w}^\mu \hat{w}^\nu. \tag{3.33}$$

for which, applying the result (3.28), gives the norm bound

$$|v| \leq 1 + \lambda (R - 2R_{\mu\nu} \hat{v}^\mu \hat{v}^\nu) - 8\lambda^2 P^{\mu\nu}[\hat{v}] R_{\mu\rho} R_{\nu\sigma} \hat{v}^\rho \hat{v}^\sigma + \mathcal{O}(\lambda^3). \tag{3.34}$$

In fact, the $\mathcal{O}(\lambda^2)$ term in the above norm bound can be removed as they vanish on m_λ^*, which follows from the vanishing of extrinsic curvature terms in the identity

$$R_{iz} = \nabla_j K_i^j - \partial_i K. \tag{3.35}$$

Thus the norm bound is simply

$$|v| \leq 1 + \lambda (R - 2R_{\mu\nu} \hat{v}^\mu \hat{v}^\nu) + \mathcal{O}(\lambda^3). \tag{3.36}$$

There is perfect agreement between the norm bound found using Lagrange dualization (3.36) and the norm bound found using the bottleneck method (2.97) in their overlapping regimes of validity: when m_0^* has no boundary or extrinsic curvature. The non-trivial part of the agreement is that the second-order correction to the norm bound derived using the two methods both vanish.

4 Maximization over bit thread paths

There are special cases in which the corrected norm bound takes the form $|v| \leq F_\lambda[\hat{v}]$, with the right-hand side depending only on the direction of v. An example is the one discussed at the end of the previous section, in which the unperturbed minimal surface m_0^* has no extrinsic curvature, and the norm bound is given by (3.36). This suggests a decoupling of the norm $|v|$ and direction $\hat{v}$ of the vector field. However, the two are coupled by the divergencelessness constraint $\nabla_\mu v^\mu = 0$. Here we will show that one can nonetheless decouple the direction and norm. Thus the problem of maximizing the flow can be decomposed into two steps: for a given $\hat{v}$, maximum the norm $|v|$; then maximize over $\hat{v}$.

In the language of bit threads, the direction field $\hat{v}$ specifies the potential thread configurations, while the norm bound fixes the maximum density.

Consider a particular thread originating from a boundary point $x^i \in A$. Define a path $x^\mu(x^i, s)$ along the thread as the integral curve along $\hat{v}^\mu$: the solution to

$$\frac{d}{ds} x^\mu(x^i, s) = \hat{v}^\mu, \tag{4.1}$$

with $x^\mu(x^i, s = 0)$ the boundary point. The claim is that given knowledge only of the direction field $\hat{v}$, and the fact that we want to maximize the flux through A, we can find the thread number density everywhere in the bulk, and hence know everything about v.

First we show that if we know the thread density at any point on the thread, we know it for the whole thread. The divergencelessness of v can be written as

$$\hat{v}^\mu \nabla_\mu \ln |v| = -\nabla_\mu \hat{v}^\mu \tag{4.2}$$

Integrating this along the bit thread from the boundary at $s = 0$ to a point $s = s'$ gives

$$|v|_{(x^i, s')} = |v|_{(x^i, 0)} \exp\left(-\int_0^{s'} \nabla_\mu \hat{v}^\mu ds\right) \tag{4.3}$$

From (4.3) we see that, in order for $|v|$ to be single-valued, any loops of bit threads must obey $\oint \nabla_\mu \hat{v}^\mu = 0$. In fact, as any loops of bit threads in the bulk can only impede threads leaving A and contribute nothing to the flux, we can assume without loss of generality that the direction field is free of loops. (Given a direction field containing loops, we can simply set v to 0, making $\hat{v}$ undefined, on every point through which a loop passes.)

Next, we use the fact that, in order to maximize the flux out of A, for each point $x_i \in A$ we should increase $|v|_{(x_i, 0)}$ until there is a point along the bit thread which saturates the norm bound, which occurs for

$$|v|_{(x^i, 0)} = \min_{s'} F_\lambda[\hat{v}]_{(x^i, s)} \exp\left(\int_0^{s'} \nabla_\mu \hat{v}^\mu ds\right). \tag{4.4}$$

Thus $|v|_{(x^i, 0)}$ is known, which in turn tells us the thread density everywhere.

Threads are always maximally packed on the minimal surface, and generally spread out towards the boundaries. (4.3) says that when $\nabla_\mu \hat{v}^\mu < 0$ the threads are coming closer

together, and when $\nabla_\mu \hat{v}^\mu > 0$ the threads are moving apart. For RT bit threads, the minimal surface has $|v| = 1$ and hence $\nabla_\mu \hat{v}^\mu = 0$ on it. In most of the bulk, the threads are free to come together or move apart, but in the neighborhood of either side of the minimal surface, there must be non-zero regions of $\nabla_\mu \hat{v}^\mu$, one side which is a source for the direction field, and the other a sink. The minimal surface thus emerges in this direction field picture as the surface which separates the two source and sink regions. For GB bit threads, there is a correction to this: the minimal surface will not perfectly demarcate bands of source and sink regions, as $\nabla_\mu \hat{v}^\mu$ does not necessariy vanish on m_λ^*.

Suppose one has specified a direction field $\hat{v}$ and this gives a set of integral curves. Each integral curve has its own bottleneck, at the value of s' for which the exponential factor in (4.4) is smallest. We increase the value of $|v|$ on the boundary until the norm bound is saturated at that s'. For general direction fields, the union of neighbouring integral curve's bottleneck points won't be continuous, more like a random set of points, but for the special direction fields which give m_λ^* that union of points is in fact the continuous minimal surface we are looking for[2]. This is another way of seeing how m_λ^* appears in the bit thread picture. Finally, we note that while the higher curvature corrections to the bit threads were incorporated by altering the norm bound, there are equivalent alternatives. The bit thread prescription is simple and has few components to it, there are only three aspects the corrections can affect: the divergence of v, the norm bound, or the objective functional. By a change of variables, redefining $v^\mu \to F_\lambda[\hat{v}]v^\mu$ we regain the constant norm bound $|v| \leq 1$ at the cost of replacing the divergencelessness condition with $\nabla_\mu v^\mu = -v^\mu \partial_\mu F_\lambda[\hat{v}]$ and the objective functional with $\int_A F_\lambda[\hat{v}]v$. This field redefinition exchanges bit threads whose thickness varies with position and orientation, but must end on the boundary with threads that have constant thickness, but can start and end in the bulk. We should emphasize that this is only a change of variables. Even though the divergencesslessness condition has changed, it has nothing to do with quantum corrections. The specific form of the divergence here forces the new threads to follow the same integral curve on which they are created, effectively adding thickness to the thread. A general quantum correction would also give rise to a corrected divergencelessness condition, but would presumably allow threads to be created in the bulk which would flow more independently of the threads around it.

Acknowledgments

---

[2]In that union there will generally also be other scattered points.

A First order of quadratic obstruction equation

The first order in (2.15) is

$$
\begin{aligned}
(v_\mu &\nabla_z^2 v^\mu + (\nabla_z v^\mu)^2 - F_\lambda \nabla_z^2 F_\lambda - (\nabla_z F_\lambda)^2)^{(1)} \\
&= \left(g_{\mu\nu} v^\mu \nabla_z^2 v^\nu + g_{\mu\nu} \nabla_z v^\mu \nabla_z v^\nu\right)^{(1)} - (\nabla_z^2 F)^{(1)} \\
&= g_{\mu\nu}^{(0)} \left(v^\mu \nabla_z^2 v^\nu + \nabla_z v^\mu \nabla_z v^\nu\right)^{(1)} + g_{\mu\nu}^{(1)} \left(v^\mu \nabla_z^2 v^\nu + \nabla_z v^\mu \nabla_z v^\nu\right)^{(0)} - (\nabla_z^2 f_1[v])^{(0)} \\
&= g_{\mu\nu}^{(0)} \left(v_0^\mu (\nabla_z^2 v^\nu)^{(1)} + v_1^\mu (\nabla_z^2 v^\nu)^{(0)} + 2(\nabla_z v^\mu)^{(0)} (\nabla_z v^\nu)^{(1)}\right) \\
&\quad + \tilde{g}_{ij}^{(1)} (\nabla_z v^i)^{(0)} (\nabla_z v^j)^{(0)} - (\nabla_z^2 f_1[v])^{(0)} \\
&= (\nabla_z^2 v^z)^{(1)} + v_1^z (\nabla_z^2 v^z)^{(0)} + \tilde{g}_{ij}^{(0)} v_1^i (\nabla_z^2 v^j)^{(0)} + 2\tilde{g}_{ij}^{(0)} (\nabla_z v^i)^{(0)} (\nabla_z v^j)^{(1)} \\
&\quad + \tilde{g}_{ij}^{(1)} (\nabla_z v^i)^{(0)} (\nabla_z v^j)^{(0)} - (\nabla_z^2 f_1[v])^{(0)} \leq 0
\end{aligned}
\tag{A.1}
$$

When we reach the second-order calculation we will be performing a pointwise maximization with respect to v_1^i. If the above simplifies to a constraint purely on v_1^i and its derivative tangential to m then it reduces the space of feasible v_1^i and will be important in the second-order calculation. However, if unconstrained variables such as $\partial_z v_1^i$ do not vanish then this bound places no real constraint on the value of v_1 on m. We need to be especially careful to make full use of the divergenceless of v, which relates derivative of v in different directions. The evaluation of each term in (A.1) gives

$$
\begin{aligned}
\left(\nabla_z v^i\right)^{(0)} &= \partial_z v_0^i + K^{(0)i}{}_j v_0^j = \partial_z v_0^i \\
\left(\nabla_z v^i\right)^{(1)} &= \partial_z v_1^i + (\Gamma_{\mu z}^i v^\mu)^{(1)} = \partial_z v_1^i + K^{(0)i}{}_j v_1^j \\
\left(\nabla_z^2 v^i\right)^{(0)} &= \partial_z(\partial_z v_0^i + K^{(0)i}{}_j v_0^j) + K^{(0)i}{}_l(\partial_z v_0^l + K^{(0)l}{}_j v_0^j) = \partial_z^2 v_0^i + 2K^{(0)i}{}_j \partial_z v_0^j \\
\nabla_z^2 v^z &= \nabla_z(\nabla_\mu v^\mu - \nabla_i v^i) \\
&= -\nabla_z \nabla_i v^i = -\partial_i \partial_z v^i - v^z \partial_z K - v^j \partial_z \Gamma_{ij}^i - K \partial_z v^z - \Gamma_{ij}^i \partial_z v^j \\
\left(\nabla_z^2 v^z\right)^{(0)} &= -\partial_i \partial_z v_0^i - \partial_z K^{(0)} - \Gamma_{ij}^{(0)i} \partial_z v_0^j \\
\left(\nabla_z^2 v^z\right)^{(1)} &= -\partial_i \partial_z v_1^i - \tilde{R}^{(0)} \partial_z K^{(0)} - \partial_z K^{(1)} - v_1^j \partial_z \Gamma_{ij}^{(0)i} - \Gamma_{ij}^{(0)i} \partial_z v_1^j - \Gamma_{ij}^{(1)i} \partial_z v_0^j \\
\left(\nabla_z^2 f_1[\hat{v}]\right)^{(0)} &= \partial_z^2 f_1[v_0]
\end{aligned}
\tag{A.2}
$$

which allows us to write (A.1) as

$$
-\partial_i \partial_z v_1^i + A_i[v_0] \partial_z v_1^i + B_i[v_0] v_1^i + C[v_0] \leq 0,
\tag{A.3}
$$

with the definitions

$$
\begin{aligned}
A_i[v_0] &:= (2\tilde{g}_{ij}^{(0)} \partial_z v_0^j - \Gamma_{ij}^{(0)j}) \\
B_i[v_0] &:= 2\tilde{g}_{ij}^{(0)}(\partial_z^2 v_0^j + 4K^{(0)j}{}_l \partial_z v_0^l - \partial_z \Gamma_{ij}^{(0)i}) \\
C[v_0] &:= -\partial_z K^{(1)} - \Gamma_{ij}^{(1)i} \partial_z v_0^j + \tilde{R}^{(0)}(-\partial_i \partial_z v_0^i - 2\partial_z K^{(0)} - \Gamma_{ij}^{(0)i} \partial_z v_0^j) \\
&\quad + \tilde{g}_{ij}^{(1)} \partial_z v_0^i \partial_z v_0^j - \partial^2 f_1[v_0].
\end{aligned}
\tag{A.4}
$$

Eq. (A.3) contains $\partial_z v_1^i$, so there is no real constraint on v_1^i from this obstruction equation. In contrast the first-order linear obstruction equation has no such $\partial_z v_1^i$ terms and the set of obstructionless v_1^i which the flow maximizes over is generally a subset of all v_1^i.

References

[1] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [`hep-th/0603001`] [ɪɴSPIRE].

[2] M. Freedman and M. Headrick, *Bit threads and holographic entanglement*, *Commun. Math. Phys.* **352** (2017) 407 [`arXiv:1604.00354`] [ɪɴSPIRE].

[3] A. Bhattacharyya, M. Sharma and A. Sinha, *On generalized gravitational entropy, squashed cones and holography*, *JHEP* **01** (2014) 021 [`arXiv:1308.5748`] [ɪɴSPIRE].

[4] A. Bhattacharyya, A. Kaviraj and A. Sinha, *Entanglement entropy in higher derivative holography*, *JHEP* **08** (2013) 012 [`arXiv:1305.6694`] [ɪɴSPIRE].

[5] A. Bhattacharyya and M. Sharma, *On entanglement entropy functionals in higher derivative gravity theories*, *JHEP* **10** (2014) 130 [`arXiv:1405.3511`] [ɪɴSPIRE].

[6] T. Jacobson and R.C. Myers, *Black hole entropy and higher curvature interactions*, *Phys. Rev. Lett.* **70** (1993) 3684 [`hep-th/9305016`] [ɪɴSPIRE].

[7] E. Caceres, M. Sanchez and J. Virrueta, *Holographic entanglement entropy in time dependent Gauss-Bonnet gravity*, *JHEP* **09** (2017) 127 [`arXiv:1512.05666`] [ɪɴSPIRE].

[8] J. Erdmenger, M. Flory and C. Sleight, *Conditions on holographic entangling surfaces in higher curvature gravity*, *JHEP* **06** (2014) 104 [`arXiv:1401.5075`] [ɪɴSPIRE].

[9] B. Chen and J.-j. Zhang, *Note on generalized gravitational entropy in Lovelock gravity*, *JHEP* **07** (2013) 185 [`arXiv:1305.6767`] [ɪɴSPIRE].

[10] M.R. Mohammadi Mozaffar, A. Mollabashi, M.M. Sheikh-Jabbari and M.H. Vahidinia, *Holographic entanglement entropy, field redefinition invariance and higher derivative gravity theories*, *Phys. Rev.* **D 94** (2016) 046002 [`arXiv:1603.05713`] [ɪɴSPIRE].

[11] X. Dong, *Holographic entanglement entropy for general higher derivative gravity*, *JHEP* **01** (2014) 044 [`arXiv:1310.5713`] [ɪɴSPIRE].

[12] J. Camps, *Generalized entropy and higher derivative gravity*, *JHEP* **03** (2014) 070 [`arXiv:1310.6659`] [ɪɴSPIRE].

[13] F.M. Haehl, E. Hijano, O. Parrikar and C. Rabideau, *Higher curvature gravity from entanglement in conformal field theories*, *Phys. Rev. Lett.* **120** (2018) 201602 [`arXiv:1712.06620`] [ɪɴSPIRE].

[14] J. de Boer, M. Kulaxizi and A. Parnachev, *Holographic entanglement entropy in Lovelock gravities*, *JHEP* **07** (2011) 109 [`arXiv:1101.5781`] [ɪɴSPIRE].

[15] L.-Y. Hung, R.C. Myers and M. Smolkin, *On holographic entanglement entropy and higher curvature gravity*, *JHEP* **04** (2011) 025 [`arXiv:1101.5813`] [ɪɴSPIRE].

[16] M. Headrick and V.E. Hubeny, *Riemannian and Lorentzian flow-cut theorems*, *Class. Quant. Grav.* **35** (2018) 10 [`arXiv:1710.09516`] [ɪɴSPIRE].