

Correlation Boundaries and the First Law: Entanglement Anisotropy, Holographic Screens, and the Thermodynamics of Surfaces

Work in progress Goleudy.ai February 2026

Abstract

We introduce the concept of *correlation boundaries* — surfaces across which the total pairwise correlation vanishes — and show that these are the only surfaces on which an analog of the first law of thermodynamics holds generically. For surfaces with nonzero cross-boundary correlation $P(\Sigma)$, we derive a thermodynamic deficit $\Delta = P \cdot A$ proportional to the product of the cross-boundary entanglement entropy and the entanglement anisotropy. This result resolves the finding of Wang and Braunstein [*Nat. Commun.* **9**, 2977 (2018)] that holographic screens away from horizons are not thermodynamic: horizons are correlation boundaries ($P = 0$), stretched horizons have $P \approx 0$, and spherically symmetric screens have vanishing anisotropy ($A = 0$) despite nonzero P . We identify the cross-boundary correlation $P(\Sigma)$ with the entanglement entropy S_{EE} across Σ , and show that the Ryu-Takayanagi formula emerges as the minimization of P over surfaces homologous to a boundary region. The framework requires only pairwise correlations with a monogamy (finite capacity) constraint, making it independent of specific microscopic dynamics. We discuss implications for the emergent gravity program, the quantum extremal surface prescription, and the relationship between entanglement structure and gravitational thermodynamics.

I. INTRODUCTION

The deep connection between gravity and thermodynamics, first glimpsed in the laws of black hole mechanics [1] and the Bekenstein-Hawking entropy formula [2,3], has been developed into a powerful program for understanding spacetime as emergent. Jacobson's 1995 derivation of the Einstein equations from thermodynamic relations at local Rindler horizons [4] showed that gravitational dynamics follow from assuming that $\delta Q = TdS$ holds at causal horizons, with entropy proportional to horizon area.

Verlinde [5,6] extended this program by proposing that gravity is an entropic force arising from thermodynamic properties of holographic screens — surfaces of constant Newtonian potential that encode information about the enclosed region. This entropic gravity program makes strong claims: it aims to derive Newton's laws, reproduce general relativity, and explain dark matter effects from the competition between area-law and volume-law entanglement entropy at cosmological scales [6].

A central assumption of Verlinde's program is that holographic screens — ordinary surfaces away from horizons — are intrinsically thermodynamic, possessing a well-defined temperature $T = \kappa/(2\pi)$ and an entropy proportional to area. Wang and Braunstein [7] tested this assumption rigorously. For static asymptotically flat

spacetimes, they proved that the first law holds on horizons, holds approximately on stretched horizons, but *fails* on ordinary surfaces away from horizons. The sole exception is spherically symmetric configurations, where ordinary surfaces do satisfy the first law.

This result has been widely interpreted as a serious challenge to the emergent gravity program. We argue that it is not a challenge but a *prediction* — one that follows from a simple structural principle: surfaces are thermodynamic if and only if they are *correlation boundaries*, defined as surfaces across which the total pairwise correlation vanishes. We develop this principle, derive the thermodynamic deficit for general surfaces, connect it to entanglement entropy, and show that the Ryu-Takayanagi formula [8] for holographic entanglement entropy emerges naturally.

The paper is organized as follows. Section II introduces correlation fields with monogamy constraints and defines correlation boundaries. Section III derives the thermodynamic deficit theorem. Section IV establishes the identification between cross-boundary correlation and entanglement entropy. Section V derives the Ryu-Takayanagi formula from this framework. Section VI presents the resolution of the Wang-Braunstein result. Section VII discusses implications and open questions.

II. CORRELATION FIELDS AND CORRELATION BOUNDARIES

A. Correlation fields with monogamy

Consider a system of N degrees of freedom (which we call *features*) labeled $i = 1, \dots, N$. Each pair (i, j) carries a correlation strength $\lambda_{ij} \in [0, \Lambda]$, where Λ is the maximum correlation capacity. We impose the *monogamy constraint*: each feature has finite total correlation capacity,

$$\sum_{j \neq i} \lambda_{ij} \leq \Lambda \quad \forall i. \quad (1)$$

This constraint is not an assumption about a specific physical system — it is a structural feature shared by quantum entanglement [9], mutual information in classical systems [10], and correlation measures in tensor networks [11]. In quantum mechanics, monogamy of entanglement ensures that if subsystem A is maximally entangled with B, it cannot be significantly entangled with C [12]. The constraint (1) is the general form of this principle.

In the continuum limit (large N , with features densely distributed), the discrete correlation field λ_{ij} becomes a two-point function $\lambda(x, y)$ on a manifold, with the monogamy constraint becoming

$$\int \lambda(x, y) d\mu(y) \leq \Lambda \quad \forall x. \quad (2)$$

Throughout, we work with both discrete and continuum descriptions as appropriate.

B. Surfaces and cross-boundary correlation

A *surface* Σ is a codimension-1 partition of the features into two sets: Inside (I) and Outside (O). For any such surface, define the *cross-boundary correlation*:

$$\mathcal{P}(\Sigma) = \sum_{i \in I} \sum_{j \in O} \lambda_{ij}. \quad (3)$$

This non-negative scalar measures the total correlation threading through Σ . In the continuum:

$$\mathcal{P}(\Sigma) = \int_{\Sigma} \int_O \lambda(x, y) d\mu_{\Sigma}(x) d\mu_O(y). \quad (4)$$

C. Correlation boundaries

Definition 1. A surface Σ is a *correlation boundary* if $\mathcal{P}(\Sigma) = 0$.

At a correlation boundary, every feature inside is uncorrelated with every feature outside. The surface completely isolates the interior from the exterior in terms of pairwise correlations.

Definition 2. A surface Σ is an ϵ -*approximate correlation boundary* (or *stretched correlation boundary*) if $\mathcal{P}(\Sigma)/\mathcal{P}_{\max}(\Sigma) < \epsilon$, where \mathcal{P}_{\max} is the maximum cross-boundary correlation the surface could support.

Definition 3. A surface Σ is *ordinary* if $\mathcal{P}(\Sigma)/\mathcal{P}_{\max}(\Sigma) \sim O(1)$.

The monogamy constraint (1) ensures that correlation boundaries exist. A feature's finite correlation budget Λ cannot extend to arbitrarily many partners. At some point, the budget is exhausted and correlations reach zero — creating a boundary.

D. Connection to horizons

In a gravitational context, black hole horizons are correlation boundaries: an observer outside the horizon has zero quantum correlation with degrees of freedom behind it (up to Hawking radiation, which represents the small but nonzero \mathcal{P} of an approximate correlation boundary). The Rindler horizon of an accelerating observer is also a correlation boundary — by construction, it is the limit of the observer's causal (and therefore correlational) reach.

This identification is not new. The connection between causal horizons and entanglement structure has been developed extensively in the context of holographic entanglement entropy [8,13,14] and the ER=EPR conjecture [15]. What we add is a precise criterion — $\mathcal{P}(\Sigma) = 0$ — that can be tested for any surface, not just horizons, and that makes quantitative predictions about the *degree* to which a surface fails to be thermodynamic.

III. THE THERMODYNAMIC DEFICIT

A. Energy partition across a surface

For a system partitioned by a surface Σ into I and O, the total energy decomposes as

$$E_{total} = E_I + E_O + E_{cross}(\{\lambda_{ij}\}_{i \in I, j \in O}), \quad (5)$$

where E_{cross} is the energy associated with cross-boundary correlations. Similarly, for any extensive potential Φ :

$$\Phi_{total} = \Phi_I + \Phi_O + \Phi_{cross}. \quad (6)$$

When $P(\Sigma) = 0$, there are no cross-boundary correlations, so $E_{cross} = 0$ and $\Phi_{cross} = 0$. The partition is *informationally clean*: $\Phi_{total} = \Phi_I + \Phi_O$ exactly.

When $P(\Sigma) > 0$, the cross term is nonzero. To leading order in the cross-boundary correlations:

$$\Phi_{cross} = \sum_{i \in I, j \in O} \lambda_{ij} \cdot \phi_{ij} + O(\lambda^2), \quad (7)$$

where ϕ_{ij} is the contribution to the potential per unit correlation between features i and j .

B. Perturbations and the first law

Consider a perturbation that displaces the surface Σ by $\delta\varepsilon$ in the outward normal direction $\hat{n}(x)$ at each point $x \in \Sigma$. This perturbation changes three quantities:

- (i) **Surface flux** δQ : the energy crossing the surface due to the displacement.
- (ii) **Entropy change** TdS : the change in entropy of the interior, related to internal reconfiguration.
- (iii) **Cross-boundary correlation response** $\delta\Phi_{cross}$: the change in the cross term due to the perturbation.

The perturbed energy balance is:

$$\delta Q = T dS + \delta\Phi_{cross}. \quad (8)$$

The *thermodynamic deficit* is defined as

$$\Delta_{FL}(\Sigma) \equiv \delta Q - T dS = \delta\Phi_{cross}. \quad (9)$$

The first law $\delta Q = TdS$ holds if and only if $\delta\Phi_{cross} = 0$.

C. The correlation anisotropy

When the surface Σ is displaced by $\delta\varepsilon$ at point x , features near the surface are reassigned between I and O. The change in cross-boundary correlation depends on *how correlations are distributed directionally* at x .

Define the *correlation current density* at $x \in \Sigma$:

$$\vec{\mathcal{J}}(x) = \sum_j \lambda_{xj} \hat{r}_{xj} \delta(x \in \Sigma), \quad (10)$$

where \hat{r}_{xj} is the unit vector from x toward feature j . The *correlation anisotropy* at x is the projection of the net current onto the surface normal:

$$\mathcal{A}(x) = \hat{n}(x) \cdot \left(\vec{\mathcal{J}}^{out}(x) - \vec{\mathcal{J}}^{in}(x) \right), \quad (11)$$

where the superscripts denote contributions from outside and inside features respectively.

The anisotropy measures the directional asymmetry of correlations at the surface. If correlations are distributed isotropically, \mathcal{J}^{out} and \mathcal{J}^{in} are balanced and $\mathcal{A} = 0$. If there is a preferred direction, $\mathcal{A} \neq 0$.

D. The thermodynamic deficit theorem

Theorem 1. For a surface Σ in a correlation field satisfying the monogamy constraint (1), the thermodynamic deficit under a normal displacement $\delta\varepsilon(x)$ is

$$\Delta_{FL}(\Sigma) = \int_{\Sigma} \mathcal{P}(x) \cdot \mathcal{A}(x) \cdot h[\delta\varepsilon(x)] d\mu_{\Sigma} + O(\mathcal{P}^2), \quad (12)$$

where $\mathcal{P}(x)$ is the local cross-boundary correlation density, $\mathcal{A}(x)$ is the local correlation anisotropy, and h is a function of the perturbation profile.

Proof sketch. Displace Σ by $\varepsilon \cdot \delta\varepsilon(x)$ in the outward normal. Features in a shell of width $\varepsilon \cdot \delta\varepsilon$ near the surface transfer from O to I, changing the cross-boundary correlation set. The change in Φ_{cross} receives a dominant contribution (Effect 1) from the reassignment of correlations and a subleading contribution (Effect 2, suppressed as $1/N$) from the change in susceptibilities. Effect 1 evaluates to an integral over Σ of three factors: the local feature density $n(x)$, the net directional asymmetry of correlations at x (which is $\mathcal{A}(x)$), and the perturbation profile $\delta\varepsilon(x)$. Under the assumption of approximately uniform susceptibility ($\varphi^0_{ij} \approx \bar{\varphi}$), the integrand factors into $\mathcal{P}(x) \cdot \hat{\mathcal{A}}(x) \cdot h(x)$, where $\hat{\mathcal{A}} = \mathcal{A}/\mathcal{P}$ is the normalized anisotropy and h absorbs the perturbation-dependent quantities. The product structure ensures the deficit vanishes whenever either factor is zero. The complete proof, including error bounds and the uniform susceptibility condition, is given in the Supporting Information (Secs. S1-S4). ■

For uniform perturbations, Eq. (12) simplifies to

$$\Delta_{FL} \propto \langle \mathcal{P} \cdot \mathcal{A} \rangle_{\Sigma}. \quad (13)$$

E. The anisotropy tensor and multipole decomposition

Promoting the scalar anisotropy to a tensor:

$$\mathcal{A}_{ab}(x) = \sum_j \lambda_{xj} \left(\hat{r}_{xj,a} \hat{r}_{xj,b} - \frac{1}{d} \delta_{ab} \right), \quad (14)$$

the scalar anisotropy is $A(x) = n^a n^b \mathcal{A}_{ab}(x)$. Expanding in spherical harmonics on a surface of approximate spherical topology:

$$\mathcal{A}(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \mathcal{A}_{lm} Y_{lm}(\hat{x}). \quad (15)$$

This decomposition is central to understanding the spherical symmetry exception (Sec. VI).

IV. IDENTIFICATION WITH ENTANGLEMENT ENTROPY

A. Entanglement entropy and cross-boundary correlations

The entanglement entropy of a bipartite quantum system in pure state $|\Psi\rangle$, partitioned into A and B, is

$$S_{EE}(A) = -\text{Tr}(\rho_A \ln \rho_A), \quad (16)$$

where $\rho_A = \text{Tr}_B |\Psi\rangle\langle\Psi|$. For a system described by pairwise correlations λ_{ij} , the entanglement entropy decomposes as

$$S_{EE} = \sum_{i \in I, j \in O} f(\lambda_{ij}) + \text{multi-body corrections}, \quad (17)$$

where $f(\lambda)$ is a monotonically increasing function with $f(0) = 0$ and $f(\lambda) \rightarrow \ln d$ as $\lambda \rightarrow \Lambda$, where d is the local Hilbert space dimension.

B. The pairwise-dominated regime

In the regime where most cross-boundary correlations are weak — which is ensured at large N by the monogamy constraint, since each feature's budget Λ is shared among $O(N)$ partners, giving $\lambda_{ij} \sim \Lambda/N$ — the function f is approximately linear:

$$f(\lambda) \approx c_1 \lambda + O(\lambda^2). \quad (18)$$

The multi-body corrections (genuine multi-partite entanglement involving three or more features simultaneously) are suppressed as $O(1/N)$ relative to the pairwise terms. Therefore:

$$S_{EE}(\Sigma) = c_1 \cdot \mathcal{P}(\Sigma) + O(1/N). \quad (19)$$

This identifies the cross-boundary correlation $\mathcal{P}(\Sigma)$ with the entanglement entropy S_{EE} , up to normalization and corrections that vanish at large N .

C. Where the identification is exact

The identification $\mathcal{P} \propto S_{EE}$ is *exact* in three important cases:

- (i) **Tensor networks.** In discrete tensor network models of holography (e.g., the HaPPY code [16]), each bond carries a fixed entanglement $\ln d$. The entanglement entropy of a boundary region equals the number of bonds cut by the minimal separating surface, times $\ln d$. This is precisely $\mathcal{P}(\Sigma)$ with uniform bond weights.
- (ii) **Free (Gaussian) field theories.** For Gaussian states, the entanglement entropy is entirely determined by the two-point correlation function [17], which maps directly to the pairwise λ_{ij} . Multi-body corrections vanish identically.
- (iii) **Holographic CFTs at leading order in $1/G_N$.** The Ryu-Takayanagi entropy is a geometric (area-based) quantity that captures the leading-order contribution to entanglement entropy. This leading-order contribution is pairwise-dominated [18].

D. Multi-body corrections as irreducible content

The discrepancy between $\mathcal{P}(\Sigma)$ and S_{EE} at finite N measures genuine multi-partite entanglement — correlations that cannot be decomposed into pairwise contributions. These corrections, while important for precision, do not affect the qualitative structure of the thermodynamic deficit theorem: the first law still fails as $\Delta \propto S_{EE} \cdot A$ whether S_{EE} is computed from \mathcal{P} alone or includes multi-body terms.

E. The rewritten deficit

With the identification (19), the thermodynamic deficit theorem becomes:

$$\Delta_{FL}(\Sigma) \propto S_{EE}(\Sigma) \cdot \langle \mathcal{A} \rangle_{\Sigma}. \quad (20)$$

The first law fails in proportion to the entanglement entropy across the surface, weighted by the entanglement anisotropy. This is the central result of this paper.

V. DERIVATION OF THE RYU-TAKAYANAGI FORMULA

A. Minimal P surfaces

Given a boundary region A in a holographic setting, consider all bulk surfaces Σ homologous to A (i.e., sharing the boundary ∂A and continuously deformable to A). The surface that correctly computes the entanglement entropy of A is the one that minimizes the cross-boundary correlation:

$$S_{EE}(A) = c_1 \cdot \min_{\Sigma \sim A} \mathcal{P}(\Sigma). \quad (21)$$

The minimization follows from strong subadditivity of entanglement entropy: the surface that captures the actual entanglement content is the minimum cut through the correlation network. Any other surface overcounts by including correlations internal to one side.

B. Area law in the vacuum

For the vacuum state of a conformal field theory, the correlation function falls off as a power law:

$$\lambda(x, y) \sim \frac{\Lambda_0}{|x - y|^{2\Delta}}, \quad (22)$$

where Δ is the scaling dimension. The cross-boundary correlation integral is dominated by short-distance pairs near the surface, giving

$$\mathcal{P}(\Sigma) \sim \text{Area}(\Sigma) \cdot \Lambda_{UV}^{d-2}, \quad (23)$$

where Λ_{UV} is the UV cutoff. In the vacuum, the correlation density is uniform, so minimizing $\mathcal{P}(\Sigma)$ is equivalent to minimizing $\text{Area}(\Sigma)$:

$$S_{EE}(A) = \frac{1}{4G_N} \min_{\Sigma \sim A} \text{Area}(\Sigma). \quad (24)$$

This is the Ryu-Takayanagi formula [8], with the identification $c_1 \Lambda_0 \Lambda_{UV}^{d-2} \rightarrow 1/(4G_N)$.

C. Non-vacuum generalization

In non-vacuum states, the correlation density is non-uniform. The cross-boundary correlation integral receives a bulk contribution from correlations between features deep inside I and deep inside O that happen to thread through Σ . The correct expression is:

$$S_{EE}(A) = \min_{\Sigma \sim A} \left[\frac{\text{Area}(\Sigma)}{4G_N} + S_{bulk}(\Sigma) \right], \quad (25)$$

where S_{bulk} captures the non-geometric contribution from bulk entanglement. This is the *quantum extremal surface* prescription of Engelhardt and Wall [19], derived here from the structure of $\mathcal{P}(\Sigma)$ in a non-uniform correlation field.

The derivation makes clear why the quantum correction S_{bulk} appears: it represents correlations whose spatial distribution does not follow the vacuum area law but instead reflects the specific entanglement pattern of the excited state.

D. The RT surface as approximate correlation boundary

The Ryu-Takayanagi surface γ_A is the surface that minimizes \mathcal{P} , not the surface where \mathcal{P} vanishes. It is the *best approximation to a correlation boundary* for the given boundary region:

$$\gamma_A = \arg \min_{\Sigma \sim A} \mathcal{P}(\Sigma). \quad (26)$$

This perspective places the RT surface in a hierarchy between true correlation boundaries (horizons, where $\mathcal{P} = 0$) and ordinary surfaces (where $\mathcal{P} \sim O(1)$):

Surface type	$\mathcal{P}(\Sigma)$	Thermodynamic?
Horizon	0	Exactly
RT surface (symmetric region)	$\min, A \approx 0$	Approximately
RT surface (generic region)	$\min, A \neq 0$	With corrections
Stretched horizon	$\varepsilon \approx 0$	Approximately
Ordinary (spherical)	$O(1), A = 0$	Degenerate
Ordinary (generic)	$O(1), A \neq 0$	No

A new prediction follows: the thermodynamic properties of RT surfaces depend on the *symmetry of the boundary region*, not just the bulk geometry. Highly symmetric boundary regions (e.g., half-spaces) produce RT surfaces with small anisotropy and good thermodynamic behavior. Irregular boundary regions produce RT surfaces with larger anisotropy and greater thermodynamic deficit.

VI. RESOLUTION OF WANG-BRAUNSTEIN

A. Summary of their result

Wang and Braunstein [7] consider static asymptotically flat spacetimes with Killing vector ξ . For a surface Σ of

constant Newtonian potential φ , they test whether perturbations satisfy the first law in the form

$$\delta M = \frac{\kappa}{8\pi} \delta A + \text{matter terms}, \quad (27)$$

where κ is the surface gravity and A the area. They find:

- (a) On horizons, the first law holds exactly.
- (b) On stretched horizons, it holds to excellent approximation.
- (c) On ordinary surfaces, additional terms appear that depend on metric perturbations k_1 , which cannot be eliminated by any choice of surface perturbation.
- (d) Exception: for spherically symmetric configurations, the first law holds on any surface of constant φ .

B. Resolution from the correlation boundary formalism

Each of these results follows from Theorem 1 (Eq. 12):

Result (a): Horizons. A horizon is a correlation boundary: $P = 0$. By Eq. (12), $\Delta_{\text{FL}} = 0$ regardless of A . The first law holds exactly.

Result (b): Stretched horizons. A stretched horizon has $P(\Sigma) = \varepsilon \ll 1$. The deficit is $\Delta_{\text{FL}} \sim \varepsilon \cdot |A| \cdot h$, which is small. The first law holds approximately, with corrections of order ε .

Result (c): Ordinary surfaces. For generic surfaces, $P \sim O(1)$ and $A \sim O(1)$. The deficit Δ_{FL} is order unity. The first law fails. The "unwanted terms" in Wang-Braunstein's Eq. (3) correspond to the $\delta\Phi_{\text{cross}}$ contribution from cross-boundary correlations responding to the perturbation.

Result (d): Spherical symmetry. This is the most interesting case and requires detailed analysis.

C. The spherical symmetry degeneracy

For a spherically symmetric mass distribution, every surface of constant φ (i.e., every concentric sphere) has the following property: the correlations threading through it are *isotropic*.

At any point x on the sphere, the correlations from x to outside features are distributed uniformly over solid angle (by spherical symmetry). The correlation current J has no preferred angular direction; it is purely radial. Therefore the anisotropy $A(x)$ — the directional asymmetry of correlations — decomposes into spherical harmonics with:

$$\mathcal{A}_{lm} = 0 \quad \text{for all } l \geq 1. \quad (28)$$

The $l = 0$ (monopole) component may be nonzero, but it has the same angular structure as the standard thermodynamic terms δQ and TdS . It can therefore be absorbed into redefined thermodynamic quantities T' and S' (see Supporting Information Sec. S6 for the complete absorption argument):

$$\delta Q = T' dS', \quad (29)$$

where T' and S' incorporate the monopole contribution from $\delta\Phi_{\text{cross}}$. The first law holds with *renormalized* thermodynamic quantities.

D. Breaking spherical symmetry: multipole corrections

For a nearly spherical mass distribution with multipole moments Q_l , the anisotropy acquires nonzero $l \geq 1$ components:

$$\mathcal{A}_{lm} \propto Q_l \cdot r^{-(l+1)}. \quad (30)$$

These components have angular dependence that cannot be absorbed into scalar thermodynamic quantities. The thermodynamic deficit becomes:

$$\langle |\Delta_{FL}|^2 \rangle_{\Sigma} \propto \sum_{l \geq 2} \frac{Q_l^2}{r^{2(l+1)}} \cdot \mathcal{P}(r), \quad (31)$$

where the $l = 1$ (dipole) term vanishes for surfaces centered on the center of mass.

The dominant correction is the quadrupole ($l = 2$):

$$\langle |\Delta_{FL}|^2 \rangle \sim \frac{Q_2^2}{r^6} \cdot \mathcal{P}(r). \quad (32)$$

This is a *testable prediction*: the thermodynamic deficit on holographic screens around oblate or prolate mass distributions should scale with the quadrupole moment squared (see Supporting Information Sec. S7 for the detailed multipole analysis).

E. Why this resolves the tension

Wang and Braunstein's result was interpreted as evidence *against* emergent gravity. The correlation boundary formalism shows that it is instead evidence *for a more refined version* of emergent gravity. The original programs of Jacobson [4] and Verlinde [5] both involve thermodynamic properties of surfaces, but they differ in which surfaces:

- **Jacobson** applies $\delta Q = TdS$ at local Rindler horizons — surfaces that *are* correlation boundaries by construction. His derivation is unaffected by Wang-Braunstein.
- **Verlinde** extends thermodynamic properties to arbitrary holographic screens — surfaces that are *not* correlation boundaries. Wang-Braunstein correctly identifies that this extension fails.

The resolution is that gravitational thermodynamics is a property of *correlation boundaries specifically*, not of surfaces in general. The thermodynamic structure exists everywhere in the correlation field (as a geometric relation), but the thermodynamic *interpretation* — temperature, entropy, first law — applies only at surfaces where the partition into inside and outside is informationally clean ($P = 0$) or degenerate by symmetry ($A = 0$).

VII. DISCUSSION

A. Relation to Jacobson's entanglement equilibrium

Jacobson [20] argued that the Einstein equations follow from the condition that entanglement entropy across local Rindler horizons is in equilibrium. In our language, this is the statement that cross-boundary correlation P is stationary at correlation boundaries:

$$\delta\mathcal{P}(\Sigma_H) = 0 \quad \text{for all perturbations of the horizon } \Sigma_H. \quad (33)$$

Since $P = 0$ at the horizon, stationarity means that P remains zero under small perturbations — the horizon is a *stable* correlation boundary. This condition, imposed at every point, constrains the geometry and reproduces the Einstein equations.

The connection between $P = S_{EE}$ and Jacobson's entanglement equilibrium makes this precise: the equilibrium condition (33) is the entanglement first law [21], applied at the special surface where entanglement across the partition vanishes.

B. Implications for the emergent gravity program

The results of this paper suggest that the emergent gravity program can be *salvaged* by restricting thermodynamic assumptions to correlation boundaries:

Strong version (Jacobson). Apply $\delta Q = TdS$ only at local Rindler horizons. This remains valid and leads to the Einstein equations.

Modified Verlinde. The geometric structure underlying Verlinde's derivation — the relationship between energy, entropy, and surface properties — exists everywhere. But the thermodynamic *interpretation* should be restricted to surfaces with high thermodynamic index (Sec. III). This preserves Verlinde's mathematics where it works while correctly predicting where it fails.

Verlinde's dark matter proposal. The cosmological horizon is a true correlation boundary (it defines the limit of causal correlation in de Sitter space). The thermodynamic properties Verlinde assigns to it [6] are legitimate. Whether the quantitative predictions for galaxy rotation curves are correct is a separate question, but the cosmological horizon is not affected by Wang-Braunstein's objection.

C. The non-thermodynamic contribution to gravity

The cross-boundary correlation response $\delta\Phi_{\text{cross}}$ represents a contribution to gravitational dynamics that does

not have a thermodynamic interpretation. It is the energy associated with correlations threading through surfaces, responding to perturbations through channels that bypass the surface flux.

For an observer who assumes all gravitational dynamics are thermodynamic (i.e., who assumes $\delta Q = TdS$ holds on all surfaces), this non-thermodynamic contribution would appear as unaccounted energy — dynamics that cannot be attributed to visible matter through standard gravitational accounting. This suggests a natural research direction connecting $\delta\Phi_{\text{cross}}$ to modified gravity phenomenology, though developing this connection quantitatively requires further work.

D. Computational testability

The thermodynamic deficit theorem makes predictions that can be tested in computational systems:

- (i) In tensor networks, define surfaces and compute $P(\Sigma)$ and $A(x)$. Verify that thermodynamic behavior (well-defined temperature, first-law compliance) correlates with low P and low A .
- (ii) In lattice gauge theories, compute the anisotropy of entanglement entropy distribution across surfaces of different geometry. The deficit should scale as $S_{EE} \cdot A$.
- (iii) In holographic models, test the prediction that RT surface thermodynamic properties depend on boundary region symmetry.

E. Open questions

Several directions for future work emerge:

- (i) **Quantitative mapping.** The derivation establishes $\Delta \propto P \cdot A$ but does not fix the proportionality constant. This requires specifying the relationship between the correlation field λ_{ij} and the gravitational variables (metric, curvature).
- (ii) **Volume-law entropy.** Verlinde's 2016 proposal distinguishes area-law entropy (horizons) from volume-law entropy (cosmological de Sitter space). The correlation boundary formalism should accommodate this distinction, potentially through the relationship between P and the bulk correlation density.
- (iii) **Finite- N corrections.** At small N , the distinction between correlation boundary and ordinary surface becomes fuzzy (correlation strengths are discrete). Developing the formalism for mesoscopic N could connect to quantum gravity corrections in the Bekenstein-Hawking entropy.
- (iv) **Derivation from first principles.** This paper takes the monogamy constraint (1) as given. A deeper question is whether it can be derived from more primitive principles — for instance, from conditions on what must hold for physical degrees of freedom to be distinguishable from each other.

VIII. CONCLUSION

We have introduced the concept of correlation boundaries — surfaces across which the total pairwise correlation vanishes — and shown that these are the only surfaces on which the first law of thermodynamics

holds generically. The thermodynamic deficit of any surface is controlled by two quantities: the cross-boundary entanglement entropy $P \propto S_{EE}$ and the entanglement anisotropy A . The product $P \cdot A$ determines the magnitude of the deficit.

This result resolves the tension identified by Wang and Braunstein between the emergent gravity program and the non-thermodynamic behavior of holographic screens. Horizons are thermodynamic because they are correlation boundaries. Stretched horizons are approximately thermodynamic because they are near-correlation boundaries. Spherically symmetric screens are thermodynamic by a degeneracy — isotropic entanglement mimics a correlation boundary. Generic screens fail because they have nonzero entanglement that is anisotropically distributed.

The identification $P = S_{EE}$ connects this formalism to the holographic entanglement entropy program, with the Ryu-Takayanagi formula emerging as the minimization of cross-boundary correlation over surfaces homologous to a boundary region. The quantum extremal surface prescription follows from non-uniformity of the correlation field in excited states.

These results follow from a minimal set of assumptions — pairwise correlations with finite capacity — making them independent of specific microscopic theories and applicable across quantum gravity, condensed matter, and quantum information contexts.

ACKNOWLEDGMENTS

Goleudy.ai

REFERENCES

- [1] J. M. Bardeen, B. Carter, and S. W. Hawking, "The four laws of black hole mechanics," *Commun. Math. Phys.* **31**, 161 (1973).
- [2] J. D. Bekenstein, "Black holes and entropy," *Phys. Rev. D* **7**, 2333 (1973).
- [3] S. W. Hawking, "Particle creation by black holes," *Commun. Math. Phys.* **43**, 199 (1975).
- [4] T. Jacobson, "Thermodynamics of Spacetime: The Einstein Equation of State," *Phys. Rev. Lett.* **75**, 1260 (1995).
- [5] E. Verlinde, "On the Origin of Gravity and the Laws of Newton," *JHEP* **2011**, 29 (2011). arXiv:1001.0785.
- [6] E. Verlinde, "Emergent Gravity and the Dark Universe," *SciPost Phys.* **2**, 016 (2017). arXiv:1611.02269.
- [7] Z.-W. Wang and S. L. Braunstein, "Surfaces away from horizons are not thermodynamic," *Nat. Commun.* **9**, 2977 (2018).

- [8] S. Ryu and T. Takayanagi, "Holographic Derivation of Entanglement Entropy from AdS/CFT," *Phys. Rev. Lett.* **96**, 181602 (2006).
- [9] V. Coffman, J. Kundu, and W. K. Wootters, "Distributed entanglement," *Phys. Rev. A* **61**, 052306 (2000).
- [10] T. J. Osborne and F. Verstraete, "General Monogamy Inequality for Bipartite Qubit Entanglement," *Phys. Rev. Lett.* **96**, 220503 (2006).
- [11] G. Vidal, "Entanglement Renormalization," *Phys. Rev. Lett.* **99**, 220405 (2007).
- [12] B. M. Terhal, "Is entanglement monogamous?," *IBM J. Res. Dev.* **48**, 71 (2004).
- [13] A. Lewkowycz and J. Maldacena, "Generalized gravitational entropy," *JHEP* **2013**, 90 (2013).
- [14] M. Van Raamsdonk, "Building up spacetime with quantum entanglement," *Gen. Rel. Grav.* **42**, 2323 (2010).
- [15] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes," *Fortschr. Phys.* **61**, 781 (2013).
- [16] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, "Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence," *JHEP* **2015**, 149 (2015).
- [17] I. Peschel, "Calculation of reduced density matrices from correlation functions," *J. Phys. A* **36**, L205 (2003).
- [18] T. Faulkner, A. Lewkowycz, and J. Maldacena, "Quantum corrections to holographic entanglement entropy," *JHEP* **2013**, 74 (2013).
- [19] N. Engelhardt and A. C. Wall, "Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime," *JHEP* **2015**, 73 (2015).
- [20] T. Jacobson, "Entanglement Equilibrium and the Einstein Equation," *Phys. Rev. Lett.* **116**, 201101 (2016).
- [21] D. D. Blanco, H. Casini, L.-Y. Hung, and R. C. Myers, "Relative entropy and holography," *JHEP* **2013**, 60 (2013).