Eur. Phys. J. Plus (2026) 141:180
https://doi.org/10.1140/epjp/s13360-026-07398-x

THE EUROPEAN PHYSICAL JOURNAL PLUS

Regular Article

Thermodynamics reconstructed from information theory: an axiomatic framework via information-volume constraints and path-space KL divergence

Tatsuaki Tsuruyama[1,2,a]

[1] Department of Physics, Tohoku University, Sendai 980-8578, Japan
[2] Department of Drug Discovery Medicine, Kyoto University, Kyoto 606-8501, Japan

Received: 29 December 2025 / Accepted: 2 February 2026

Abstract We develop an axiomatic reconstruction of thermodynamics based entirely on two primitive components: a description of what aspects of a system are observed and a reference measure that encodes the underlying descriptive convention. These ingredients define an information volume for each observational cell. By incorporating the logarithm of this volume as an additional constraint in a minimum-relative entropy inference scheme, temperature, chemical potential, and pressure arise as conjugate variables of a single information-theoretic functional. This leads to a Legendre-type structure and a first-law-like relation in which pressure corresponds to information volume rather than geometric volume. For nonequilibrium dynamics, entropy production is characterized through the relative entropy asymmetry between forward and time-reversed stochastic evolutions. A decomposition using observational entropy then separates total dissipation into system and environment contributions. Heat is defined as the part of dissipation not accounted for by the system-entropy change, yielding a representation that does not rely on local detailed balance or a specific bath model. We further show that the difference between joint and partially observed dissipation equals the average of conditional relative entropies, providing a unified interpretation of hidden dissipation and information-flow terms as projection-induced gaps.

1 Introduction

Thermodynamic quantities (temperature, chemical potential, pressure, heat, work, etc.) have traditionally been defined either (i) through equilibrium statistical mechanics, via partition functions and thermodynamic potentials, or (ii) on specific dynamical models describing energy exchange with a heat bath [e.g., Markov processes satisfying local detailed balance (LDB)]. In contrast, from the point of view of information theory, entropy and free energy are naturally characterized by the Kullback–Leibler (KL) divergence. In this view, fluctuation theorems and the second law are often expressed as time-reversal asymmetry of path measures [1–4].

The aim of this paper is to push this perspective one step further:

(1) We introduce an *information–volume* variable derived from the primitive data (M, τ)—a measurement map M and a reference measure τ—and define *pressure* as a primary dual variable, on par with temperature and chemical potential, via an expectation constraint on the log-volume.

(2) We introduce dissipation in a primary manner as the path-space KL divergence between the forward path measure $\mathbb{P}$ and its time-reversed counterpart $\mathbb{P}^{\mathrm{R}}$,

$$\Sigma_{0,t} := D_{\mathrm{KL}}\big(\mathbb{P} \,\|\, \mathbb{P}^{\mathrm{R}}\big).$$

Combining this with observational entropy $S_{M,\tau}$, we obtain a representation theorem that identifies the *environmental entropy change* ΔS_{env}.

(3) We integrate these static and dynamic constructions under a single reference measure τ, and organize their relations to existing information thermodynamics (observational entropy, mutual information, hidden dissipation, etc.) in an axiomatic way.

First, starting only from the primitive descriptive data (M, τ) and KL-based inference principles, we *derive* (i) an equilibrium state family as a Min-KL solution, (ii) a Legendre-type dual structure, (iii) the intensive variables (T, μ, P) as Lagrange multipliers conjugate to $(U, \bar{N}, \bar{\ell}_V)$, and (iv) a model-independent representation of heat from the pair (path-KL dissipation, observational entropy). Second, we identify these derived quantities with the conventional thermodynamic nomenclature in regimes where additional structure is available (e.g., geometric volume or local detailed balance), while outside such regimes, the same definitions remain internally consistent and comparable across descriptions. Consequently, because both static and dynamic notions are anchored to (M, τ) and

[a] e-mail: tsuruyam@kuhp.kyoto-u.ac.jp (corresponding author)

path measures, the framework (a) defines pressure and heat even when geometric volume or a heat bath model are unavailable, (b) makes the system–environment entropy split explicitly *gauge* dependent while keeping total dissipation *gauge* invariant, (c) unifies hidden dissipation and information-flow terms as KL projection gaps, and (d) turns observation design into a principled problem (choosing (M, τ) to tighten observable dissipation bounds). We use *gauge* only in the bookkeeping/descriptive sense associated with (M, τ), not in the sense of a physical gauge redundancy. In what follows, we present static Min-KL inference, dynamic path-KL dissipation, and the resulting information-theoretic representation of heat in as model-independent a form as possible.

2 Primitive data, information volume, and observational entropy

We begin by specifying the minimal descriptive ingredients required in our framework. Let Ω denote the set of microscopic states of the system, equipped with a σ-algebra $\mathcal{F}$. Rather than assuming direct access to the full microscopic description, we assume that the observer has access only to a coarse-grained description of states through a measurable observation map

$$M : \Omega \to \mathcal{Y},$$

where $\mathcal{Y}$ denotes the set of possible observation outcomes (macrostates). Thus, the map M specifies which aspects of the microscopic state are accessible to observation and which are ignored.

In addition, we fix a σ-finite reference measure τ on $(\Omega, \mathcal{F})$. This reference measure encodes the underlying descriptive convention for counting or weighting microscopic states; depending on the context, it may represent a counting measure, a Lebesgue measure, or another natural baseline. Throughout the paper, we refer to the pair (M, τ) as the *primitive descriptive data*, as it fully specifies both what is observed and how microscopic states are measured.

For each observation outcome $y \in \mathcal{Y}$, the observation map M induces a cell

$$C_y := M^{-1}(y) \subset \Omega,$$

consisting of all microscopic states that are observationally indistinguishable. The τ-measure of this cell

$$\upsilon(y) := \tau(C_y), \tag{1}$$

defines the *information volume* associated with the observation outcome y. This quantity measures the degeneracy of the microscopic description compatible with a given observation.

Correspondingly, we define the log-volume variable on Ω by

$$\ell_V(x) := \log \tau(C_{M(x)}) = \log \upsilon(M(x)), \tag{2}$$

which quantifies how much microscopic information is collapsed by the observation map at the state x.

Given the primitive descriptive data (M, τ), a natural notion of system entropy should quantify the uncertainty of the observed outcomes $y = M(x)$ while explicitly accounting for the information volume $\upsilon(y) = \tau(C_y)$ associated with each observational cell. This requirement leads uniquely to the notion of *observational entropy* introduced by Safránek et al. [5–9], which we adopt as the system entropy in the present framework.

2.1 Observational entropy

For $p \ll \tau$, define $p_M(y) := p(C_y)$. For finite $\mathcal{Y}$, the observational entropy is

$$S_{M,\tau}(p) := -\sum_{y\in\mathcal{Y}} p_M(y) \log\left(\frac{p_M(y)}{\tau(C_y)}\right) = -\sum_y p_M(y) \log p_M(y) + \sum_y p_M(y) \log \tau(C_y). \tag{3}$$

We adopt $S_{M,\tau}$ as the primary definition of system entropy.

2.2 Gauge freedom in the choice of (M, τ)

Changes in (M, τ) represent *changes of description*. Two basic operations are relevant.

2.2.1 (i) Refinement/coarse-graining of M.

If M_1 is a refinement of M_2 then (Safránek et al.)

$$S_{M_1,\tau}(p) \le S_{M_2,\tau}(p).$$

Hence, the system-entropy change

$$\Delta S_{\text{sys}} = S_{M,\tau}(p_t) - S_{M,\tau}(p_0)$$

depends on the chosen observation M. In contrast, the path-KL dissipation

$$\Sigma_{0,t} = D_{\mathrm{KL}}(\mathbb{P}\|\mathbb{P}^{\mathrm{R}})$$

is independent of M.

(ii) Changing the reference measure τ. Even when the observation map M is fixed, the description may be altered by changing the reference measure τ. Let τ' be another reference measure that is absolutely continuous with respect to τ, with

$$\mathrm{d}\tau' = e^{-\varphi}\ \mathrm{d}\tau.$$

By definition of the KL divergence,

$$D_{\mathrm{KL}}\big(p \parallel \tau'\big) = \int \log\!\left(\frac{\mathrm{d}p}{\mathrm{d}\tau'}\right) \mathrm{d}p = \int \log\!\left(\frac{\mathrm{d}p}{\mathrm{d}\tau}\right) \mathrm{d}p + \int \varphi\ \mathrm{d}p,$$

up to an additive constant independent of p arising from normalization. Hence

$$D_{\mathrm{KL}}\big(p \parallel \tau'\big) = D_{\mathrm{KL}}(p \parallel \tau) + \int \varphi\ \mathrm{d}p + \text{const.}$$

Thus, changing τ induces an additive gauge transformation of entropy-like functionals (up to an additive constant independent of p, when finite). Here "const" denotes an additive term independent of p, arising only from the choice of normalization of the reference measures.

Remark 1 In the decomposition

$$\Sigma_{0,t} = \Delta S_{\mathrm{sys}} + \Delta S_{\mathrm{env}},$$

both ΔS_{sys} and ΔS_{env} depend on $(M,\ \tau)$, whereas $\Sigma_{0,t}$ is invariant. Our framework separates gauge-invariant dissipation from the gauge-dependent bookkeeping of system/environment entropy.

3 Static theory: Min-KL inference and dual variables

We define thermodynamic states not by dynamical assumptions but as outcomes of an information-update rule. The basic principle is minimum cross-entropy (minimum KL) with respect to the reference measure τ, justified information-theoretically by Jaynes' MaxEnt interpretation [10] and the Shore–Johnson axiomatization [11].

As constraints, we specify expectations of energy $E(x)$, particle number $N(x)$, and the information-volume code length $\ell_V(x)$

$$\mathbb{E}_p[E] = U, \qquad \mathbb{E}_p[N] = \bar{N}, \qquad \mathbb{E}_p[\ell_V] = \bar{\ell}_V. \tag{4}$$

We then define the state p^* by

$$p^*(U, \bar{N}, \bar{\ell}_V) := \arg\min_p\ D_{\mathrm{KL}}(p \parallel \tau) \quad \text{s.t. (4)}. \tag{5}$$

Here, the KL divergence (using the Radon–Nikodym density $r = \mathrm{d}p/\mathrm{d}\tau$) is

$$D_{\mathrm{KL}}(p \parallel \tau) := \int_\Omega \log r(x)\, p(\mathrm{d}x) = \int_\Omega r(x) \log r(x)\, \tau(\mathrm{d}x). \tag{6}$$

Under convexity and regularity assumptions, the minimizer belongs to the exponential family

$$p^*(\mathrm{d}x) = \frac{1}{Z(\beta,\alpha,\pi)}\ \exp\!\Big(-\beta E(x) - \alpha N(x) - \pi\,\ell_V(x)\Big)\, \tau(\mathrm{d}x), \tag{7}$$

where Z is the normalization constant.

3.1 Definitions of temperature, chemical potential, and pressure

We *primarily* define the intensive variables as dual variables

$$\frac{1}{T} := \beta, \qquad -\frac{\mu}{T} := \alpha, \qquad \frac{P}{T} := \pi. \tag{8}$$

Hence, temperature is the multiplier for the energy constraint, chemical potential corresponds to the particle-number constraint, and pressure is the shadow price of the information–volume constraint.

3.1.1 Intuition and a non-geometric example

The *information volume* $v(y)$ is the size (the τ-measure) of the microstate set C_y that is consistent with an observation outcome y. When a geometric volume is available, v reproduces the usual volume factor; however, even for discrete, combinatorial, or non-spatial state spaces, $v(y)$ has a clear meaning as a degeneracy (the number of microstates yielding the same observation). The log-volume

$$\ell_V(x) = \log \tau(C_{M(x)})$$

quantifies how much micro-information is collapsed by the observation (the log-size of the equivalence class), and imposing the constraint $\mathbb{E}[\ell_V] = \bar{\ell}_V$ fixes, at the primary level, the typical degeneracy (or code length) of the observation cells visited by the system. In the Min-KL solution (7), $\pi = P/T$ appears as the Lagrange multiplier conjugate to ℓ_V, in exactly the same sense that $\beta = 1/T$ and $\alpha = -\mu/T$ are conjugate multipliers for the energy and particle-number constraints. Thus, P is an intensive variable that quantifies how strongly the equilibrium family suppresses or enhances observation outcomes with large degeneracy, even when no geometric container volume is available.

Example (frequency statistics and type-class degeneracy) Consider a code sequence $x = (x_1, \ldots, x_N)$ of length N, where each symbol takes values in a finite alphabet $\mathcal{A}$ with $|\mathcal{A}| = q$. Let the state space be $\Omega = \mathcal{A}^N$ and the reference measure be the counting measure τ. Take the observation M to output the *empirical distribution* (type)

$$\hat{p}_a(x) := \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}\{x_i = a\}, \qquad a \in \mathcal{A}.$$

Then, the cell $C_{\hat{p}}$ associated with $y = M(x) = \hat{p}$ consists of all sequences having the same frequency vector (the same type), i.e., the *type class*. Its size is given by the multinomial coefficient

$$v(\hat{p}) = \tau(C_{\hat{p}}) = |C_{\hat{p}}| = \frac{N!}{\prod_{a\in\mathcal{A}}(N\hat{p}_a)!}, \qquad \ell_V(x) = \log v(\hat{p}(x)).$$

Therefore, ℓ_V is the logarithm of the number of micro-sequences compatible with a given empirical distribution, quantifying the combinatorial volume (degeneracy) of the macro-description without assuming any geometric structure.

Relation to Shannon entropy: Using Stirling's approximation, one obtains

$$\log v(\hat{p}) = N\,H(\hat{p}) \;-\; \frac{q-1}{2}\log N \;+\; O(1), \qquad H(\hat{p}) := -\sum_{a\in\mathcal{A}} \hat{p}_a \log \hat{p}_a,$$

where $H(\hat{p})$ is the Shannon entropy of the empirical distribution on $\mathcal{A}$. For $q = 2$ and $\hat{p} = (p, 1-p)$, it reduces to the binary entropy $h(p) := -p\log p - (1-p)\log(1-p)$.

Thus, for large N, the leading term satisfies

$$\frac{1}{N}\,\ell_V(x) \approx H(\hat{p}(x)).$$

In other words, the more entropic (more uniform) the empirical distribution is, the exponentially larger the number of micro-sequences consistent with it, and hence the larger the degeneracy.

Meaning of the $\mathbb{E}[\ell_V]$*constraint and of* P: The constraint $\mathbb{E}[\ell_V] = \bar{\ell}_V$ controls the typical degeneracy of the macrostates that are observed, and P is the intensive variable conjugate to this non-geometric volume ℓ_V, in the same dual sense as T and μ.

The log-partition function is defined by

$$\Psi(\beta,\alpha,\pi) := \log \int_\Omega \exp\Big(-\beta E - \alpha N - \pi\ell_V\Big)\,\tau(\mathrm{d}x) = \log Z(\beta,\alpha,\pi) \tag{9}$$

which implies

$$U = -\partial_\beta \Psi, \qquad \bar{N} = -\partial_\alpha \Psi, \qquad \bar{\ell}_V = -\partial_\pi \Psi. \tag{10}$$

Furthermore, defining

$$S^*(U, \bar{N}, \bar{\ell}_V) := -D_{\mathrm{KL}}\big(p^* \,\|\, \tau\big), \tag{11}$$

convex duality yields

$$\mathrm{d}S^* = \beta\;\mathrm{d}U + \alpha\;\mathrm{d}\bar{N} + \pi\;\mathrm{d}\bar{\ell}_V. \tag{12}$$

Substituting (8), we obtain a differential relation similar to the first law

$$\mathrm{d}U = T\;\mathrm{d}S^* - P\;\mathrm{d}\bar{\ell}_V + \mu\;\mathrm{d}\bar{N}. \tag{13}$$

Proposition 1 *(Information-theoretic derivation of a first-law-type relation) Assume the Min-KL problem (5) has a unique solution p^*, and that $\Psi(\beta, \alpha, \pi)$ is C^2 and convex. Let S^* be defined by (11). Then, the total differential satisfies (12). With the intensive variables defined by (8), the internal energy differential satisfies (13).*

Proof By convex duality

$$S^*(U, \bar{N}, \bar{\ell}_V) = \inf_{\beta,\alpha,\pi} \{\beta U + \alpha \bar{N} + \pi \bar{\ell}_V - \Psi(\beta, \alpha, \pi)\}.$$

The minimizer satisfies the stationarity conditions (10), giving

$$\partial_U S^* = \beta, \qquad \partial_{\bar{N}} S^* = \alpha, \qquad \partial_{\bar{\ell}_V} S^* = \pi,$$

hence (12). Inserting (8) and solving for $\mathrm{d}U$ yield (13). □

4 Nonequilibrium extension: path-space KL and hidden dissipation

We extend the preceding information-theoretic framework to time-evolving nonequilibrium processes. Irreversibility (entropy production, dissipation) is defined as the time-reversal asymmetry of *path measures*. The most general formulation—independent of coordinate choices and independent of any model-specific assumptions such as local detailed balance—is the path-space KL divergence [1–3].

4.1 Definition of path dissipation (entropy production)

Let $\mathbb{P}$ be the forward path measure of a stochastic process on $[0, t]$, and $\mathbb{P}^{\mathrm{R}}$ the time-reversed path measure induced by the reversal map. Define dissipation by

$$\Sigma_{0,t} := D_{\mathrm{KL}}\big(\mathbb{P} \,\|\, \mathbb{P}^{\mathrm{R}}\big) \geq 0. \tag{14}$$

This definition requires no thermodynamic model, no heat bath, and no LDB assumptions.

4.2 Lower bound under observation (coarse-graining) and hidden dissipation

Observation in time induces a map on paths

$$\Pi: \ (x_{0:t}) \mapsto (y_{0:t}), \qquad y_s = M(x_s). \tag{15}$$

Define observed path measures $\Pi\mathbb{P}$ and $\Pi\mathbb{P}^{\mathrm{R}}$, and observed dissipation

$$\Sigma_{\mathrm{obs}} := D_{\mathrm{KL}}\big(\Pi\mathbb{P} \,\|\, \Pi\mathbb{P}^{\mathrm{R}}\big). \tag{16}$$

By the data-processing inequality

$$\Sigma_{0,t} = D_{\mathrm{KL}}\big(\mathbb{P} \,\|\, \mathbb{P}^{\mathrm{R}}\big) \geq D_{\mathrm{KL}}\big(\Pi\mathbb{P} \,\|\, \Pi\mathbb{P}^{\mathrm{R}}\big) = \Sigma_{\mathrm{obs}}. \tag{17}$$

4.2.1 Joint/marginal dissipations and hidden dissipation

Consider a joint system (X, Y) with joint forward and reverse path measures $\mathbb{P}_{XY}$, $\mathbb{P}^{\mathrm{R}}_{XY}$, and marginal measures $\mathbb{P}_X$, $\mathbb{P}^{\mathrm{R}}_X$. Define

$$\Sigma_{XY} := D_{\mathrm{KL}}\big(\mathbb{P}_{XY} \,\|\, \mathbb{P}^{\mathrm{R}}_{XY}\big), \tag{18}$$

$$\Sigma_X := D_{\mathrm{KL}}\big(\mathbb{P}_X \,\|\, \mathbb{P}^{\mathrm{R}}_X\big). \tag{19}$$

The difference

$$\Sigma^{(X)}_{\mathrm{hidden}} := \Sigma_{XY} - \Sigma_X \tag{20}$$

quantifies the dissipation invisible from subsystem X.

Proposition 2 *(Projection-gap representation and nonnegativity of hidden dissipation) Assume that joint and marginal path measures admit densities with respect to common references and that conditional path measures $\mathbb{P}_{Y|X}$, $\mathbb{P}^R_{Y|X}$ exist. Then*

$$\Sigma^{(X)}_{hidden} = \int D_{KL}\Big(\mathbb{P}_{Y|X=\gamma} \,\|\, \mathbb{P}^R_{Y|X=\gamma}\Big)\mathbb{P}_X(d\gamma) \geq 0. \tag{21}$$

Proof Write

$$\mathrm{d}\mathbb{P}_{XY} = \mathrm{d}\mathbb{P}_X \; \mathrm{d}\mathbb{P}_{Y|X}, \qquad \mathrm{d}\mathbb{P}^{\mathrm{R}}_{XY} = \mathrm{d}\mathbb{P}^{\mathrm{R}}_X \; \mathrm{d}\mathbb{P}^{\mathrm{R}}_{Y|X}.$$

Then apply the chain rule for KL divergence

$$\begin{aligned}\Sigma_{XY} &= \int \log \frac{\mathrm{d}\mathbb{P}_X}{\mathrm{d}\mathbb{P}^{\mathrm{R}}_X}\, \mathrm{d}\mathbb{P}_X + \int \left[\int \log \frac{\mathrm{d}\mathbb{P}_{Y|X}}{\mathrm{d}\mathbb{P}^{\mathrm{R}}_{Y|X}}\, \mathrm{d}\mathbb{P}_{Y|X} \right] \mathrm{d}\mathbb{P}_X \\ &= \Sigma_X + \int D_{\mathrm{KL}}\Big(\mathbb{P}_{Y|X=\gamma} \,\|\, \mathbb{P}^{\mathrm{R}}_{Y|X=\gamma}\Big) \mathbb{P}_X(\mathrm{d}\gamma). \end{aligned} \tag{22}$$

The KL divergence is nonnegative, completing the proof. □

This result shows that hidden dissipation is not an ad hoc correction but a KL projection gap. It provides a geometric foundation for mutual-information-flow terms in frameworks of Ito–Sagawa, Horowitz–Esposito, etc. [12–14].

4.3 Mutual information and correlation free energy

For a static joint distribution p_{XY} with marginals $p_X,\ p_Y$, the mutual information is

$$I(X;Y) = D_{\mathrm{KL}}(p_{XY} \| p_X p_Y). \tag{23}$$

Assume the Hamiltonian is additive

$$E_{XY}(x, y) = E_X(x) + E_Y(y), \tag{24}$$

and fix an isothermal reference temperature T (so $\beta := 1/T$). Let $q_X,\ q_Y$ be canonical reference distributions

$$q_X(x) = \frac{e^{-\beta E_X(x)}}{Z_X(\beta)}, \qquad q_Y(y) = \frac{e^{-\beta E_Y(y)}}{Z_Y(\beta)}. \tag{25}$$

(Any additive constants in $E_X,\ E_Y$ are absorbed into $Z_X,\ Z_Y$.)

4.4 Nonequilibrium free energies

4.4.1 4.4.1

For generic distributions $p_X,\ p_Y,\ p_{XY}$, define the (Shannon-based) nonequilibrium free energies

$$F_X[p_X] := \sum_x p_X(x)\, E_X(x) - T\, H(p_X), \tag{26}$$

$$F_Y[p_Y] := \sum_y p_Y(y)\, E_Y(y) - T\, H(p_Y), \tag{27}$$

$$F_{XY}[p_{XY}] := \sum_{x,y} p_{XY}(x, y)\, (E_X(x) + E_Y(y)) - T\, H(p_{XY}), \tag{28}$$

where $H(p) := -\sum_i p_i \log p_i$ denotes the Shannon entropy.

Proposition 3 *(Free-energy decomposition with a correlation term) Under the above assumptions*

$$F_{XY}[p_{XY}] = F_X[p_X] + F_Y[p_Y] + T\, I(X;Y). \tag{29}$$

Moreover, the relative entropy to the product canonical reference decomposes as

$$D_{\mathrm{KL}}(p_{XY} \| q_X q_Y) = I(X;Y) + D_{\mathrm{KL}}(p_X \| q_X) + D_{\mathrm{KL}}(p_Y \| q_Y). \tag{30}$$

Proof Using $H(p_{XY}) = H(p_X) + H(p_Y) - I(X;Y)$ and additivity of the mean energy under (24),

$$F_{XY}[p_{XY}] = \sum_x p_X(x) E_X(x) + \sum_y p_Y(y) E_Y(y) - T(H(p_X) + H(p_Y) - I(X;Y)),$$

which is exactly (29). For (30), expand

$$D_{\mathrm{KL}}(p_{XY} \| q_X q_Y) = \sum_{x,\, y} p_{XY}(x,\ y) \log \frac{p_{XY}(x,\ y)}{q_X(x) q_Y(y)}$$

and add/subtract $\log(p_X p_Y)$

$$\log \frac{p_{XY}}{q_X q_Y} = \log \frac{p_{XY}}{p_X p_Y} + \log \frac{p_X}{q_X} + \log \frac{p_Y}{q_Y}.$$

Averaging termwise over p_{XY} yields $I(X;Y) + D_{\mathrm{KL}}(p_X \| q_X) + D_{\mathrm{KL}}(p_Y \| q_Y)$. □

4.5 Example: a two-state Markov-chain toy model

We illustrate the general theory using the simplest irreversible stochastic dynamics: a two-state, time-homogeneous Markov chain. Despite its minimal structure, this model already displays key features of path-space dissipation, its dependence on observational resolution, and the decomposition into an observed part and an observation-induced gap.

Let the state space be {0, 1}, and let the transition matrix be

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}, \qquad 0 < a, b < 1, \tag{31}$$

so that both states communicate. The stationary distribution is

$$\pi_0 = \frac{b}{a+b}, \qquad \pi_1 = \frac{a}{a+b}.$$

We consider a single time step and set $t = 1$; hence, $\Sigma_{0,t}$ becomes $\Sigma_{0,1}$ in this example. A one-step trajectory is (x_0, x_1) drawn from $\mathbb{P}(x_0 = i, x_1 = j) = \pi_i P_{ij}$. Stationarity implies that the corresponding time-reversed one-step trajectory distribution is

$$\mathbb{P}^{\mathrm{R}}(x_0 = i, x_1 = j) = \pi_j P_{ji}.$$

Thus, the dissipation (entropy production) over one step is the path-space KL divergence

$$\Sigma_{0,1} = \sum_{i,j \in \{0,1\}} \pi_i P_{ij} \log \frac{\pi_i P_{ij}}{\pi_j P_{ji}}, \tag{32}$$

which is strictly positive unless detailed balance holds ($\pi_i P_{ij} = \pi_j P_{ji}$). This quantity provides the benchmark against which the effects of coarse-graining will be measured.

4.6 Full observation

We continue to use the same two-state Markov-chain toy model introduced in Sect. 4.5, now to illustrate the dependence of observable dissipation on the descriptive gauge (M, τ).

4.6.1 (a) Full observation M_{id}

To quantify dissipation that is lost under observational coarse-graining, we define the *coarse-graining gap*

$$\Sigma_{\mathrm{cg}} := \Sigma_{0,1} - \Sigma_{\mathrm{obs}} \geq 0. \tag{33}$$

(Here, we avoid the term "hidden dissipation" for Σ_{cg} and reserve it for the *joint–marginal* gap $\Sigma^{(X)}_{\mathrm{hidden}} = \Sigma_{XY} - \Sigma_X$ in Sect. 4.2.)

Under full observation, each microstate transition is visible, so the observational channel is the identity. Therefore $\Pi\mathbb{P} = \mathbb{P}$ and $\Pi\mathbb{P}^{\mathrm{R}} = \mathbb{P}^{\mathrm{R}}$, giving

$$\Sigma_{\mathrm{obs}} = \Sigma_{0,1}, \qquad \Sigma_{\mathrm{cg}} = 0.$$

If the reference measure τ is uniform on {0, 1}, the observational entropy $S_{M_{\mathrm{id}},\tau}$ reduces to the Shannon entropy (up to an additive constant). Because the chain is stationary, the one-step process starts and ends in the same distribution, and hence,

$$\Delta S_{\mathrm{sys}} = 0.$$

Thus, all dissipation is allocated to the environment

$$\Delta S_{\mathrm{env}} = \Sigma_{0,1}, \qquad Q = T\,\Sigma_{0,1}.$$

This case shows that in a fully resolved description, dissipation is entirely revealed at the observational level; no coarse-graining gap remains.

4.6.2 (b) Extreme coarse-graining M_{triv}

Consider the opposite observational limit, in which the measurement map collapses both states into a single outcome: $M_{\mathrm{triv}}(0) = M_{\mathrm{triv}}(1) = y_0$. Then, every path is mapped to the same constant observed trajectory (y_0, y_0), both forward and backward, so

$$\Pi\mathbb{P} = \Pi\mathbb{P}^{\mathrm{R}}, \qquad \Sigma_{\mathrm{obs}} = 0.$$

Thus, the entire dissipation becomes invisible at the observational scale

$$\Sigma_{\mathrm{cg}} = \Sigma_{0,1}.$$

With a single observational cell, $S_{M_{\mathrm{triv}},\tau}(p)$ is constant in p (and may be set to zero), so again

$$\Delta S_{\mathrm{sys}} = 0, \qquad \Delta S_{\mathrm{env}} = \Sigma_{0,1}.$$

Unlike the fully observed case, the system now appears perfectly reversible at the observational level, even though the underlying dynamics is irreversible. This explicitly illustrates that the system–environment split in entropy accounting is *gauge dependent* (via M), whereas the total dissipation $\Sigma_{0,1}$ is invariant.

This toy model highlights three essential principles of the general theory:

1. Even the simplest Markov chain exhibits nonzero dissipation whenever detailed balance is broken.
2. Observational resolution controls how much dissipation is observable: coarse-graining reduces Σ_{obs}, while the total dissipation remains invariant.
3. The decomposition of dissipation into system and environment terms depends on the observational gauge, whereas the path-space KL divergence provides the unique gauge-invariant notion of irreversibility.

These features extend directly to high-dimensional or continuous-state dynamics, where coarse-graining may hide large amounts of dissipation, making the KL projection-gap interpretation crucial for analyzing nonequilibrium systems.

5 Heat as a representation theorem

Here, we connect the intensive variable T obtained from the static Min-KL construction to the dissipation $\Sigma_{0,t}$ defined dynamically on path-space. The quantity $\Sigma_{0,t}$ is model-independent and uniquely determined by the time-reversal asymmetry of path measures, whereas the system entropy $S_{M,\tau}$ is fixed once the descriptive gauge (M, τ) is chosen. Therefore, any consistent definition of heat must be constructed from these two ingredients.

5.1 Definition of heat

Consider a process that starts from p_0 and ends at p_t. We use the standard split of total dissipation into system and environment contributions

$$\Sigma_{0,t} = \Delta S_{\mathrm{sys}} + \Delta S_{\mathrm{env}}, \qquad \Delta S_{\mathrm{sys}} := S_{M,\tau}(p_t) - S_{M,\tau}(p_0), \tag{34}$$

which equivalently *defines* the environmental entropy change as

$$\Delta S_{\mathrm{env}} := \Sigma_{0,t} - \Delta S_{\mathrm{sys}}. \tag{35}$$

Note that ΔS_{sys} (hence ΔS_{env}) is gauge dependent through (M, τ), while $\Sigma_{0,t}$ is gauge invariant.

Definition 1 *(Heat (dissipated to the environment))* In this paper, we define heat Q as the heat *dissipated to the environment* (i.e., flowing into the bath) by

$$Q := T\,\Delta S_{\mathrm{env}} = T(\Sigma_{0,t} - \Delta S_{\mathrm{sys}}). \tag{36}$$

If one prefers the convention heat absorbed by the system, set $Q_{\mathrm{sys}} := -Q$.

Hence, once the gauge (M, τ) is fixed, and given the endpoints (or, more generally, the full process), heat is uniquely determined by $(T, \Sigma_{0,t}, \Delta S_{\mathrm{sys}})$, i.e., by $(T, \Sigma_{0,t})$ together with the entropy bookkeeping induced by $S_{M,\tau}$.

Remark 2 *(Connection to cross-entropy formulations of dissipated heat)* Some works write the entropy production as $\sigma = \Delta S - \beta q$ in terms of the heat *q into the system* [15]. Then, the heat dissipated to the surroundings is $q_{\mathrm{ex}} := -q$, so that $\beta q_{\mathrm{ex}} = \sigma - \Delta S$. Our definition (36) corresponds to this q_{ex} (heat dissipated to the environment), up to the choice of units (here $k_B = 1$).

Moreover, in settings where the path-KL dissipation reduces to a distribution-level KL (e.g. in relaxation/erasure-type settings or in bounds saturated by optimal protocols), $\Sigma_{0,t} = D_{\mathrm{KL}}(p \,\|\, p^*)$ for some reference (e.g. equilibrium/target) distribution p^*. If additionally $S_{M,\tau}$ reduces to the Shannon entropy $H(p) := -\sum_i p_i \log p_i$ (up to an additive constant), then we obtain the identity

$$\frac{Q}{T} = D_{\mathrm{KL}}\big(p \,\|\, p^*\big) - (S(p^*) - S(p)) = H_{cross}(p, p^*) - S(p^*), \tag{37}$$

where $H_{cross}(p, p^*) := -\sum_i p_i \log p_i^*$ is the (ordinary) cross-entropy. In particular, if p^* has full support, Q/T is finite and can be expressed via cross-entropy up to an additive term. The deterministic-target case should be understood as a limiting (regularized) sharp-peak regime: if p^* is strictly supported on a subset, then $H_{cross}(p, p^*)$ diverges unless p is supported on the same subset.

Remark 3 *(Relation to standard stochastic-thermodynamic heat under LDB)* In conventional stochastic thermodynamics, one often assumes local detailed balance (LDB) and defines a trajectory-level heat $Q_{\mathrm{LDB}}[\gamma]$ via explicit energy exchange with a heat bath. In that setting, the path-probability ratio takes the form

$$\log \frac{\mathrm{d}\mathbb{P}[\gamma]}{\mathrm{d}\mathbb{P}^{\mathrm{R}}[\gamma^{\mathrm{R}}]} = \Delta s_{\mathrm{sys}}[\gamma] + \beta\, Q_{\mathrm{LDB}}[\gamma],$$

and averaging yields $\Sigma_{0,t} = \Delta S_{\mathrm{sys}} + \beta \langle Q_{\mathrm{LDB}} \rangle$. Hence, when (M, τ) is chosen, so that $S_{M,\tau}$ reduces to the usual Gibbs–Shannon entropy, our definition $Q := T(\Sigma_{0,t} - \Delta S_{\mathrm{sys}})$ coincides with the standard mean heat $\langle Q_{\mathrm{LDB}} \rangle$. Our framework preserves this consistency while *not* assuming LDB or any specific bath model: Eq. (36) provides the unique scalar heat compatible with (i) temperature from Min-KL inference, (ii) path-KL dissipation as irreversibility, and (iii) the chosen entropy bookkeeping $S_{M,\tau}$.

Proposition 4 *(Representation theorem for heat) Heat is uniquely determined by the following axioms:*

(i) *(**Static consistency**) For quasistatic processes connecting Min-KL equilibrium states, $Q = T\,\Delta S_{env}$ reproduces the standard thermodynamic relation.*
(ii) *(**Dynamic consistency**) For an arbitrary process, irreversibility is quantified by $\Sigma_{0,t}$.*
(iii) *(**Gauge covariance**) The system entropy depends on (M, τ) through $S_{M,\tau}$, whereas $\Sigma_{0,t}$ is invariant, and heat transforms as a scalar of the form $Q = T(\Sigma_{0,t} - \Delta S_{sys})$.*

Then, Q must be given by (36).

Proof Given the decomposition (34), the only scalar combination of $(T, \Sigma_{0,t}, \Delta S_{\mathrm{sys}})$ that satisfies (i)–(iii) is $Q = T(\Sigma_{0,t} - \Delta S_{\mathrm{sys}})$. Any alternative expression violates either static consistency in the quasistatic limit, monotonicity with respect to dissipation (dynamic consistency), or gauge covariance. □

5.2 Cross-entropy identity and nonequilibrium deviation

To relate our entropy bookkeeping $S_{M,\tau}$ to cross-entropy expressions, it is natural to introduce a cross-entropy adapted to the same description (M, τ). For distributions p, q on Ω, define the (M, τ)*-cross* entropy

$$H_{cross:M,\tau}(p\|q) := -\sum_{y\in\mathcal{Y}} p_M(y)\,\log\!\left(\frac{q_M(y)}{v(y)}\right), \qquad v(y) := \tau(C_y), \tag{38}$$

where $p_M(y) := p(C_y)$ and $q_M(y) := q(C_y)$.

Lemma 1 *(A (M, τ)-cross-entropy identity) For any p, q, we have the exact decomposition*

$$H_{cross:M,\tau}(p\|q) = S_{M,\tau}(p) + D_{KL}(p_M \,\|\, q_M), \tag{39}$$

where $D_{KL}(p_M \,\|\, q_M) = \sum_y p_M(y) \log\!\big(p_M(y)/q_M(y)\big)$.

Proof By definition

$$H_{cross:M,\tau}(p\|q) = -\sum_y p_M(y) \log q_M(y) + \sum_y p_M(y) \log v(y).$$

On the other hand

$$S_{M,\tau}(p) = -\sum_y p_M(y) \log p_M(y) + \sum_y p_M(y) \log v(y).$$

Subtracting yields $H_{cross:M,\tau}(p\|q) - S_{M,\tau}(p) = \sum_y p_M(y) \log\!\big(p_M(y)/q_M(y)\big) = D_{\mathrm{KL}}(p_M \,\|\, q_M)$. □

Applying (39) at the initial and final times with any reference distributions q_0, q_T gives the following:

$$H_{cross:M,\tau}(p_t\|q_t) - H_{cross:M,\tau}(p_0\|q_0) = \Delta S_{\mathrm{sys}} + \Big[D_{\mathrm{KL}}\big(p_{t,M} \,\|\, q_{t,M}\big) - D_{\mathrm{KL}}\big(p_{0,M} \,\|\, q_{0,M}\big)\Big]. \tag{40}$$

In particular, choosing $q_t = p_t^*$ (the Min-KL equilibrium associated with the same (M, τ) and the corresponding constraint values at time t) isolates the deviation from instantaneous equilibrium at the level of observation M: the bracketed term measures the change of the observed KL distance to the instantaneous equilibrium.

Remark 4 *(Nonequilibrium generalized grand potential and the KL gap)* For later use, it is convenient to introduce the τ-entropy

$$S_\tau(p) := -D_{\mathrm{KL}}(p \,\|\, \tau), \tag{41}$$

which is canonically associated with the Min-KL inference. We define the (description-dependent) nonequilibrium *generalized grand potential* at given intensives (T, μ, P) by

$$\Phi_\tau(p; T, \mu, P) := U[p] - \mu\,\bar{N}[p] + P\,\bar{\ell}_V[p] - T\,S_\tau(p) = U[p] - \mu\,\bar{N}[p] + P\,\bar{\ell}_V[p] + T\,D_{\mathrm{KL}}(p\,\|\,\tau), \tag{42}$$

where $U[p] = \mathbb{E}_p[E]$, $\bar{N}[p] = \mathbb{E}_p[N]$, and $\bar{\ell}_V[p] = \mathbb{E}_p[\ell_V]$. Let p^* be the equilibrium state Min-KL (7) in the same intensives $(\beta, \alpha, \pi) = (1/T, -\mu/T, P/T)$. Then, we have the exact variational identity

$$\Phi_\tau(p; T, \mu, P) = -T\,\Psi(\beta, \alpha, \pi) + T\,D_{\mathrm{KL}}\big(p\,\|\,p^*\big), \tag{43}$$

and in particular, the KL gap quantifies the *excess generalized grand potential*

$$\Phi_\tau(p; T, \mu, P) - \Phi_\tau(p^*; T, \mu, P) = T\,D_{\mathrm{KL}}\big(p\,\|\,p^*\big)\ \geq\ 0. \tag{44}$$

When τ is a counting/Lebesgue measure and the description is sufficiently fine, $S_\tau(p)$ reduces to the usual Gibbs–Shannon entropy up to an additive constant, so that Φ_τ recovers the conventional thermodynamic potentials (up to the same constant shift).

6 Relation to existing theories and features of the present framework

This section summarizes how the present framework relates to existing information thermodynamics and highlights conceptual features and open problems.

6.1 Dissipation measures

6.1.1 Path-space KL as a model-independent dissipation

Expressing dissipation as a path-space KL divergence

$$\Sigma_{0,t} = D_{\mathrm{KL}}\big(\mathbb{P}\,\|\,\mathbb{P}^{\mathrm{R}}\big),$$

is established for reversible phase-space dynamics [1] and for general stochastic processes [2, 3]. The lower bound of coarse-graining $\Sigma_{0,t} \geq \Sigma_{\mathrm{obs}}$ (Eq. (17)) and related hidden dissipation gaps appear in the discussions of dissipation under coarse-graining and partial observation [4, 16]. Our contribution is to treat $\Sigma_{0,t}$ as the *primary* gauge-invariant notion of irreversibility and to interpret all loss-of-resolution corrections as KL projection gaps (Sect. 4).

Observational entropy and descriptive gauge On the static side, observational entropy $S_{M,\tau}(p)$—including cell volume $\tau(C_y)$ in the definition of entropy—has been systematically developed by Šafránek et al. in classical and quantum settings [5–9]. We adopt $S_{M,\tau}$ as the entropy of the system and emphasize that the split $\Sigma_{0,t} = \Delta S_{\mathrm{sys}} + \Delta S_{\mathrm{env}}$ is *gauge dependent* through (M, τ), while $\Sigma_{0,t}$ is invariant (Sect. 2.2). This viewpoint separates (i) objective dissipation (time-reversal asymmetry on path-space) from (ii) description-dependent bookkeeping.

Correlation free energy and KL geometry. The idea that correlations between a system and memory (or environment) store usable free energy quantified by $k_{\mathrm{B}}T$ times mutual information $I(X; Y)$ has been emphasized in the literature [12–14]. We embed these discussions into a KL-geometric decomposition of $D_{\mathrm{KL}}(p_{XY}\|q_X q_Y)$ (Eq. (30)), thus treating local and correlation free energies within a unified information-geometric framework that is compatible with the Min-KL prior τ.

6.2 Conceptual features

Our main conceptual contributions can be summarized as follows.

(i) *Primary definition of pressure via information-volume constraints.* We introduce volume not as geometric container volume but as the cell volume $v(y) = \tau(C_y)$ determined by the descriptive gauge (M, τ) and the associated log-volume $\ell_V(x)$. Incorporating the constraint $\mathbb{E}[\ell_V] = \bar{\ell}_V$ into Min-KL inference defines pressure P as the dual variable conjugate to $\bar{\ell}_V$, on equal footing with temperature T and chemical potential μ (Eq. (8)). This repositions pressure as an intensive variable conjugate to a *description-dependent information volume*, enabling extensions to non-geometric state spaces (combinatorial, discrete, or otherwise non-spatial) without changing the formal thermodynamic structure.

(ii) *Hidden dissipation and information terms as projection gaps.* At the path level, the difference between joint dissipation Σ_{XY} and marginal dissipation Σ_X admits an explicit conditional-KL representation (Eq. (21)), showing that hidden dissipation $\Sigma^{(X)}_{\mathrm{hidden}}$ is an intrinsic KL gap induced by projection from the joint description to a partial one. Information-flow terms in existing information thermodynamics can then be understood as special evaluations or bounds on this projection gap, rather than as additional postulates.

(iii) *Heat reconstructed as a representation theorem.* With dissipation defined by path-KL, system entropy defined by observational entropy, and temperature defined as a Min-KL dual variable, heat is defined by

$$\Delta S_{\rm env} = \Sigma_{0,t} - \Delta S_{\rm sys}, \qquad Q = T\,\Delta S_{\rm env}$$

(Sect. 5). Thus, Q is not a primitive dynamical input, but a secondary representation fixed by (a) gauge-invariant dissipation and (b) gauge-fixed entropy bookkeeping, and it remains meaningful even when local detailed balance or an explicit bath model is unavailable.

6.3 Open problems and directions

Several theoretical and applied directions arise naturally.

(a) *Phase-transition structure with information-volume constraints.* In Min-KL inference with $\mathbb{E}[\ell_V] = \bar{\ell}_V$, it is important to analyze when the dual map $(\beta, \alpha, \pi) \mapsto (U, \bar{N}, \bar{\ell}_V)$ loses regularity (e.g., non-invertibility, non-differentiability) and when ensemble nonequivalence occurs, directly relating to phase-transition phenomena involving information volume. Effective models without geometric volume, or with nonlocal observations M, may display critical behavior that is invisible in the conventional ensembles.
(b) *Optimization of observational dissipation limits and observation design,* The gap between observable dissipation $\Sigma_{\rm obs}$ and true dissipation $\Sigma_{0,t}$ is sensitive to the choice of (M, τ). Designing (M, τ) to tighten lower bounds (e.g., maximize $\Sigma_{\rm obs}$ under experimental constraints) is a natural optimization problem that links dissipation inference to measurement design.
(c) *Second-law-type inequalities including information-volume work.* Combining the static first-law type relation (13) with the path-KL dissipation suggests new decompositions and inequalities that involve changes in U, $\bar{\ell}_V$, and $\bar{N}$ together with dissipation and correlation generation. For example, deriving bounds that explicitly include an information-volume work-like term and correlation free energy $TI(X;Y)$ under a general prior τ would strengthen the link between nonequilibrium thermodynamics and information geometry.
(d) *Applications and empirical comparability across descriptions.* In applications to RNA polymerase–DNA systems and to chemical reaction networks analyzed via large-deviation theory [17], explicit computations of correlation free energy, hidden dissipation, and the heat representation (36) will clarify agreement with, or deviation from, conventional thermodynamic descriptions, such as fluctuation theorem and Landauer-type bounds, while keeping the descriptive gauge (M, τ) explicit.

Closing remarks Taken together, these results redefine pressure, dissipation, and heat as quantities that can be defined and compared across descriptions through primitive observational data (M, τ) and relative entropy of the path-space. By making the information-volume constraint and the path-KL definition of dissipation primary, we obtain a unified and representation-theoretic framework in which conventional thermodynamic notions emerge as secondary objects, while projection-induced gaps quantify the thermodynamic cost of partial observation. We expect this viewpoint to be especially useful when geometric volume, energy-bath idealizations, or microscopic reversibility assumptions are not directly available, and when experimental access is limited, so that observation design becomes essential.

Data availability No datasets were generated or analysed during the current study.

References

1. R. Kawai, J.M.R. Parrondo, C. Van den Broeck, Phys. Rev. Lett. **98**, 080602 (2007)
2. É. Roldán, J.M.R. Parrondo, Phys. Rev. Lett. **105**, 150607 (2010)
3. Q. Zhang, Y. Lu, Chaos **35**, 103104 (2025)
4. M. Esposito, Phys. Rev. E **85**, 041125 (2012)
5. D. Šafránek, J.M. Deutsch, A. Aguirre, Phys. Rev. A **99**, 010101 (2019)
6. D. Šafránek, J.M. Deutsch, A. Aguirre, Phys. Rev. A **99**, 012103 (2019)
7. D. Šafránek, A. Aguirre, J. Schindler, J.M. Deutsch, Found. Phys. **51**, 101 (2021)
8. D. Šafránek, A. Aguirre, J.M. Deutsch, Phys. Rev. E **102**, 032106 (2020)
9. G. Bai, D. Šafránek, J. Schindler, F. Buscemi, V. Scarani, Quantum **8**, 1524 (2024)
10. E.T. Jaynes, Phys. Rev. **106**, 620–630 (1957)
11. J.E. Shore, R.W. Johnson, IEEE Trans. Inf. Theory **26**(1), 26–37 (1980)
12. S. Ito, T. Sagawa, Phys. Rev. Lett. **111**, 180603 (2013)
13. J.M. Horowitz, M. Esposito, Phys. Rev. X **4**, 031015 (2014)
14. J.M.R. Parrondo, J.M. Horowitz, T. Sagawa, Nat. Phys. **11**, 131–139 (2015)
15. T. Tsuruyama, Eur. Phys. J. Plus **140**, 620 (2025)
16. A. Gomez-Marin, J.M.R. Parrondo, C. Van den Broeck, Phys. Rev. E **78**, 011107 (2008)
17. T. Tsuruyama, Eur. Phys. J. Plus **138**, 604 (2023)