Supporting Information: Full Proof of the Thermodynamic Deficit Theorem

For: "Correlation Boundaries and the First Law"

Overview

This Supporting Information provides the complete proof of Theorem 1 from the main text, which states that the thermodynamic deficit on a surface Σ is

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

The proof proceeds through five stages:

- **Stage 1** (Sec. S1): Decomposition of the system potential across a partition
- **Stage 2** (Sec. S2): Response of the cross term to surface perturbations
- **Stage 3** (Sec. S3): Identification of the anisotropy factor
- **Stage 4** (Sec. S4): Assembly of the theorem and error bounds
- **Stage 5** (Sec. S5): Verification of special cases (horizons, stretched horizons, spherical symmetry)

Supplementary sections cover the harmonic absorption argument (Sec. S6), the multipole prediction (Sec. S7), and the connection to entanglement entropy (Sec. S8).

S1. Decomposition of the System Potential

S1.1 Setup

Consider N features {1, ..., N} with pairwise correlation field λ_ij ∈ [0, Λ] satisfying the monogamy constraint

$$\sum_{j \neq i} \lambda_{ij} \leq \Lambda \quad \forall\, i. \tag{S1}$$

Each feature carries a local state characterized by some internal configuration. The total system state determines a *potential* Φ_total that quantifies the system's thermodynamic character. We

require only the following structural properties of Φ:

(P1) Additivity under independence. If two subsystems have no correlations between them (λ_ij = 0 for all i in one and j in the other), then Φ_total = Φ_1 + Φ_2.

(P2) Smoothness. Φ is a smooth function of the correlation field {λ_ij}.

(P3) Monotonicity. Φ depends monotonically on the accessible configurations (increasing Ω increases Φ) and inversely on the descriptive complexity (increasing K decreases Φ).

These properties are satisfied by a wide class of potentials, including the von Neumann entropy, Helmholtz free energy, and the efficiency potential Φ = ln(Ω/K). The proof does not require a specific choice.

S1.2 The partition

Let Σ be a surface partitioning the features into Inside I = {i_1, ..., i_n} and Outside O = {j_1, ..., j_{N-n}}. Define:

- **Internal correlations of I:** {λ_{ab} : a, b ∈ I}
- **Internal correlations of O:** {λ_{ab} : a, b ∈ O}
- **Cross-boundary correlations:** {λ_{ij} : i ∈ I, j ∈ O}

S1.3 Decomposition

Lemma S1 (Potential Decomposition). Under properties (P1)-(P2), the total potential decomposes as

$$\Phi_{total} = \Phi_I + \Phi_O + \Phi_{cross}(\{\lambda_{ij}\}_{i\in I, j\in O}), \tag{S2}$$

where Φ_I depends only on the internal state and correlations of I, Φ_O depends only on those of O, and Φ_cross depends on the cross-boundary correlations.

Proof. Consider the family of correlation fields parameterized by t ∈ [0, 1]:

$$\lambda_{ij}(t) = t \cdot \lambda_{ij} \quad \text{for } i \in I, j \in O, \tag{S3}$$

with all internal correlations held fixed. At t = 0, there are no cross-boundary correlations, so by (P1):

$$\Phi_{total}(t = 0) = \Phi_I + \Phi_O. \tag{S4}$$

By (P2), Φ_total is smooth in t. Define:

$$\Phi_{cross} \equiv \Phi_{total}(t=1) - \Phi_{total}(t=0) = \int_0^1 \frac{d\Phi_{total}}{dt}\, dt. \tag{S5}$$

This gives Φ_total = Φ_I + Φ_O + Φ_cross, where Φ_cross depends on all cross-boundary correlations through the integral. ■

S1.4 First-order expansion of Φ_cross

By the fundamental theorem of calculus applied to (S5):

$$\Phi_{cross} = \int_0^1 \sum_{i\in I, j\in O} \left.\frac{\partial \Phi_{total}}{\partial \lambda_{ij}}\right|_t \cdot \lambda_{ij}\, dt. \tag{S6}$$

Define the *correlation susceptibility*:

$$\phi_{ij}(t) \equiv \left.\frac{\partial \Phi_{total}}{\partial \lambda_{ij}}\right|_t. \tag{S7}$$

This measures how much the potential responds to a unit change in the (i,j) correlation. At t = 0 (no cross-boundary correlations), define φ⁰_ij ≡ φ_ij(0). Then:

$$\Phi_{cross} = \sum_{i\in I, j\in O} \lambda_{ij}\cdot \phi^0_{ij} + \frac{1}{2}\sum_{i,j}\sum_{k,l} \lambda_{ij}\lambda_{kl}\cdot \left.\frac{\partial \phi_{ij}}{\partial \lambda_{kl}}\right|_0 + O(\lambda^3). \tag{S8}$$

The first-order approximation:

$$\Phi_{cross} = \sum_{i\in I, j\in O} \lambda_{ij}\cdot \phi^0_{ij} + O(\lambda^2), \tag{S9}$$

is valid when individual cross-boundary correlations are small compared to Λ. The monogamy constraint (S1) ensures that at large N, each λ_ij ~ Λ/(N-1), making this regime natural.

S2. Response to Surface Perturbations

S2.1 Surface displacement

Consider a one-parameter family of surfaces Σ(ε) obtained by displacing each point x ∈ Σ by ε · δε(x) in the outward normal direction n̂(x), where δε(x) is a smooth profile function on Σ. At ε = 0, we recover the original surface Σ(0) = Σ.

The displacement changes the partition: features in a thin shell of width ε · δε(x) near the surface transfer from O to I (for outward displacement where δε > 0) or from I to O (for inward displacement where δε < 0).

S2.2 The reassigned features

For an infinitesimal displacement ε, the reassigned features near point x ∈ Σ occupy a shell of volume

$$dV_{shell}(x) = \varepsilon \cdot \delta\varepsilon(x) \cdot d\mu_{\Sigma}(x), \tag{S10}$$

where dμ_Σ is the surface measure. In the continuum, the number of features in this shell is

$$dN_{shell}(x) = n(x) \cdot dV_{shell}(x), \tag{S11}$$

where n(x) is the feature density near the surface.

For outward displacement (δε > 0): features in the shell move from O to I.
For inward displacement (δε < 0): features in the shell move from I to O.

S2.3 Change in the cross-boundary correlation

When a feature α at position x near the surface transfers from O to I:

Correlations that become internal to I: All λ_{αi} for i ∈ I. These were previously cross-boundary; they are now internal to I. This *reduces* P by Σ_{i∈I} λ_{αi}.

Correlations that become cross-boundary: All λ_{αj} for j ∈ O, j ≠ α. These were previously internal to O; they are now cross-boundary. This *increases* P by Σ_{j∈O} λ_{αj}.

The net change in cross-boundary correlation from reassigning feature α is:

$$\delta\mathcal{P}_{\alpha} = \sum_{j \in O} \lambda_{\alpha j} - \sum_{i \in I} \lambda_{\alpha i}. \tag{S12}$$

S2.4 Directional structure

We now express (S12) in terms of the directional distribution of correlations. For feature α at position x near the surface, the correlations with all other features can be decomposed by direction:

$$\sum_{j \in O} \lambda_{\alpha j} = \sum_{j} \lambda_{\alpha j} \cdot 1[j \in O], \tag{S13}$$

$$\sum_{i \in I} \lambda_{\alpha i} = \sum_{j} \lambda_{\alpha j} \cdot 1[j \in I]. \tag{S14}$$

For a feature α very close to the surface Σ, whether a partner j is in I or O depends on the angle between the vector from α to j and the surface normal n̂(x). Specifically, in the local neighborhood of x:

- Partners on the outward side of Σ (r̂_{αj} · n̂ > 0) are in O
- Partners on the inward side of Σ (r̂_{αj} · n̂ < 0) are in I

This is a local approximation valid for partners near the surface; distant partners' membership in I or O depends on the global surface topology, not just the local normal. However, because correlations fall off with distance (a consequence of the monogamy constraint — each feature's budget Λ is shared among all partners), the dominant contributions to (S12) come from nearby partners, where the local approximation holds.

Define the *outward* and *inward* correlation sums at x:

$$\mathcal{C}^{out}(x) = \sum_{j} \lambda_{xj} \cdot \Theta(\hat{r}_{xj} \cdot \hat{n}), \tag{S15}$$

$$\mathcal{C}^{in}(x) = \sum_{j} \lambda_{xj} \cdot \Theta(-\hat{r}_{xj} \cdot \hat{n}), \tag{S16}$$

where Θ is the Heaviside step function. The net change in P from reassigning one feature at x is:

$$\delta\mathcal{P}(x) = \mathcal{C}^{out}(x) - \mathcal{C}^{in}(x). \tag{S17}$$

S2.5 Change in Φ_cross

The change in Φ_cross due to the surface displacement involves two effects:

(Effect 1: Topological) The set of cross-boundary pairs changes. Correlations involving reassigned features switch status (cross-boundary ↔ internal). From (S9), this contributes:

$$\delta\Phi^{(1)}_{cross} = \sum_{\alpha \in \text{shell}} \left[\sum_{j \in O'} \lambda_{\alpha j} \cdot \phi^0_{\alpha j} - \sum_{i \in I} \lambda_{\alpha i} \cdot \phi^0_{\alpha i} \right], \tag{S18}$$

where O' = O \ {shell} is the set of features remaining outside after reassignment.

(Effect 2: Geometric) The susceptibilities φ_ij change because the partition has changed. For correlations that remain cross-boundary, their contribution to Φ_cross shifts:

$$\delta\Phi^{(2)}_{cross} = \sum_{\substack{i\in I, j\in O \\ \text{not reassigned}}} \lambda_{ij} \cdot \delta\phi_{ij}. \tag{S19}$$

This second effect is higher order: δφ_ij is of order ε (the perturbation size) while the sum is over O(N²) terms with individual λ_ij ~ O(1/N). The total contribution is O(ε/N) relative to Effect 1, which is O(ε). At large N, Effect 2 is subleading.

Retaining only Effect 1:

$$\delta\Phi_{cross} = \varepsilon \int_\Sigma n(x)\, \delta\varepsilon(x) \left[\sum_j \lambda_{xj}\, \phi^0_{xj} \cdot \mathrm{sgn}(\hat{r}_{xj} \cdot \hat{n})\right] d\mu_\Sigma + O(\varepsilon^2, \varepsilon/N). \tag{S20}$$

S3. Identification of the P · A Structure

S3.1 Decomposing the integrand

The bracketed expression in (S20) can be written as:

$$\sum_j \lambda_{xj}\, \phi^0_{xj} \cdot \mathrm{sgn}(\hat{r}_{xj} \cdot \hat{n}) = \sum_j \lambda_{xj}\, \phi^0_{xj} \cdot (\hat{r}_{xj} \cdot \hat{n}) \cdot \frac{\mathrm{sgn}(\hat{r}_{xj} \cdot \hat{n})}{\hat{r}_{xj} \cdot \hat{n}}. \tag{S21}$$

This is not immediately illuminating. Instead, separate the sum into a magnitude factor and a directional factor.

Step 1: Define the *local cross-boundary correlation density* at x:

$$\mathcal{P}(x) = \sum_{j\in O} \lambda_{xj} = \mathcal{C}^{out}(x), \tag{S22}$$

which is the total correlation from x to all outside features. (For x near the surface, this is the contribution of x to the global P(Σ).)

Step 2: Define the *correlation-weighted mean susceptibility*:

$$\bar{\phi}(x) = \frac{\sum_j \lambda_{xj} |\phi^0_{xj}|}{\sum_j \lambda_{xj}}. \tag{S23}$$

This is the average susceptibility, weighted by correlation strength.

Step 3: Rewrite the integrand by separating the isotropic and anisotropic parts. For any function g_j defined on the partners of x:

$$\sum_j \lambda_{xj}\, g_j = \left(\sum_j \lambda_{xj}\right) \cdot \langle g \rangle_\lambda + \sum_j \lambda_{xj}\, (g_j - \langle g \rangle_\lambda), \tag{S24}$$

where ⟨g⟩_λ = (Σ_j λ_xj g_j)/(Σ_j λ_xj) is the correlation-weighted average.

Apply this with g_j = φ°_xj · sgn(r̂_xj · n̂):

$$\sum_j \lambda_{xj}\, \phi^0_{xj} \cdot \mathrm{sgn}(\hat{r}_{xj} \cdot \hat{n}) = \Lambda_{tot}(x) \cdot \langle \phi \cdot \mathrm{sgn} \rangle_\lambda + \text{fluctuation term}, \tag{S25}$$

where Λ_tot(x) = Σ_j λ_xj is the total correlation strength at x.

S3.2 The isotropic and anisotropic contributions

Now consider the weighted average ⟨φ · sgn⟩_λ. This splits into contributions from outside and inside partners:

$$\langle \phi \cdot \mathrm{sgn} \rangle_\lambda = \frac{1}{\Lambda_{tot}} \left[\sum_{j \in O} \lambda_{xj} \phi^0_{xj} - \sum_{i \in I} \lambda_{xi} \phi^0_{xi} \right]. \tag{S26}$$

Case 1: Isotropic correlations. If the correlation field is isotropic at x — meaning the statistical properties of {λ_xj, φ°_xj} are the same in all directions — then the outside and inside sums are proportional to the number of partners on each side. For x on a closed surface separating comparable regions, these sums are approximately equal, giving ⟨φ · sgn⟩_λ ≈ 0.

Case 2: Anisotropic correlations. If the correlation field has a preferred direction at x — say, more strongly correlated partners outside than inside — then ⟨φ · sgn⟩_λ ≠ 0.

Define the *effective anisotropy*:

$$\mathcal{A}_{eff}(x) \equiv \frac{1}{\Lambda_{tot}(x)} \left[\sum_{j \in O} \lambda_{xj} \phi^0_{xj} - \sum_{i \in I} \lambda_{xi} \phi^0_{xi} \right]. \tag{S27}$$

This combines the directional asymmetry of the correlations (which partners are outside vs. inside) with the susceptibility weighting (how much each correlation contributes to the

potential).

S3.3 Simplification under uniform susceptibility

In the general case, the effective anisotropy A_eff (Eq. S27) couples the directional asymmetry of correlations to the susceptibility weighting φ⁰_xj, producing a factorization where the "anisotropy" includes both geometric and susceptibility contributions. The qualitative structure — deficit requires both nonzero P and nonzero directional asymmetry — is unchanged, but the quantitative expression is more complex. We defer the full non-uniform susceptibility analysis to future work and proceed with the uniform case, which captures the essential structure and all qualitative conclusions of the theorem.

In the important special case where the susceptibility is approximately uniform — φ⁰_xj ≈ φ̄ for all partners j of x — the effective anisotropy simplifies to:

$$\mathcal{A}_{eff}(x) \approx \bar{\phi} \cdot \frac{\mathcal{C}^{out}(x) - \mathcal{C}^{in}(x)}{\Lambda_{tot}(x)} = \bar{\phi} \cdot \frac{\delta\mathcal{P}(x)}{\Lambda_{tot}(x)}. \tag{S28}$$

The numerator is the net directional asymmetry of correlations — precisely the correlation anisotropy A(x) defined in the main text (Eq. 11). So:

$$\mathcal{A}_{eff}(x) \approx \bar{\phi} \cdot \frac{\mathcal{A}(x)}{\Lambda_{tot}(x)}. \tag{S29}$$

S3.4 Assembly

Substituting back into (S20):

$$\delta\Phi_{cross} = \varepsilon \int_{\Sigma} n(x)\, \delta\varepsilon(x) \cdot \Lambda_{tot}(x) \cdot \bar{\phi} \cdot \frac{\mathcal{A}(x)}{\Lambda_{tot}(x)}\, d\mu_{\Sigma} + O(\varepsilon^2). \tag{S30}$$

The Λ_tot factors cancel:

$$\delta\Phi_{cross} = \varepsilon\, \bar{\phi} \int_{\Sigma} n(x)\, \delta\varepsilon(x) \cdot \mathcal{A}(x)\, d\mu_{\Sigma} + O(\varepsilon^2). \tag{S31}$$

Now, the anisotropy A(x) is nonzero only where cross-boundary correlations exist — at points where P(x) = 0, the anisotropy is trivially zero (no correlations to be anisotropic). To make the P-dependence explicit, write:

$$\mathcal{A}(x) = \frac{\mathcal{P}(x)}{\mathcal{P}(x)} \cdot \mathcal{A}(x) \equiv \mathcal{P}(x) \cdot \hat{\mathcal{A}}(x), \tag{S32}$$

where $\hat{A}(x) = A(x)/P(x)$ is the *anisotropy per unit cross-boundary correlation* — a normalized measure of directional asymmetry that is well-defined wherever $P(x) > 0$.

Alternatively, and more directly, note that $A(x)$ as defined in (S27) can be bounded:

$$|\mathcal{A}_{eff}(x)| \leq \bar{\phi} \cdot \frac{|\mathcal{C}^{out} - \mathcal{C}^{in}|}{\Lambda_{tot}} \leq \bar{\phi} \cdot \frac{\min(\mathcal{C}^{out}, \mathcal{C}^{in})}{\Lambda_{tot}} \leq \bar{\phi}. \tag{S33}$$

And A_eff is zero when either (a) C^out = C^in = 0 (i.e., P(x) = 0), or (b) C^out = C^in (isotropic). This gives the product structure: the deficit requires both nonzero P *and* nonzero directional asymmetry.

Defining $h(x) = \varepsilon \cdot n(x) \cdot \delta\varepsilon(x) \cdot \bar{\varphi}$ as the perturbation-dependent factor, and keeping both P and A explicit:

$$\boxed{\delta\Phi_{cross} = \int_{\Sigma} \mathcal{P}(x) \cdot \hat{\mathcal{A}}(x) \cdot h(x)\, d\mu_{\Sigma} + O(\varepsilon^2, 1/N)} \tag{S34}$$

This is Theorem 1.

S4. Error Analysis and Domain of Validity

S4.1 Sources of error

The derivation introduces approximations at three points:

Approximation 1: First-order expansion (S9). The cross-boundary potential Φ_{cross} is expanded to first order in the cross-boundary correlations λ_{ij}. The error is $O(\lambda^2)$, i.e., $O(\Lambda^2/N^2)$ per pair, summed over $O(n(N-n))$ pairs, giving total error $O(\Lambda^2 n(N-n)/N^2)$. For a balanced partition ($n \sim N/2$), this is $O(\Lambda^2/4)$.

This approximation is valid when individual cross-boundary correlations are weak: $\lambda_{ij} \ll \Lambda$. The monogamy constraint ensures this at large N but not at small N. For $N \sim O(1)$, the second-order term in (S8) can be comparable to the first-order term.

Approximation 2: Local partner classification (S15-S16). The classification of partners as inside/outside based on the local surface normal is exact for partners near x but approximate for distant partners. The error affects partners at distance r from x comparable to or larger than the surface curvature radius R_{Σ}.

Since correlations decay with distance (due to monogamy — a feature's budget is shared among all partners), the dominant contributions come from partners within a correlation length ξ. The

approximation is valid when $\xi \ll R_\Sigma$ — i.e., the surface is smooth on the correlation scale.

Approximation 3: Neglect of Effect 2 (S19). The geometric effect (change in susceptibilities due to the partition shift) is $O(\varepsilon/N)$ relative to the topological effect. This is valid at large N.

S4.2 Combined validity condition

The theorem is valid in the regime:

$$N \gg 1, \quad \xi \ll R_\Sigma, \quad \varepsilon \ll R_\Sigma. \tag{S35}$$

This is the standard *thermodynamic* regime: many degrees of freedom, smooth surfaces, small perturbations. The theorem breaks down at small N (where the first-order expansion fails), at surfaces with sharp features (where the local classification fails), or for large perturbations (where nonlinear effects dominate).

S4.3 Bound on the deficit

From (S33), the deficit is bounded by:

$$|\Delta_{FL}(\Sigma)| \leq \bar{\phi} \cdot \varepsilon \int_\Sigma n(x)\, |\delta\varepsilon(x)| \cdot \mathcal{P}(x)\, d\mu_\Sigma, \tag{S36}$$

with equality when the anisotropy is maximal (all correlations point in one direction). The deficit vanishes when $P(x) = 0$ everywhere (correlation boundary) and is bounded by the total cross-boundary correlation times the perturbation size.

S5. Verification of Special Cases

S5.1 Correlation boundary (P = 0)

If $P(x) = 0$ for all $x \in \Sigma$, then by (S34):

$$\delta\Phi_{cross} = \int_\Sigma \underbrace{\mathcal{P}(x)}_{=0} \cdot \hat{\mathcal{A}}(x) \cdot h(x)\, d\mu_\Sigma = 0. \tag{S37}$$

The first law holds exactly: $\Delta_{FL} = 0$. ■

This result is *exact* — it does not depend on the first-order approximation, because $\Phi_{cross} = 0$ exactly when $P = 0$ (by Lemma S1 and property P1). There are no cross-boundary correlations to expand, so no expansion error arises.

S5.2 Approximate correlation boundary (P ≈ 0)

If P(x) < ε_P ≪ 1 for all x ∈ Σ, then:

$$|\Delta_{FL}| \leq \bar{\phi} \cdot \varepsilon \cdot \varepsilon_P \int_\Sigma n(x)\, |\delta\varepsilon(x)|\, d\mu_\Sigma = O(\varepsilon \cdot \varepsilon_P). \tag{S38}$$

The first law holds to accuracy ε_P. ■

S5.3 Isotropic correlations (A = 0)

If the correlation field is isotropic at every point x ∈ Σ — meaning C^out(x) = C^in(x) — then A(x) = 0 everywhere. By (S34):

$$\delta\Phi_{cross} = \int_\Sigma \mathcal{P}(x) \cdot \underbrace{\hat{\mathcal{A}}(x)}_{=0} \cdot h(x)\, d\mu_\Sigma = 0. \tag{S39}$$

The first law holds despite P ≠ 0. ■

However, this result is *not* exact in the same way as S5.1. The A = 0 condition eliminates the first-order deficit, but the second-order term in (S8) may be nonzero. The first law holds to first order in ε but may acquire corrections at O(ε²) even for isotropic correlations.

This distinction matters physically: a correlation boundary (P = 0) is *absolutely* thermodynamic, while an isotropic surface (A = 0) is thermodynamic only at leading order. The isotropic case is a degeneracy, not a fundamental property.

S6. The Spherical Symmetry Absorption Argument

This section provides the detailed argument for why spherically symmetric ordinary surfaces satisfy the first law, despite having nonzero cross-boundary correlation.

S6.1 Setup

Consider a spherically symmetric correlation field — one where λ(x, y) depends only on |x| and |y| (the radial positions). Let Σ be a sphere of radius R centered at the origin.

S6.2 Harmonic decomposition of the deficit

The deficit (S34) involves the anisotropy A(x) evaluated on the sphere. Expand A in spherical harmonics:

$$\mathcal{A}(x) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \mathcal{A}_{lm}(R)\, Y_{lm}(\hat{x}). \tag{S40}$$

Claim: For a spherically symmetric correlation field, A_lm = 0 for all l ⩾ 1.

Proof. The anisotropy A(x) at point x on the sphere depends on the distribution of correlations from x to all other features. By spherical symmetry, the correlation from x to a partner at position y depends only on |x|, |y|, and the angle between x and y:

$$\lambda(x, y) = \lambda(|x|, |y|, \hat{x} \cdot \hat{y}). \tag{S41}$$

For x on the sphere at radius R, this becomes λ(R, |y|, x̂ · ŷ).

The classification of partner y as inside (|y| < R) or outside (|y| > R) depends only on |y|, not on ŷ. Therefore:

$$\mathcal{C}^{out}(x) = \int_{|y|>R} \lambda(R, |y|, \hat{x} \cdot \hat{y})\, n(|y|)\, d^3y, \tag{S42}$$

$$\mathcal{C}^{in}(x) = \int_{|y|<R} \lambda(R, |y|, \hat{x} \cdot \hat{y})\, n(|y|)\, d^3y. \tag{S43}$$

Now perform the angular integration over ŷ. Because λ depends on x̂ · ŷ, the angular integral of λ(R, |y|, x̂ · ŷ) over ŷ is independent of x̂ (by rotational invariance of the angular integration domain). This makes both C^out(x) and C^in(x) independent of x̂ — they depend only on R.

Therefore A(x) = C^out(x) − C^in(x) = const on the sphere. In terms of harmonics, only l = 0 is nonzero:

$$\mathcal{A}_{lm} = 0 \quad \text{for } l \geq 1. \tag{S44}$$

■

S6.3 The monopole contribution

The l = 0 term A_{00} is generally nonzero. This means the deficit (S34) does not vanish:

$$\delta\Phi_{cross} = \mathcal{A}_{00} \int_{\Sigma} \mathcal{P}(x) \cdot Y_{00} \cdot h(x)\, d\mu_{\Sigma}. \tag{S45}$$

Since Y_{00} = 1/√(4π) is constant and P(x) = P(R) is constant on the sphere (by the same symmetry argument), this becomes:

$$\delta\Phi_{cross} = \mathcal{P}(R) \cdot \mathcal{A}_{00}(R) \cdot \frac{1}{\sqrt{4\pi}} \int_\Sigma h(x)\, d\mu_\Sigma. \tag{S46}$$

The integral ∫h dμ has the same form as the standard thermodynamic terms δQ and TdS evaluated on the sphere. Both of these are also integrals of scalar quantities over the sphere — they have only l = 0 content by the symmetry of the perturbation (for perturbations that respect the spherical symmetry, such as uniform radial displacements).

S6.4 Absorption into redefined thermodynamic quantities

The perturbed energy balance is:

$$\delta Q = T\, dS + \delta\Phi_{cross}. \tag{S47}$$

For the spherically symmetric case, all three terms have the same angular structure (l = 0 only). Define:

$$T' = T + T \cdot \frac{\partial(\delta\Phi_{cross}/\delta Q)}{\partial S}, \tag{S48}$$

$$S' = S + \frac{\delta\Phi_{cross}}{T}. \tag{S49}$$

More precisely, the redefinition works as follows. We can write:

$$\delta Q = T\, dS + \mathcal{P}(R) \cdot \mathcal{A}_{00}(R) \cdot g(R) \cdot \delta\varepsilon, \tag{S50}$$

where g(R) absorbs the integration constants. Since δQ itself is proportional to δε (it's a linear response), the extra term is proportional to δε with a coefficient that depends only on R. We can therefore absorb it:

$$\delta Q = \left[T + \frac{\mathcal{P}(R) \cdot \mathcal{A}_{00}(R) \cdot g(R)}{dS/d\varepsilon}\right] dS \equiv T'(R)\, dS. \tag{S51}$$

The first law holds with a *renormalized temperature* T'(R) that differs from the naive T by a term proportional to $P \cdot A_{00}$. This renormalization is possible because the extra term has the same functional form as the standard term — both are proportional to (area) × (radial function) × δε.

S6.5 Why absorption fails without spherical symmetry

For a non-spherically-symmetric configuration, the anisotropy has l ⩾ 1 components:

$$\delta\Phi_{cross} = \sum_{l\geq 0}\sum_{m} \mathcal{D}_{lm}(R)\, Y_{lm}(\hat{x}). \tag{S52}$$

The standard thermodynamic terms δQ and TdS, for a spherically symmetric background with perturbation δε(x), contain only l = 0 components (if the perturbation is uniform) or match the perturbation's harmonic content.

The l ⩾ 1 components of δΦ_cross have angular dependence *not present* in δQ or TdS. No scalar redefinition of T or S can accommodate angular dependence in the deficit. The equation

$$\delta Q(x) = T'(x)\, dS(x) \quad \forall x \in \Sigma \tag{S53}$$

would require T' to vary over the surface in a way that is not consistent with a single well-defined temperature.

Therefore, the first law fails for generic surfaces around non-spherically-symmetric distributions, with the failure concentrated in the l ⩾ 1 multipole components of the entanglement anisotropy. ■

S7. Multipole Predictions

S7.1 Nearly spherical configurations

For a mass distribution with small multipole moments Q_l (l ⩾ 1), the correlation field is nearly spherically symmetric. The anisotropy acquires small l ⩾ 1 components proportional to Q_l.

Dipole (l = 1). The l = 1 anisotropy arises from an off-center distribution. For a surface centered on the center of mass, the dipole contribution vanishes by the definition of center of mass (the first moment of the mass distribution about the center of mass is zero). Therefore A_{1m} = 0 for surfaces centered on the center of mass.

Quadrupole (l = 2). The leading nonvanishing correction. For a distribution with quadrupole tensor Q_{2m}, the anisotropy on a sphere of radius R scales as:

$$\mathcal{A}_{2m}(R) \propto \frac{Q_{2m}}{R^3}. \tag{S54}$$

This scaling follows from the standard multipole expansion: the contribution of a quadrupole to the potential at distance R falls off as R^{-3}.

S7.2 The deficit scaling

The surface-averaged deficit magnitude:

$$\langle |\Delta_{FL}|^2 \rangle_\Sigma^{1/2} = \left[\int_\Sigma |\Delta_{FL}(x)|^2 \, d\mu_\Sigma \right]^{1/2}. \tag{S55}$$

Using orthonormality of spherical harmonics:

$$\langle |\Delta_{FL}|^2 \rangle = \sum_{l \geq 1} \sum_m |\mathcal{D}_{lm}|^2 \approx \sum_m |\mathcal{D}_{2m}|^2 + \text{higher } l, \tag{S56}$$

where D_{lm} ∝ P(R) · A_{lm}(R) · h. Therefore:

$$\langle |\Delta_{FL}|^2 \rangle \propto \mathcal{P}(R)^2 \sum_m \frac{|Q_{2m}|^2}{R^6} = \frac{\mathcal{P}(R)^2 \cdot |Q_2|^2}{R^6}. \tag{S57}$$

Prediction: The root-mean-square thermodynamic deficit on a sphere of radius R around a nearly spherical mass distribution scales as:

$$\Delta_{FL}^{rms} \propto \frac{\mathcal{P}(R) \cdot |Q_2|}{R^3}. \tag{S58}$$

This is testable in numerical relativity: compute the first law violation on constant-potential surfaces around oblate or prolate mass distributions and verify the Q_2/R^3 scaling.

S7.3 Higher multipoles

The l-th multipole contributes:

$$\Delta_{FL}^{(l)} \propto \frac{\mathcal{P}(R) \cdot |Q_l|}{R^{l+1}}. \tag{S59}$$

Higher multipoles are suppressed at large R by higher powers of 1/R. At the surface of a compact object (R ~ R_object), all multipoles contribute comparably, and the total deficit is dominated by the overall shape asymmetry.

S8. Connection to Entanglement Entropy

S8.1 P(Σ) as entanglement entropy

The identification between P(Σ) and entanglement entropy S_EE is developed in the main text (Sec. IV). Here we provide the technical details.

For a system in a pure quantum state |Ψ⟩ with density matrix ρ = |Ψ⟩⟨Ψ|, partition into I and O with reduced density matrix ρ_I = Tr_O ρ. The entanglement entropy is:

$$S_{EE} = -\mathrm{Tr}(\rho_I \ln \rho_I). \tag{S60}$$

S8.2 Expansion in terms of correlation functions

For a state described by pairwise correlations, the reduced density matrix can be expressed as:

$$\rho_I = \rho_I^{(0)} + \sum_{i\in I, j\in O} \lambda_{ij} \cdot \rho_I^{(1)}(i,j) + O(\lambda^2), \tag{S61}$$

where ρ_I^(0) is the reduced density matrix when all cross-boundary correlations are zero (a product state projected to I) and ρ_I^(1)(i,j) is the first-order correction from the (i,j) correlation.

The von Neumann entropy expanded to first order:

$$S_{EE} = S_{EE}^{(0)} + \sum_{i\in I, j\in O} \lambda_{ij} \cdot s_{ij} + O(\lambda^2), \tag{S62}$$

where S_EE^(0) = 0 (a pure state has zero entanglement entropy when no correlations cross the boundary) and

$$s_{ij} = -\mathrm{Tr}\left[\rho_I^{(1)}(i,j) \cdot (\ln \rho_I^{(0)} + \mathbf{1})\right]. \tag{S63}$$

For weak correlations, s_ij > 0 (each correlation adds entanglement). Therefore:

$$S_{EE} = \sum_{i\in I, j\in O} \lambda_{ij} \cdot s_{ij} + O(\lambda^2) = c_1 \mathcal{P}(\Sigma) + O(\lambda^2), \tag{S64}$$

where c_1 is a characteristic susceptibility (the correlation-weighted average of s_ij).

This confirms P ∝ S_EE at leading order, with corrections that are quadratic in the cross-boundary correlations.

S8.3 The deficit in entanglement language

With P = S_EE/c_1, the deficit theorem becomes:

$$\Delta_{FL} = \frac{1}{c_1} \int_{\Sigma} S_{EE}(x) \cdot \hat{\mathcal{A}}(x) \cdot h(x)\, d\mu_{\Sigma} + O(S_{EE}^2). \tag{S65}$$

The thermodynamic deficit is proportional to the entanglement entropy weighted by the entanglement anisotropy. This is Eq. (20) of the main text.

S9. Summary of Assumptions and Results

S9.1 Assumptions

Label	Assumption	Where used
(P1)	Additivity of Φ under independence	Lemma S1
(P2)	Smoothness of Φ in λ_ij	Lemma S1, Eqs. S6-S9
(P3)	Monotonicity of Φ	Not used in proof (only for interpretation)
(S1)	Monogamy constraint	Validity conditions, S4.2
Large N	$N \gg 1$	Approx. 1, 3 (Sec. S4.1)
Smooth Σ	$\xi \ll R_\Sigma$	Approx. 2 (Sec. S4.1)
Small ε	$\varepsilon \ll R_\Sigma$	Linear response regime
Uniform φ	$\varphi^0_{ij} \approx \bar{\varphi}$	Simplification in S3.3 (qualitative results hold without; see S3.3 remark)

S9.2 Results

Result	Equation	Status
Potential decomposition	(S2)	Exact
First-order Φ_{cross}	(S9)	Approx. $O(\lambda^2)$
Thermodynamic deficit = $\delta\Phi_{cross}$	(S34)	Approx. $O(\varepsilon^2, 1/N)$

Result	Equation	Status
Deficit = 0 at correlation boundaries	(S37)	Exact
Deficit ~ ε_P at stretched boundaries	(S38)	Approx. $O(\varepsilon^2\_P)$
Deficit = 0 for isotropic A	(S39)	Approx. $O(\varepsilon^2)$
Spherical: only l=0 survives	(S44)	Exact (by symmetry)
l=0 absorbable into T'	(S51)	Exact (by functional form)
l⩾1 not absorbable	(S53)	Exact (by harmonic orthogonality)
Quadrupole scaling Q_2/R^3	(S58)	Leading order in Q_l/R
$P \propto S\_EE$	(S64)	Approx. $O(\lambda^2)$

S9.3 What is proven vs. conjectured

Proven within the stated assumptions:

- The product structure $\Delta \propto P \cdot A$ (Theorem 1)
- The vanishing at correlation boundaries (exact)
- The spherical symmetry exception via harmonic absorption (exact given spherical symmetry)
- The failure for generic surfaces (exact given $l \geqslant 1$ content)

Derived but requiring additional input for quantitative predictions:

- The proportionality constant relating P to S_EE (requires specifying the microscopic theory)
- The numerical coefficient in the multipole scaling (requires specifying $\lambda(x,y)$ in gravitational contexts)

Not proven here:

- That the monogamy constraint (S1) holds for the physical degrees of freedom relevant to gravity (assumed, motivated by entanglement monogamy)
- That the potential Φ satisfying (P1)-(P3) exists for the gravitational system (assumed, motivated by black hole thermodynamics)

References

[S1] See main text Refs. [1]-[21].

[S2] E. H. Lieb and M. B. Ruskai, "Proof of the Strong Subadditivity of Quantum-Mechanical Entropy," J. Math. Phys. **14**, 1938 (1973).

[S3] M. M. Wolf, F. Verstraete, M. B. Hastings, and J. I. Cirac, "Area Laws in Quantum Systems: Mutual Information and Correlations," Phys. Rev. Lett. **100**, 070502 (2008).

[S4] J. Eisert, M. Cramer, and M. B. Plenio, "Colloquium: Area laws for the entanglement entropy," Rev. Mod. Phys. **82**, 277 (2010).