# Unified framework for the entropy production and the stochastic interaction based on information geometry 

Sosuke Ito, ${ }^{1,2}$ Masafumi Oizumi, ${ }^{3,4}$ and Shun-ichi Amari ${ }^{4}$<br>${ }^{1}$ Universal Biology Institute, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0031, Japan<br>${ }^{2}$ JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan<br>${ }^{3}$ Department of General Systems Studies, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan<br>${ }^{4}$ RIKEN CBS Hirosawa 2-1, Wako-shi, Saitama 351-0198, Japan

(Received 5 December 2018; accepted 16 June 2020; published 9 July 2020)


#### Abstract

We show a relation between the entropy production in stochastic thermodynamics and the stochastic interaction in the integrated information theory. To clarify this relation, we introduce an information-geometric interpretation of the entropy production for a total system and the partial entropy productions for subsystems. We show that the violation of the additivity of the entropy productions is related to the stochastic interaction. This framework is a thermodynamic foundation of the integrated information theory. We also show that our information-geometric formalism leads to an expression of the entropy production related to an optimization problem minimizing the Kullback-Leibler divergence. We analytically illustrate this interpretation by using the spin model.


DOI: 10.1103/PhysRevResearch.2.033048

## I. INTRODUCTION

Information geometry [1,2] is a differential geometric theory for elucidating various results in information theory and probability theory. Applications of information geometry have been found in a variety of fields, including machine learning [3], neuroscience [4], statistical physics [5,6], and thermodynamics [7-22]. The projection theorem [23,24] plays a crucial role in applications of information geometry. For example, the projection theorem unifies the conventional definitions of information measures such as the mutual information, the transfer entropy, and several measures in the integrated information theory [2,25,26].

The integrated information theory seeks for measures of inseparability of networks [25-35]. Several measures have been proposed by considering different ways of dividing networks [25,26,34]. A possible promising measure of information integration is the stochastic interaction [29,31], which quantifies the inseparability of stochastic dynamics in two interacting systems.

In the field of stochastic thermodynamics [36,37], a similar problem of inseparability takes place. For example, in the context of Maxwell's demon, information thermodynamic measures of the correlation between two interacting dynamics have been discussed [38-50]. For two interacting dynamics, we introduce a measure of information thermodynamics, namely, the partial entropy production for the subsystem

[^0][42-44]. If two interacting dynamics are well separated, the sum of the partial entropy productions for each subsystem is equivalent to the total entropy production. This fact is known as the additivity of the entropy productions. If two interacting dynamics are not well separated, this additivity is generally violated.

In this paper, we introduce a framework of stochastic thermodynamics based on information geometry. We introduce several submanifolds related to backward dynamics, and the total entropy production and the partial entropy production can be considered to be given by the projections of the entire system onto these submanifolds. From the inclusion property of these submanifolds, we obtain a geometric interpretation of the additivity of the entropy productions. This interpretation clarifies a relation between the violation of the additivity and the stochastic interaction. Additionally, our framework leads to an expression of the entropy production by considering an optimization problem to minimize the Kullback-Leibler divergence. We analytically illustrate our results by using the spin models.

## II. THE PROJECTION THEOREM

We first introduce the projection theorem in information geometry, which is a differential geometrical theory for the manifold of the probability distribution [23,24]. In information geometry, a Riemannian metric is given by the Fisher information matrix, and a dual pair of affine connections is defined [1]. Let $p_{S}(s)$ be the joint probability, where $\boldsymbol{S}=\left\{S_{1}, \ldots, S_{N}\right\}$ is the set of random variables and $\boldsymbol{s}=$ $\left\{s_{1}, \ldots, s_{N}\right\}$ is the set of events. In information geometry, the set of the joint probabilities is considered a manifold. A subset of probabilities gives a submanifold $\mathcal{M}$, and a probability $p_{S}(\boldsymbol{s})$ corresponds to a point.


FIG. 1. Schematic of the projection theorem. The subset of probabilities gives a submanifold $\mathcal{M}$, and the probability $p$ corresponds to a point. If $\mathcal{M}$ is flat, we have a unique solution $q_{S}^{*}$ of the optimization problem to minimize the Kullback-Leibler divergence between the probability $p$ and the probability $q_{S} \in \mathcal{M}$. The flatness of the manifold is given by the Pythagorean theorem, and the solution $q_{S}^{*}$ is the projection onto the flat submanifold $\mathcal{M}$.

We now consider an optimization problem to minimize the Kullback-Leibler divergence between two probabilities $p_{S}(\boldsymbol{s})$ and $q_{S}(\boldsymbol{s})$,

$$
\begin{align*}
D^{\mathrm{opt}}\left(p_{S} \| \mathcal{M}\right) & :=\min _{q_{S} \in \mathcal{M}} D\left(p_{S} \| q_{S}\right),  \tag{1}\\
D\left(p_{S} \| q_{S}\right) & :=\sum_{s} p_{S}(s) \ln \frac{p_{S}(\boldsymbol{s})}{q_{S}(\boldsymbol{s})} \tag{2}
\end{align*}
$$

when $q_{S}(\boldsymbol{s})$ is in a submanifold $\mathcal{M}$. If the submanifold $\mathcal{M}$ is flat, we have the unique solution $q_{S}^{*} \in \mathcal{M}$ that satisfies $D^{\text {opt }}\left(p_{S} \| \mathcal{M}\right)=D\left(p_{S} \| q_{S}^{*}\right)$. This unique solution $q_{S}^{*}$ can be interpreted as the projection from point $p_{S}$ onto the flat submanifold $\mathcal{M}$. In Fig. 1, we show an intuitive schematic of the projection theorem.

This projection can be understood by considering the Pythagorean theorem

$$
\begin{equation*}
D\left(p_{S} \| q_{S}\right)=D\left(p_{S} \| q_{S}^{*}\right)+D\left(q_{S}^{*} \| q_{S}\right) \tag{3}
\end{equation*}
$$

for any probability $q_{S}$ on the flat submanifold $\mathcal{M}$ [1]. This Pythagorean theorem can be regarded as the definition of the flatness of a submanifold $\mathcal{M}$. In information geometry, the Pythagorean theorem holds when the geodesic connecting $p_{S}$ and $q_{S}^{*}$ is orthogonal to the dual geodesic connecting $q_{S}^{*}$ and $q_{S}$. From the nonnegativity of the Kullback-Leibler divergence $D\left(q_{S}^{*}| | q_{S}\right) \geqslant 0$, we obtain the fact that $q_{S}^{*}$ is the unique solution of an optimization problem,

$$
\begin{equation*}
D\left(p_{S} \| q_{S}\right) \geqslant D\left(p_{S} \| q_{S}^{*}\right)=D^{\mathrm{opt}}\left(p_{S} \| \mathcal{M}\right) \tag{4}
\end{equation*}
$$

## III. THE ENTROPY PRODUCTION AND THE PROJECTION THEOREM

## A. The total entropy production

We here consider a Markov process. Let $\boldsymbol{Z}$ and $\boldsymbol{Z}^{\prime}$ be random variables of the state of a system $\mathcal{Z}$ at times $t$ and $t+d t$, respectively. Let $p_{Z, Z^{\prime}}\left(z, z^{\prime}\right)$ be the joint probability of the states $s=\left\{z, z^{\prime}\right\}$ corresponding to random variables $\boldsymbol{S}=\left\{\boldsymbol{Z}, \boldsymbol{Z}^{\prime}\right\}$. The transition probability is given by
$T\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right):=p_{Z^{\prime} \mid \mathbf{Z}}\left(z^{\prime} \mid \boldsymbol{z}\right)$, where the conditional probability is defined as $p_{Z^{\prime} \mid Z}\left(z^{\prime} \mid z\right):=p_{Z^{\prime}, Z}\left(z^{\prime}, z\right) / p_{Z}(z)=p_{S}(\boldsymbol{s}) /\left[\sum_{z^{\prime}} p_{S}(\boldsymbol{s})\right]$. Because the transition probability $T\left(z^{\prime}, z\right)$ is a function of $\left(z^{\prime}, z\right)$, we can define a new quantity $T\left(z, z^{\prime}\right)$ by replacing $z$ with $z^{\prime}$. We remark that $T\left(z^{\prime}, z\right)$ is not equal to the conditional probability $p_{Z \mid Z^{\prime}}\left(z \mid z^{\prime}\right):=p_{S}(\boldsymbol{s}) /\left[\sum_{z} p_{S}(\boldsymbol{s})\right]$.

In stochastic thermodynamics [37], the total entropy production $\sigma_{\text {tot }}^{\mathcal{Z}}$ is defined as the sum of the entropy changes,

$$
\begin{equation*}
\sigma_{\text {tot }}^{\mathcal{Z}}:=\sigma_{\text {sys }}^{\mathcal{Z}}+\sigma_{\text {bath }}^{\mathcal{Z}} \tag{5}
\end{equation*}
$$

The entropy change of the system $\sigma_{\text {sys }}^{\mathcal{Z}}$ is defined as the Shannon entropy change from time $t$ to $t+d t$,

$$
\begin{equation*}
\sigma_{\text {sys }}^{\mathcal{Z}}:=H\left(\boldsymbol{Z}^{\prime}\right)-H(\boldsymbol{Z}) \tag{6}
\end{equation*}
$$

where $H(\boldsymbol{Z})=-\sum_{z} p_{Z}(z) \ln p_{Z}(z)$ is the Shannon entropy. The entropy change of the heat bath $\sigma_{\text {bath }}^{\mathcal{Z}}$ is defined as

$$
\begin{equation*}
\sigma_{\text {bath }}^{\mathcal{Z}}:=\mathbb{E}\left[\ln \frac{T\left(z^{\prime}, z\right)}{T\left(z, z^{\prime}\right)}\right]=\mathbb{E}\left[-\ln T\left(z, z^{\prime}\right)\right]-H\left(\mathbf{Z}^{\prime} \mid \mathbf{Z}\right) \tag{7}
\end{equation*}
$$

where the symbol $\mathbb{E}[\cdots]:=\sum_{s} p_{S}(\boldsymbol{s}) \cdots$ denotes the expected value and $H\left(\boldsymbol{Z}^{\prime} \mid \boldsymbol{Z}\right):=H\left(\boldsymbol{Z}^{\prime}, \boldsymbol{Z}\right)-H(\boldsymbol{Z})$ is the conditional Shannon entropy. The entropy change of the heat bath can be regarded as the difference between the conditional cross entropy $\mathbb{E}\left[-\ln T\left(z, z^{\prime}\right)\right]$ and the conditional Shannon entropy. The nonnegativity of the entropy production is known as the second law of thermodynamics. If the entropy production is zero, the system is reversible, and the detailed balance $p_{Z}(\boldsymbol{z}) T\left(z^{\prime}, \boldsymbol{z}\right)=p_{Z^{\prime}}\left(z^{\prime}\right) T\left(\boldsymbol{z}, z^{\prime}\right)$ holds (see Appendix A). Hence, $\sigma_{\text {tot }}^{\mathcal{Z}}$ quantifies the irreversibility of the dynamics.

We show that the total entropy production can be obtained by the projection of $p_{S}$ onto a submanifold, called the backward manifold. The backward manifold $\mathcal{M}_{\mathrm{B}}$ is defined as the set of probabilities $q_{S}$ satisfying

$$
\begin{equation*}
\mathcal{M}_{\mathrm{B}}=\left\{q_{S} \mid q_{S}(\boldsymbol{s})=q_{Z^{\prime}}\left(z^{\prime}\right) T\left(z, z^{\prime}\right)\right\}, \tag{8}
\end{equation*}
$$

where $q_{Z^{\prime}}\left(z^{\prime}\right)=\sum_{z} q_{S}(\boldsymbol{s})$ and $T\left(z, z^{\prime}\right)$ is defined from $p_{S}(\boldsymbol{s})$. The backward manifold consists of probabilities such that backward dynamics from $\boldsymbol{Z}^{\prime}$ to $\boldsymbol{Z}$ is equal to the transition probability of $p_{s}$. The backward manifold is uniquely determined by $p_{S}$. The total entropy production of the Markov process is given by

$$
\begin{equation*}
\sigma_{\mathrm{tot}}^{\mathcal{Z}}=D^{\mathrm{opt}}\left(p_{\boldsymbol{S}} \| \mathcal{M}_{\mathrm{B}}\right), \tag{9}
\end{equation*}
$$

which is the first main result of this paper. This result means that the total entropy production can be regarded as the minimum length of $p_{S}$ to the backward manifold (see also Fig. 2). To prove Eq. (9), we introduce the joint probability $q_{S}^{*}(\boldsymbol{s}):=$ $p_{Z^{\prime}}\left(z^{\prime}\right) T\left(z, z^{\prime}\right) \in \mathcal{M}_{\mathrm{B}}$; the entropy production is given by the Kullback-Leibler divergence $\sigma_{\text {tot }}^{\mathcal{Z}}=D\left(p_{S} \| q_{S}^{*}\right)$ [51]. Because the Pythagorean theorem

$$
\begin{equation*}
D\left(p_{S} \| q_{S}\right)=D\left(p_{S} \| q_{S}^{*}\right)+D\left(q_{S}^{*} \| q_{S}\right) \tag{10}
\end{equation*}
$$

is valid for any $q_{S} \in \mathcal{M}_{\mathrm{B}}$ (see Appendix A ), we obtain the first main result, Eq. (9).

## B. The partial entropy production

We next consider the situation where the system $\mathcal{Z}$ consists of two subsystems $\mathcal{X}$ and $\mathcal{Y}$ and random variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^{\prime}$


FIG. 2. Schematic of the total entropy production and the projection onto the backward manifold $\mathcal{M}_{\mathrm{B}}$. The entropy production $\sigma_{\text {tot }}^{\mathcal{Z}}$ is given by the minimum length from the backward manifold $D^{\text {opt }}\left(p_{S} \| \mathcal{M}_{\mathrm{B}}\right)$.
are given by $\boldsymbol{Z}=\{\boldsymbol{X}, \boldsymbol{Y}\}$ and $\boldsymbol{Z}^{\prime}=\left\{\boldsymbol{X}^{\prime}, \boldsymbol{Y}^{\prime}\right\}$, respectively. The transition probability of subsystem $\mathcal{X}$ for fixed states $\left\{\boldsymbol{y}, \boldsymbol{y}^{\prime}\right\}$ is given by $T^{\mathcal{X}}\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right):=p_{X^{\prime} \mid \boldsymbol{Y}^{\prime}, \boldsymbol{Z}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{y}^{\prime}, \boldsymbol{z}\right)$.

The partial entropy production for subsystem $\mathcal{X}$ is defined as

$$
\begin{gather*}
\sigma_{\text {partial }}^{\mathcal{X}}:=\sigma_{\text {sys }}^{\mathcal{X}}+\sigma_{\text {bath }}^{\mathcal{X}}-\Theta^{\mathcal{X} \rightarrow \mathcal{Y}},  \tag{11}\\
\sigma_{\text {sys }}^{\mathcal{X}}=H\left(\boldsymbol{X}^{\prime}\right)-H(\boldsymbol{X}),  \tag{12}\\
\sigma_{\text {bath }}^{\mathcal{X}}=\mathbb{E}\left[\ln \frac{T^{\mathcal{X}}\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right)}{T^{\mathcal{X}}\left(\boldsymbol{z}, \boldsymbol{z}^{\prime}\right)}\right],  \tag{13}\\
\Theta^{\mathcal{X} \rightarrow \mathcal{Y}}=I\left(\boldsymbol{X}^{\prime} ;\left\{\boldsymbol{Y}, \boldsymbol{Y}^{\prime}\right\}\right)-I\left(\boldsymbol{X} ;\left\{\boldsymbol{Y}, \boldsymbol{Y}^{\prime}\right\}\right), \tag{14}
\end{gather*}
$$

where $I\left(\boldsymbol{Z} ; \boldsymbol{Z}^{\prime}\right)=H(\boldsymbol{Z})-H\left(\boldsymbol{Z} \mid \boldsymbol{Z}^{\prime}\right)$ is the mutual information between two random variables $\boldsymbol{Z}$ and $\boldsymbol{Z}^{\prime}$. The additional term $\Theta^{\mathcal{X} \rightarrow \mathcal{Y}}$ quantifies dynamic information flow from subsystem $\mathcal{X}$ to subsystem $\mathcal{Y}$. Thus, the nonnegativity of the partial entropy production can be regarded as the second law of information thermodynamics for the subsystem $\sigma_{\text {sys }}^{\mathcal{X}}+$ $\sigma_{\text {bath }}^{\mathcal{X}} \geqslant \Theta^{\mathcal{X} \rightarrow \mathcal{Y}}$, which implies a trade-off relation between the entropy changes $\sigma_{\text {sys }}^{\mathcal{X}}+\sigma_{\text {bath }}^{\mathcal{X}}$ and information flow $\Theta^{\mathcal{X} \rightarrow \mathcal{Y}}$. The partial entropy production for subsystem $\mathcal{X}$ quantifies the local irreversibility of the dynamics in system $\mathcal{X}$. The partial entropy production vanishes if the dynamics in system $\mathcal{X}$ are locally reversible, that is, $p_{Z, Y^{\prime}}\left(\boldsymbol{z}, \boldsymbol{y}^{\prime}\right) T^{\mathcal{X}}\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right)=$ $p_{Z^{\prime}, Y}\left(z^{\prime}, \boldsymbol{y}\right) T^{\mathcal{X}}\left(z, z^{\prime}\right)$.

We here show that the partial entropy production can also be derived from the projection of $p_{S}$ onto the local backward manifold. The local backward manifold of system $\mathcal{X}$ is defined as the set of probabilities such that

$$
\begin{equation*}
\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}=\left\{q_{S} \mid q_{S}(s)=q_{Y, Z^{\prime}}\left(\boldsymbol{y}, z^{\prime}\right) T^{\mathcal{X}}\left(z, z^{\prime}\right)\right\}, \tag{15}
\end{equation*}
$$

where $q_{Y, Z^{\prime}}\left(\boldsymbol{y}, z^{\prime}\right)=\sum_{x} q_{S}(\boldsymbol{s})$ and $T^{\mathcal{X}}\left(z, z^{\prime}\right)$ is defined from $p_{S}(s)$. The local backward manifold means the set of probabilities such that local backward dynamics from $X^{\prime}$ to $X$ is equal to the transition probability in $\mathcal{X}$ of $p_{S}$. The partial entropy production of subsystem $\mathcal{X}$ is given by

$$
\begin{equation*}
\sigma_{\text {partial }}^{\mathcal{X}}=D^{\mathrm{opt}}\left(p_{\boldsymbol{S}} \| \mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}\right) \tag{16}
\end{equation*}
$$

which is the second main result of this paper. To prove Eq. (16), we introduce the probability $q_{S}^{\mathcal{X} *}(\boldsymbol{s})=$


Hierarchy of the entropy productions: $D^{\mathrm{opt}}\left(p_{\boldsymbol{S}} \| \mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}\right) \leq D^{\mathrm{opt}}\left(p_{\boldsymbol{S}} \| \mathcal{M}_{\mathrm{B}}\right)$
FIG. 3. Schematic of the partial entropy production and the hierarchy of the entropy productions. Because the local backward manifold includes the backward manifold, the partial entropy production is always smaller than the total entropy production.
$T^{\mathcal{X}}\left(\boldsymbol{z}, \boldsymbol{z}^{\prime}\right) p_{Y, Z^{\prime}}\left(\boldsymbol{y}, \boldsymbol{z}^{\prime}\right) \in \mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}$. Because we can show the expression

$$
\begin{equation*}
\sigma_{\text {partial }}^{\mathcal{X}}=D\left(p_{S} \| q_{S}^{\mathcal{X} *}\right) \tag{17}
\end{equation*}
$$

and the Pythagorean theorem

$$
\begin{equation*}
D\left(p_{S} \| q_{S}^{\mathcal{X}}\right)=D\left(p_{S} \| q_{S}^{\mathcal{X} *}\right)+D\left(q_{S}^{\mathcal{X} *} \| q_{S}^{\mathcal{X}}\right) \tag{18}
\end{equation*}
$$

for any $q_{S}^{\mathcal{X}} \in \mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}$, we obtain the second main result, Eq. (16). If we introduce the quantities for subsystem $\mathcal{Y}$ such as $\left(T^{\mathcal{Y}}, \sigma_{\text {partial }}^{\mathcal{Y}}, \sigma_{\text {sys }}^{\mathcal{Y}}, \sigma_{\text {bath }}^{\mathcal{Y}}, \Theta^{\mathcal{Y} \rightarrow \mathcal{X}}, \mathcal{M}_{\mathrm{LB}}^{\mathcal{Y}}\right)$ by replacing ( $\boldsymbol{X}, \boldsymbol{X}^{\prime}$ ) with $\left(\boldsymbol{Y}, \boldsymbol{Y}^{\prime}\right)$, we obtain the same results, Eqs. (11)(19) for subsystem $\mathcal{Y}$.

We note that our geometric interpretation provides the hierarchy of the entropy productions. Because the backward manifold is a submanifold of the local backward manifold $\mathcal{M}_{\mathrm{B}} \subset \mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}$, we obtain the hierarchy $D^{\mathrm{opt}}\left(p_{S} \| \mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}\right) \leqslant$ $D^{\text {opt }}\left(p_{S} \| \mathcal{M}_{\mathrm{B}}\right)$, or, equivalently,

$$
\begin{equation*}
\sigma_{\text {partial }}^{\mathcal{X}} \leqslant \sigma_{\text {tot }}^{\mathcal{Z}} \tag{19}
\end{equation*}
$$

This hierarchy of the entropy productions implies that the second law of information thermodynamics always gives a tighter bound than the second law of thermodynamics (see also Fig. 3). Moreover, if subsystem $\mathcal{X}_{1}$ includes subsystem $\mathcal{X}_{2}$, we obtain the hierarchy of the entropy productions

$$
\begin{equation*}
\sigma_{\text {partial }}^{\mathcal{X}_{2}} \leqslant \sigma_{\text {partial }}^{\mathcal{X}_{1}} \tag{20}
\end{equation*}
$$

from the inclusion property $\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}_{1}} \subset \mathcal{M}_{\mathrm{LB}}^{\mathcal{X}_{2}}$. This hierarchy clarifies the relation between the second laws of information thermodynamics in complex systems.

## IV. THE INTEGRATED INFORMATION THEORY AND THE ADDITIVITY

## A. The stochastic interaction in the integrated information theory

We here introduce the stochastic interaction $[29,31]$ as a measure of bidirectional information flow. The stochastic
interaction $[29,31]$ is defined as

$$
\begin{equation*}
\Phi_{\mathrm{SI}}:=D\left(p_{\boldsymbol{Z}, \boldsymbol{Z}^{\prime}} \| p_{X^{\prime} \mid \boldsymbol{Z}} p_{\boldsymbol{Y}^{\prime} \mid \boldsymbol{Z}} p_{\mathbf{Z}}\right) \tag{21}
\end{equation*}
$$

This quantity is zero if the stochastic process satisfies the bipartite condition $\mathcal{C}_{\mathrm{BI}}: p_{Z^{\prime} \mid Z}\left(\boldsymbol{z}^{\prime} \mid \boldsymbol{z}\right)=p_{X^{\prime} \mid \boldsymbol{Z}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}\right) p_{\boldsymbol{Y}^{\prime} \mid \boldsymbol{Z}}\left(\boldsymbol{y}^{\prime} \mid \boldsymbol{z}\right)$. The bipartite condition $\mathcal{C}_{\text {BI }}$ means that two transitions in $\mathcal{X}$ and $\mathcal{Y}$ are statistically independent because the transition probability $T^{\mathcal{X}}\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right)=p_{X^{\prime} \mid \boldsymbol{Z}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}\right)$ does not depend on $\boldsymbol{y}^{\prime}$ under the bipartite condition. We also define the stochastic interaction for backward dynamics as

$$
\begin{equation*}
\Phi_{\mathrm{SI}}^{\dagger}:=D\left(p_{Z, Z^{\prime}} \| p_{\boldsymbol{X} \mid Z^{\prime}} p_{\boldsymbol{Y} \mid Z^{\prime}} p_{\boldsymbol{Z}^{\prime}}\right) \tag{22}
\end{equation*}
$$

which exactly vanishes under the backward bipartite condition $\mathcal{C}_{\mathrm{BI}}^{*}: p_{Z \mid Z^{\prime}}\left(z \mid z^{\prime}\right)=p_{X \mid Z^{\prime}}\left(\boldsymbol{x} \mid z^{\prime}\right) p_{\boldsymbol{Y} \mid Z^{\prime}}\left(\boldsymbol{y} \mid z^{\prime}\right)$.

While the stochastic interactions are measures of bidirectional information flow, the dynamic information flow $\Theta^{\mathcal{X}} \rightarrow \mathcal{Y}$ is a measure of directed information flow. $\Theta^{\mathcal{X}} \rightarrow \mathcal{Y}$ can be decomposed into the mutual information difference $\Delta \mathcal{I}$ and the measures of directed information flow, i.e., the transfer entropy $I\left(\boldsymbol{X} ; \boldsymbol{Y}^{\prime} \mid \boldsymbol{Y}\right)[52,53]$ and the backward transfer entropy $I\left(\boldsymbol{X}^{\prime} ; \boldsymbol{Y} \mid \boldsymbol{Y}^{\prime}\right)$ [47],

$$
\begin{gather*}
\Theta^{\mathcal{X} \rightarrow \mathcal{Y}}=\Delta \mathcal{I}+I\left(\boldsymbol{X}^{\prime} ; \boldsymbol{Y} \mid \boldsymbol{Y}^{\prime}\right)-I\left(\boldsymbol{X} ; \boldsymbol{Y}^{\prime} \mid \boldsymbol{Y}\right),  \tag{23}\\
\Delta \mathcal{I}:=I\left(\boldsymbol{X}^{\prime} ; \boldsymbol{Y}^{\prime}\right)-I(\boldsymbol{X} ; \boldsymbol{Y}), \tag{24}
\end{gather*}
$$

where $I\left(\boldsymbol{Z} ; \boldsymbol{Z}^{\prime} \mid \boldsymbol{Z}^{\prime \prime}\right):=H\left(\boldsymbol{Z} \mid \mathbf{Z}^{\prime \prime}\right)-H\left(\boldsymbol{Z} \mid \boldsymbol{Z}^{\prime}, \boldsymbol{Z}^{\prime \prime}\right)$ is the conditional mutual information between $\boldsymbol{Z}$ and $\boldsymbol{Z}^{\prime}$ under the condition $Z^{\prime \prime}$. To compare the dynamic information flow with the stochastic interaction, we consider the bidirectional information flow by considering the sum of $\Theta^{\mathcal{X} \rightarrow \mathcal{Y}}$ and $\Theta^{\mathcal{Y} \rightarrow \mathcal{X}}$. The relation between the stochastic interaction and the dynamic information flow is given by

$$
\begin{equation*}
\Theta^{\mathcal{X} \rightarrow \mathcal{Y}}+\Theta^{\mathcal{Y} \rightarrow \mathcal{X}}-\Delta \mathcal{I}=\Phi_{\mathrm{SI}}-\Phi_{\mathrm{SI}}^{\dagger} . \tag{25}
\end{equation*}
$$

## B. The additivity and the stochastic interaction

We next discuss the additivity of the partial entropy productions. We show that the violation of the additivity is related to a measure of integrated information, i.e., stochastic interaction. Under the bipartite condition $\mathcal{C}_{\mathrm{BI}}$, we have the additivity of the entropy productions up to the order $O\left(d t^{2}\right)$ [43],

$$
\begin{equation*}
\sigma_{\text {tot }}^{\mathcal{Z}}=\sigma_{\text {partial }}^{\mathcal{X}}+\sigma_{\text {partial }}^{\mathcal{Y}} . \tag{26}
\end{equation*}
$$

From Eq. (26), the hierarchy equation (19) is equivalent to the second law of information thermodynamics for subsystem $\mathcal{Y}$, that is, $\sigma_{\text {partial }}^{\mathcal{Y}} \geqslant 0$. If time evolutions of two systems are strongly correlated, the assumption of the bipartite condition is not valid, and the additivity Eq. (26) is violated. The amount of the violation is given by the stochastic interactions and the additional term

$$
\begin{gather*}
\sigma_{\text {tot }}^{\mathcal{Z}}-\sigma_{\text {partial }}^{\mathcal{X}}-\sigma_{\text {partial }}^{\mathcal{Y}}=\Phi_{\text {bath }}+\Phi_{\mathrm{SI}}-\Phi_{\mathrm{SI}}^{\dagger},  \tag{27}\\
\Phi_{\text {bath }}^{\dagger}:=\sigma_{\text {bath }}^{\mathcal{Z}}-\sigma_{\text {bath }}^{\mathcal{X}}-\sigma_{\text {bath }}^{\mathcal{Y}} . \tag{28}
\end{gather*}
$$

The additional term $\Phi_{\text {bath }}$ quantifies to what extent the additivity is violated in the heat baths. This measure $\Phi_{\text {bath }}$ can be considered a measure of information integration for thermal systems because the entropy change does not attract much attention in integrated information theory.


FIG. 4. Schematic of the additivity and the rectangle. Under both bipartite conditions $\mathcal{C}_{\mathrm{BI}}$ and $\mathcal{C}_{\mathrm{BI}}^{*}$, the backward manifold is equal to the intersection of the local backward manifolds. The additivity of the entropy production indicates that the parallel sides of a quadrangle $\left(p_{S}, q_{S}^{\mathcal{X} *}, q_{S}^{*}, q_{S}^{\mathcal{Y} *}\right)$ have the same length.

We show a geometrical condition of this additivity under the bipartite conditions $\mathcal{C}_{\mathrm{BI}}$ and $\mathcal{C}_{\mathrm{BI}}^{*}$. Both bipartite conditions imply the relation between three manifolds,

$$
\begin{equation*}
\mathcal{M}_{\mathrm{B}}=\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}} \cap \mathcal{M}_{\mathrm{LB}}^{\mathcal{Y}} . \tag{29}
\end{equation*}
$$

Because Eq. (26) can be written as

$$
\begin{equation*}
D\left(p_{S} \| q_{S}^{*}\right)=D\left(p_{S} \| q_{S}^{\mathcal{X} *}\right)+D\left(p_{S} \| q_{S}^{\mathcal{Y} *}\right) \tag{30}
\end{equation*}
$$

we obtain the relations

$$
\begin{align*}
D\left(p_{S} \| q_{S}^{\mathcal{X}} *\right. & =D\left(q_{S}^{\mathcal{Y}} \| q_{S}^{*}\right)  \tag{31}\\
D\left(p_{S} \| q_{S}^{\mathcal{Y}} *\right. & =D\left(q_{S}^{\mathcal{X} *} \| q_{S}^{*}\right) \tag{32}
\end{align*}
$$

from the Pythagorean theorem Eq. (30). Equations (31) and (32) imply that the parallel sides of a quadrangle have the same length. Therefore, the additivity Eq. (26) can be understood from the rectangle condition in information geometry (Fig. 4). The measures of information integration $\Phi_{\text {bath }}+$ $\Phi_{\text {SI }}-\Phi_{\text {SI }}^{\dagger}$ quantify a distortion of this rectangle.

## V. EXAMPLES

## A. The projection theorem for a single-spin model

We illustrate the main result, Eq. (9), using the single-spin model (see Appendix B). Let $\boldsymbol{Z}=\left\{S_{1}\right\}$ and $\boldsymbol{Z}^{\prime}=\left\{S_{2}\right\}$ be random variables of the spin at times $t$ and $t+d t$, respectively. Each spin has the binary state $s_{i} \in\{0,1\}$. The joint probability is generally given by the exponential family

$$
\begin{equation*}
p_{S}^{\hat{\theta}}\left(s_{1}, s_{2}\right)=\exp \left[\sum_{i} s_{i} \hat{\theta}^{i}+\sum_{i<j} s_{i} s_{j} \hat{\theta}^{i j}-\phi_{S}(\hat{\boldsymbol{\theta}})\right], \tag{33}
\end{equation*}
$$

where $\hat{\boldsymbol{\theta}}=\left\{\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right\}$ is the set of parameters and $\phi_{\boldsymbol{S}}(\hat{\boldsymbol{\theta}})$ is the normalization factor that satisfies $\sum_{s} p_{S}^{\hat{\theta}}(\boldsymbol{s})=1$. The number of the elements in $\hat{\boldsymbol{\theta}}$ is $\left(2^{2}-1\right)=3$, so the set of
probabilities $p_{S}^{\hat{\theta}}$ can be represented by a three-dimensional submanifold. The backward manifold $\mathcal{M}_{\mathrm{B}}$ is given by the constraint of the parameters

$$
\begin{equation*}
\mathcal{M}_{\mathrm{B}}=\left\{p_{S}^{\theta} \mid \theta^{1}=\hat{\theta^{2}}, \theta^{12}=\theta^{\hat{1} 2}\right\} . \tag{34}
\end{equation*}
$$

Because the free parameter is $\theta^{2}$, the backward manifold for the single-spin model is one-dimensional.

Our result (9) can be rewritten as the optimization problem of $\theta^{2}$,

$$
\begin{align*}
\sigma_{\mathrm{tot}}^{\mathcal{Z}}= & \left.\min _{\theta^{2}} D\left(p_{S}^{\hat{\theta}} \| p_{S}^{\theta}\right)\right|_{\theta^{1}=\hat{\theta}^{2}, \theta^{12}=\theta^{\hat{1} 2}}  \tag{35}\\
= & \mathbb{E}\left[s_{1}\right]\left(\hat{\theta}^{1}-\hat{\theta}^{2}\right)-\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right) \\
& +\min _{\theta^{2}}\left\{\mathbb{E}\left[s_{2}\right]\left(\hat{\theta}^{2}-\theta^{2}\right)+\phi_{S}\left(\hat{\theta}^{2}, \theta^{2}, \hat{\theta}^{12}\right)\right\} \tag{36}
\end{align*}
$$

This problem can be numerically solved by using a conventional optimization tool.

## B. The projection theorem for a two-spin model

We next illustrate our results using the two-spin model (see Appendix C). Let $\boldsymbol{Z}=\left\{S_{1}, S_{2}\right\}$ and $\boldsymbol{Z}^{\prime}=\left\{S_{3}, S_{4}\right\}$ be random variables of two spins at times $t$ and $t+d t$, respectively. The spin has the binary state $s_{i} \in\{0,1\}$. We assume the situation where both bipartite conditions $\mathcal{C}_{\mathrm{BI}}$ and $\mathcal{C}_{\mathrm{BI}}^{*}$ hold. Under the bipartite conditions, the joint probability of the spin state is generally given by the exponential family

$$
\begin{align*}
p_{S}^{\hat{\theta}}(\boldsymbol{s})= & \exp \left[\sum_{i} s_{i} \hat{\theta}^{i}+s_{1} s_{3} \hat{\theta}^{13}+s_{1} s_{4} \hat{\theta}^{14}\right. \\
& \left.+s_{2} s_{3} \hat{\theta}^{23}+s_{2} s_{4} \hat{\theta}^{24}-\phi_{S}(\hat{\boldsymbol{\theta}})\right] \tag{37}
\end{align*}
$$

The backward manifold is given by the constraint of the parameters

$$
\begin{gather*}
\mathcal{M}_{\mathrm{B}}=\left\{p_{\boldsymbol{S}}^{\boldsymbol{\theta}} \mid \boldsymbol{\theta}^{\mathcal{X}}=\hat{\boldsymbol{\theta}}^{\mathcal{X}}, \boldsymbol{\theta}^{\mathcal{Y}}=\hat{\boldsymbol{\theta}}^{\mathcal{Y}}\right\},  \tag{38}\\
\boldsymbol{\theta}^{\mathcal{X}}=\left(\theta^{1}, \theta^{13}, \theta^{14}\right), \quad \hat{\boldsymbol{\theta}}^{\mathcal{X}}=\left(\hat{\theta}^{3}, \hat{\theta}^{13}, \hat{\theta}^{23}\right),  \tag{39}\\
\boldsymbol{\theta}^{\mathcal{Y}}=\left(\theta^{2}, \theta^{24}, \theta^{23}\right), \quad \hat{\boldsymbol{\theta}}^{\mathcal{Y}}=\left(\hat{\theta}^{4}, \hat{\theta}^{24}, \hat{\theta}^{14}\right), \tag{40}
\end{gather*}
$$

where a coordinate $\boldsymbol{\theta}$ represents a probability on the backward manifold. Because free parameters are $\left\{\theta^{3}, \theta^{4}\right\}$, the backward manifold for the two-spin model is two-dimensional. The condition of the local backward manifolds is also given by the linear constraint of $\boldsymbol{\theta}$,

$$
\begin{equation*}
\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}=\left\{p_{\boldsymbol{S}}^{\boldsymbol{\theta}} \mid \boldsymbol{\theta}^{\mathcal{X}}=\hat{\boldsymbol{\theta}}^{\mathcal{X}}\right\}, \quad \mathcal{M}_{\mathrm{LB}}^{\mathcal{Y}}=\left\{p_{S}^{\boldsymbol{\theta}} \mid \boldsymbol{\theta}^{\mathcal{Y}}=\hat{\boldsymbol{\theta}}^{\mathcal{Y}}\right\} \tag{41}
\end{equation*}
$$

Because the free parameters are $\left\{\theta^{3}, \theta^{4}, \boldsymbol{\theta}^{\mathcal{Y}}\right\}\left(\left\{\theta^{3}, \theta^{4}, \boldsymbol{\theta}^{\mathcal{X}}\right\}\right)$, the local backward manifold $\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}\left(\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}\right)$ is five-dimensional. The intersection of these two local backward manifolds is the backward manifold $\mathcal{M}_{\mathrm{B}}=\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}} \cap \mathcal{M}_{\mathrm{LB}}^{\mathcal{Y}}$. The total entropy production and the partial entropy productions are obtained from the optimization problems

$$
\begin{align*}
& \sigma_{\text {tot }}^{\mathcal{Z}}=\left.\min _{\theta^{3}, \theta^{4}} D\left(p_{S}^{\hat{\theta}} \| p_{S}^{\theta}\right)\right|_{\theta^{\mathcal{X}}=\hat{\theta}^{\mathcal{X}}, \theta^{y}=\hat{\theta}^{\mathcal{y}}},  \tag{42}\\
& \sigma_{\text {partial }}^{\mathcal{X}}=\left.\min _{\theta^{3}, \theta^{4}, \theta^{\mathcal{y}}} D\left(p_{S}^{\hat{\theta}} \| p_{S}^{\theta}\right)\right|_{\theta^{\chi}=\hat{\theta}^{\mathcal{X}}}  \tag{43}\\
& \sigma_{\text {partial }}^{\mathcal{Y}}=\left.\min _{\theta^{3}, \theta^{4}, \theta^{\chi}} D\left(p_{S}^{\hat{\theta}} \| p_{S}^{\theta}\right)\right|_{\theta^{y}=\hat{\theta}^{y}} \tag{44}
\end{align*}
$$

Without the bipartite conditions $\mathcal{C}_{\mathrm{BI}}$ and $\mathcal{C}_{\mathrm{BI}}^{*}$, the joint probability is generally given by

$$
\begin{align*}
p_{\boldsymbol{S}}^{\hat{\boldsymbol{\theta}}}(\boldsymbol{s})= & \exp \left[\sum_{i} s_{i} \hat{\theta}^{i}+\sum_{i<j} s_{i} s_{j} \hat{\theta}^{i j}+\sum_{i<j<k} s_{i} s_{j} s_{k} \hat{\theta}^{i j k}\right. \\
& \left.+\sum_{i<j<k<l} s_{i} s_{j} s_{k} s_{l} \hat{\theta}^{i j k l}-\phi_{\boldsymbol{S}}(\hat{\boldsymbol{\theta}})\right] \tag{45}
\end{align*}
$$

If the vector $\left(\hat{\theta}^{12}, \hat{\theta}^{34}, \hat{\theta}^{123}, \hat{\theta}^{134}, \hat{\theta}^{124}, \hat{\theta}^{234}, \hat{\theta}^{1234}\right)$ is nonzero, the bipartite conditions are violated, and measures of information integration $\Phi_{\mathrm{SI}}, \Phi_{\mathrm{SI}}^{\dagger}$, and $\Phi_{\text {bath }}$ have nonzero values.

## VI. CONCLUSION AND DISCUSSION

By applying the information-geometric framework, we showed the relation between the entropy production and the stochastic interaction. Our result could be the foundation of the integrated information theory based on the physical law. We may discuss the thermodynamic cost of the information integration based on this framework.

Because the second law of information thermodynamics is essential for biochemical information processing [45,54-59], this work would give geometric insight into biochemical information processing. This work provides the physical validity of the integrated information theory [25,26,30,32] for biochemical information processing.

From the viewpoint of thermodynamics, our results are complementary to other geometric expressions of the second law, such as the principle of Carathèodory [60] and the maximum entropy thermodynamics [61,62]. Our framework would be applicable to other generalizations of the entropy production, for example, thermodynamics under feedback control by selecting the backward manifolds for feedback control (see Appendix D).

## ACKNOWLEDGMENTS

We are grateful to H. Shimazaki for critical reading of the old version of this manuscript. We are also grateful to A. Dechant for the discussion of information geometry and the thermodynamic uncertainty. We thank K. Kaneko, T. Sagawa, and T. Hatakeyama for valuable comments. S.I. is supported by JSPS KAKENHI Grants No. JP16K17780 and No. JP19H05796 and JST Presto Grant No. JPMJPR18M2, Japan. M.O. is supported by JST CREST Grant No. JPMJCR1864 and JSPS KAKENHI Grant No. 18H02713, Japan.

## APPENDIX A: THE SECOND LAW OF THERMODYNAMICS FOR THE MASTER EQUATION

We here review the second law of thermodynamics in stochastic thermodynamics. We start with the master equation

$$
\begin{equation*}
\frac{d}{d t} p\left(z^{\prime} ; t\right)=\sum_{z}\left[W\left(z \rightarrow z^{\prime} ; t\right) p(z ; t)-W\left(z^{\prime} \rightarrow z ; t\right) p\left(z^{\prime} ; t\right)\right] \tag{A1}
\end{equation*}
$$

where $p(z ; t)$ is the probability of state $z$ at time $t$ and $W(z \rightarrow$ $\left.z^{\prime} ; t\right)$ is the transition rate from state $z$ to state $z^{\prime}$ at time $t$. In the notation of this paper, the probability of $z$ is given by $p_{Z}(z)=p(z ; t)$. From the master equation (A1), we obtain the probability at time $t+d t$,

$$
\begin{align*}
p\left(z^{\prime} ; t+d t\right)= & \sum_{z}\left\{W\left(z \rightarrow z^{\prime} ; t\right) p(z ; t) d t\right. \\
& \left.+\left[1-W\left(z^{\prime} \rightarrow z ; t\right) d t\right] p\left(z^{\prime} ; t\right)\right\} . \tag{A2}
\end{align*}
$$

In the notation of the main text, $p_{\boldsymbol{Z}}(z)$ and $p_{Z^{\prime}}\left(z^{\prime}\right)$ are given by $p_{\boldsymbol{Z}}(\boldsymbol{z})=p(\boldsymbol{z} ; t)$ and $p_{\boldsymbol{Z}^{\prime}}\left(\boldsymbol{z}^{\prime}\right)=p\left(\boldsymbol{z}^{\prime} ; t+d t\right)$, respectively. We also obtain the relation between $p_{\boldsymbol{Z}}$ and $p_{Z^{\prime}}$ as

$$
\begin{equation*}
p_{Z^{\prime}}\left(z^{\prime}\right)=p\left(z^{\prime} ; t\right)+O(d t)=p_{Z}\left(z^{\prime}\right)+O(d t) . \tag{A3}
\end{equation*}
$$

The transition probability $T\left(z^{\prime}, z\right)$ is given by

$$
T\left(z^{\prime}, z\right)= \begin{cases}W\left(z \rightarrow z^{\prime} ; t\right) d t & \left(z \neq z^{\prime}\right)  \tag{A4}\\ 1-\sum_{z \neq z^{\prime}} W\left(z^{\prime} \rightarrow z ; t\right) d t & \left(z=z^{\prime}\right)\end{cases}
$$

Here, we consider the detailed balance. The condition of the detailed balance is given by

$$
\begin{equation*}
W\left(z \rightarrow z^{\prime} ; t\right) p(z ; t)=W\left(z^{\prime} \rightarrow z ; t\right) p\left(z^{\prime} ; t\right) \tag{A5}
\end{equation*}
$$

for any $z$ and $z^{\prime}$. This condition is valid if the system is in equilibrium. By using the transition probability Eq. (A4), we obtain another expression of the detailed balance condition (A5) as

$$
\begin{equation*}
T\left(z^{\prime}, \boldsymbol{z}\right) p_{\boldsymbol{Z}}(\boldsymbol{z})=T\left(\boldsymbol{z}, \boldsymbol{z}^{\prime}\right) p_{Z^{\prime}}\left(z^{\prime}\right), \tag{A6}
\end{equation*}
$$

where we used $W\left(z^{\prime} \rightarrow z ; t\right) p\left(z^{\prime} ; t\right) d t=T\left(z, z^{\prime}\right) p_{Z}\left(z^{\prime}\right)=$ $T\left(z, z^{\prime}\right) p_{Z^{\prime}}\left(z^{\prime}\right)+O\left(d t^{2}\right)$. Therefore, the detailed balance condition (A5) implies the reversibility of dynamics in the transition from $t$ to $t+d t$. From the identity by Bayes's rule

$$
\begin{equation*}
p_{\mathbf{Z} \mid \mathbf{Z}^{\prime}}\left(\boldsymbol{z} \mid \boldsymbol{z}^{\prime}\right)=T\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right) \frac{p_{\mathbf{Z}}(\boldsymbol{z})}{p_{\boldsymbol{Z}^{\prime}}\left(\boldsymbol{z}^{\prime}\right)}, \tag{A7}
\end{equation*}
$$

the detailed balance condition (A5) can be rewritten as

$$
\begin{equation*}
T\left(z, z^{\prime}\right)=p_{Z \mid Z^{\prime}}\left(z \mid z^{\prime}\right) . \tag{A8}
\end{equation*}
$$

Next, we discuss the second law of thermodynamics. For the master equation, the total entropy production ratio $\sigma_{\mathrm{tot}}^{\mathcal{Z}} / d t$ is defined as

$$
\begin{equation*}
\frac{\sigma_{\mathrm{tot}}^{\mathcal{Z}}}{d t}=\sum_{z, z^{\prime}} W\left(z \rightarrow z^{\prime} ; t\right) p(z ; t) \ln \frac{W\left(z \rightarrow z^{\prime} ; t\right) p(z ; t)}{W\left(z^{\prime} \rightarrow z ; t\right) p\left(z^{\prime} ; t\right)} . \tag{A9}
\end{equation*}
$$

If the detailed balance condition is valid, the entropy production vanishes, $\sigma_{\mathrm{tot}}^{\mathcal{Z}}=0$. By using the transition probability $T\left(\boldsymbol{z}^{\prime} \mid \boldsymbol{z}\right)$, we obtain another expression of the total entropy production

$$
\begin{align*}
\sigma_{\mathrm{tot}}^{\mathcal{Z}} & =\sum_{z, z^{\prime} \mid z \neq z^{\prime}} T\left(z^{\prime}, z\right) p_{\mathbf{Z}}(z) \ln \frac{T\left(z^{\prime}, \boldsymbol{z}\right) p_{\mathbf{Z}}(\boldsymbol{z})}{T\left(\boldsymbol{z}, \boldsymbol{z}^{\prime}\right) p_{Z^{\prime}}\left(\boldsymbol{z}^{\prime}\right)}+O\left(d t^{2}\right)  \tag{A10}\\
& =\sum_{z, z^{\prime}} T\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right) p_{\boldsymbol{Z}}(\boldsymbol{z}) \ln \frac{T\left(z^{\prime}, \boldsymbol{z}\right) p_{\mathbf{Z}}(\boldsymbol{z})}{T\left(\boldsymbol{z}, \boldsymbol{z}^{\prime}\right) p_{Z^{\prime}}\left(\boldsymbol{z}^{\prime}\right)} \tag{A11}
\end{align*}
$$

To introduce two probabilities, $p_{\boldsymbol{S}}(\boldsymbol{s})=T\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right) p_{\mathrm{Z}}(\boldsymbol{z})$ and $q_{\boldsymbol{S}}^{*}(\boldsymbol{s})=T\left(\boldsymbol{z}, \boldsymbol{z}^{\prime}\right) p_{\boldsymbol{Z}^{\prime}}\left(\boldsymbol{z}^{\prime}\right)$, with $\boldsymbol{S}=\left\{\boldsymbol{Z}, \boldsymbol{Z}^{\prime}\right\}$ and $\boldsymbol{s}=\left\{\boldsymbol{z}, \boldsymbol{z}^{\prime}\right\}$, this expression of the total entropy production (A11) can be
regarded as the Kullback-Leibler divergence between two probabilities,

$$
\begin{align*}
\sigma_{\mathrm{tot}}^{\mathcal{Z}} & =\sum_{\boldsymbol{s}} p_{S}(\boldsymbol{s}) \ln \frac{p_{\boldsymbol{S}}(\boldsymbol{s})}{q_{\boldsymbol{S}}^{*}(\boldsymbol{s})}  \tag{A12}\\
& =D\left(p_{S} \| q_{S}^{*}\right) . \tag{A13}
\end{align*}
$$

From the nonnegativity of the Kullback-Leibler divergence, we obtain the second law of thermodynamics for the master equation,

$$
\begin{equation*}
\sigma_{\mathrm{tot}}^{\mathcal{Z}} \geqslant 0 . \tag{A14}
\end{equation*}
$$

## APPENDIX B: THE PROJECTION THEOREM FOR A SINGLE-SPIN MODEL

We here show a detailed calculation of the single-spin model. The spin state at time $t$ is $z=s_{1} \in\{0,1\}$, and the spin state at time $t+d t$ is $z^{\prime}=s_{2} \in\{0,1\}$. We here start with the master equation

$$
\begin{equation*}
\frac{d}{d t} p\left(s^{\prime} ; t\right)=\sum_{s}\left[W\left(s \rightarrow s^{\prime} ; t\right) p(s ; t)-W\left(s^{\prime} \rightarrow s ; t\right) p\left(s^{\prime} ; t\right)\right] \tag{B1}
\end{equation*}
$$

where $p(s ; t)$ is the probability of state $s$ at time $t$ and $W(s \rightarrow$ $\left.s^{\prime} ; t\right)$ is the transition rate from $s$ to $s^{\prime}$ at time $t$. The transition probability $T\left(s_{2}, s_{1}\right)$ is given by

$$
T\left(s_{2}, s_{1}\right)= \begin{cases}1-W(0 \rightarrow 1 ; t) d t & \left(s_{1}=0, s_{2}=0\right)  \tag{B2}\\ W(0 \rightarrow 1 ; t) d t & \left(s_{1}=0, s_{2}=1\right) \\ W(1 \rightarrow 0 ; t) d t & \left(s_{1}=1, s_{2}=0\right) \\ 1-W(1 \rightarrow 0 ; t) d t & \left(s_{1}=1, s_{2}=1\right)\end{cases}
$$

The joint probability $p_{S}(\boldsymbol{s})$ is given by

$$
\begin{gather*}
p_{S}(\boldsymbol{s})=T\left(s_{2}, s_{1}\right) p\left(s_{1} ; t\right)  \tag{B3}\\
= \begin{cases}{[1-W(0 \rightarrow 1 ; t) d t] p(0 ; t)} & \left(s_{1}=0, s_{2}=0\right), \\
W(0 \rightarrow 1 ; t) d t p(0 ; t) & \left(s_{1}=0, s_{2}=1\right), \\
W(1 \rightarrow 0 ; t) d t[1-p(0 ; t)] & \left(s_{1}=1, s_{2}=0\right), \\
{[1-W(1 \rightarrow 0 ; t) d t][1-p(0 ; t)]} & \left(s_{1}=1, s_{2}=1\right) .\end{cases} \tag{B4}
\end{gather*}
$$

Here we introduce the joint probability $p_{S}^{\hat{\boldsymbol{\theta}}}(\boldsymbol{s})$ as the exponential family

$$
\begin{gather*}
p_{\boldsymbol{S}}^{\hat{\theta}}(\boldsymbol{s})=e^{\hat{\theta}^{1} s_{1}+\hat{\theta}^{2} s_{2}+\hat{\theta}^{12} s_{1} s_{2}-\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)},  \tag{B5}\\
\phi_{\boldsymbol{S}}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)=\ln \left[1+e^{\hat{\theta}^{1}}+e^{\hat{\theta}^{2}}+e^{\hat{\theta}^{1}+\hat{\theta}^{2}+\hat{\theta}^{12}}\right], \tag{B6}
\end{gather*}
$$

which implies

$$
p_{S}^{\hat{\theta}}(\boldsymbol{s})= \begin{cases}e^{-\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)} & \left(s_{1}=0, s_{2}=0\right),  \tag{B7}\\ e^{\hat{\theta}^{2}-\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)} & \left(s_{1}=0, s_{2}=1\right), \\ e^{\hat{\theta}^{1}-\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)} & \left(s_{1}=1, s_{2}=0\right) \\ e^{\hat{\theta}^{1}+\hat{\theta}^{2}+\hat{\theta}^{12}-\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)} & \left(s_{1}=1, s_{2}=1\right)\end{cases}
$$

The transition probability $T\left(s_{2}, s_{1}\right)=p_{S}^{\hat{\theta}}(\boldsymbol{s}) /\left[\sum_{s_{2}} p_{S}^{\hat{\theta}}(\boldsymbol{s})\right]$ is
given by

$$
\begin{align*}
T\left(s_{2}, s_{1}\right) & =e^{\hat{\theta}^{2} s_{2}+\hat{\theta}^{12} s_{1} s_{2}-\phi_{s_{2} \mid s_{1}}\left(s_{1} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)} \\
\phi_{S_{2} \mid S_{1}}\left(s_{1} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right) & =\ln \left[1+e^{\hat{\theta}^{2}+\hat{\theta}^{12} s_{1}}\right] \tag{B8}
\end{align*}
$$

Because of one-to-one correspondence, we identify $p_{S}(\boldsymbol{s})$ with $p_{S}^{\hat{\theta}}(\boldsymbol{s})$. From Eqs. (B4) and (B7), we obtain the relation between $\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)$ and $(W(0 \rightarrow 1 ; t), W(1 \rightarrow 0 ; t), p(0 ; t))$ as

$$
\begin{align*}
& \phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)=\ln \frac{1}{p_{S}(0,0)} \\
& \quad=-\ln \{[1-W(0 \rightarrow 1 ; t) d t] p(0 ; t)\},  \tag{B9}\\
& \begin{aligned}
& \hat{\theta}^{1}= \phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)+\ln \{W(1 \rightarrow 0 ; t) d t[1-p(0 ; t)]\} \\
&= \ln \frac{p_{S}(1,0)}{p_{S}(0,0)} \\
&= \ln \frac{W(1 \rightarrow 0 ; t) d t[1-p(0 ; t)]}{[1-W(0 \rightarrow 1 ; t) d t] p(0 ; t)}, \\
& \hat{\theta}^{2}=\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)+\ln [W(0 \rightarrow 1 ; t) d t p(0 ; t)] \\
&=\ln \frac{p_{S}(0,1)}{p_{S}(0,0)} \\
& \quad=\ln \frac{W(0 \rightarrow 1 ; t) d t}{1-W(0 \rightarrow 1 ; t) d t}, \\
& \hat{\theta}^{12}= \phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)-\hat{\theta}^{1}-\hat{\theta}^{2} \\
& \quad+\ln \{[1-W(1 \rightarrow 0 ; t) d t][1-p(0 ; t)]\} \\
&= \ln \frac{p_{S}(0,0) p_{S}(1,1)}{p_{S}(0,1) p_{S}(1,0)} \\
&= \ln \frac{[1-W(0 \rightarrow 1 ; t) d t][1-W(1 \rightarrow 0 ; t) d t]}{[W(0 \rightarrow 1 ; t) d t][W(1 \rightarrow 0 ; t) d t]} .
\end{aligned}
\end{align*}
$$

The backward manifold is defined as

$$
\begin{equation*}
\mathcal{M}_{\mathrm{B}}=\left\{q_{S} \mid q_{S}(\boldsymbol{s})=q_{S_{2}}\left(s_{2}\right) T\left(s_{1}, s_{2}\right)\right\} . \tag{B13}
\end{equation*}
$$

If we use an expression of the exponential family for $q_{S}(\boldsymbol{s})=$ $p_{S}^{\theta}(s)$, the reversible manifold is given by

$$
\begin{equation*}
\mathcal{M}_{\mathrm{B}}=\left\{p_{S}^{\theta}(s) \mid \theta^{1}=\hat{\theta}^{2}, \theta^{12}=\hat{\theta}^{12}\right\} \tag{B14}
\end{equation*}
$$

because the condition $q_{\boldsymbol{S}}(\boldsymbol{s})=q_{S_{2}}\left(s_{2}\right) T\left(s_{1} \mid s_{2}\right)$ can be written as

$$
\begin{gather*}
e^{\theta^{1} s_{1}+\theta^{12} s_{1} s_{2}-\phi_{S_{1} \mid S_{2}}\left(s_{2} \mid \theta^{1}, \theta^{12}\right)}=e^{\hat{\theta}^{2} s_{1}+\hat{\theta}^{12} s_{2} s_{1}-\phi_{S_{2} \mid s_{1}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)},  \tag{B15}\\
\phi_{S_{1} \mid S_{2}}\left(s_{1} \mid \theta^{1}, \theta^{12}\right)=\ln \left[1+e^{\theta^{1}+\theta^{12} s_{1}}\right] \tag{B16}
\end{gather*}
$$

We obtain the following Pythagorean theorem for any $q_{S} \in$ $\mathcal{M}_{\mathrm{B}}$ :

$$
\begin{align*}
D\left(p_{S} \| q_{S}\right) & =D\left(p_{S} \| q_{S}^{*}\right)+D\left(q_{S}^{*} \| q_{S}\right) \\
q_{S}^{*}(\boldsymbol{s}) & =e^{\hat{\theta}^{2} s_{1}+\theta^{2 *} s_{2}+\hat{\theta}^{12} s_{1} s_{2}-\phi_{S}\left(\hat{\theta}^{2}, \theta^{2 *}, \hat{\theta}^{12}\right)} \tag{B17}
\end{align*}
$$

with the constraint

$$
\begin{equation*}
\sum_{s_{1}} q_{S}^{*}(\boldsymbol{s})=\sum_{s_{1}} p_{S}(\boldsymbol{s}) \tag{B18}
\end{equation*}
$$

In our main result, the total entropy production is given by the following optimization problem:

$$
\begin{equation*}
\sigma_{\mathrm{tot}}^{\mathcal{Z}}=D^{\mathrm{opt}}\left(p_{\boldsymbol{S}} \| \mathcal{M}_{\mathrm{B}}\right)=D\left(p_{\boldsymbol{S}} \| q_{\mathrm{S}}^{*}\right) \tag{B19}
\end{equation*}
$$

By using the expression by $\left(\theta_{1}, \theta_{2}, \theta_{12}\right)$, this optimization problem can be written as

$$
\begin{align*}
\sigma_{\text {tot }}^{\mathcal{Z}}= & \min _{q_{S} \in \mathcal{M}_{\mathrm{B}}} D\left(p_{S} \| q_{S}\right)  \tag{B20}\\
= & \mathbb{E}\left[s_{1}\right]\left(\hat{\theta}^{1}-\hat{\theta}^{2}\right)-\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right) \\
& +\min _{\theta^{2}}\left\{\mathbb{E}\left[s_{2}\right]\left(\hat{\theta}^{2}-\theta^{2}\right)+\phi_{S}\left(\hat{\theta}^{2}, \theta^{2}, \hat{\theta}^{12}\right)\right\}  \tag{B21}\\
= & \mathbb{E}\left[s_{1}\right]\left(\hat{\theta}^{1}-\hat{\theta}^{2}\right)+\mathbb{E}\left[s_{2}\right]\left(\hat{\theta}^{2}-\theta^{2 *}\right) \\
& -\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2 *}, \hat{\theta}^{12}\right)+\phi_{S}\left(\hat{\theta}^{2}, \theta^{2 *}, \hat{\theta}^{12}\right), \tag{B22}
\end{align*}
$$

where $\mathbb{E}$ denotes the expected value $\mathbb{E}[\cdots]=\sum_{s} p_{S}(\boldsymbol{s}) \cdots$. The constraint (B18) is calculated as

$$
\begin{align*}
& e^{\left(\hat{\theta}^{2}-\theta^{2 *}\right) s_{2}-\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)+\phi_{S}\left(\hat{\theta}^{2}, \theta^{2 *}, \hat{\theta}^{12}\right)} \\
& \quad=\exp ^{\phi_{S_{1} \mid s_{2}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)-\phi_{s_{1} \mid s_{2}}\left(s_{2} \hat{\theta}^{1}, \hat{\theta}^{12}\right)} . \tag{B23}
\end{align*}
$$

Under the constraint (B23), the optimization problem (B22) is calculated as

$$
\begin{align*}
\sigma_{\mathrm{tot}}^{\mathcal{Z}}= & \mathbb{E}\left[s_{1}\left(\hat{\theta}^{1}-\hat{\theta}^{2}\right)+s_{2}\left(\hat{\theta}^{2}-\theta^{2 *}\right)\right] \\
& -\mathbb{E}\left[\phi_{S}\left(\hat{\theta}^{1}, \hat{\theta}^{2}, \hat{\theta}^{12}\right)+\phi_{S}\left(\hat{\theta}^{2}, \theta^{2 *}, \hat{\theta}^{12}\right)\right] \\
= & \mathbb{E}\left[s_{1}\left(\hat{\theta}^{1}-\hat{\theta}^{2}\right)+\phi_{S_{1} \mid S_{2}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)-\phi_{S_{1} \mid S_{2}}\left(s_{2} \mid \hat{\theta}^{1}, \hat{\theta}^{12}\right)\right] . \tag{B24}
\end{align*}
$$

We can check the equivalence between Eq. (B24) and the original definition of the total entropy production as follows:

$$
\begin{align*}
\sigma_{\mathrm{tot}}^{\mathcal{Z}} & =\sum_{s} T\left(s_{2}, s_{1}\right) p_{S_{1}}\left(s_{1}\right) \ln \frac{T\left(s_{2}, s_{1}\right) p_{S_{1}}\left(s_{1}\right)}{T\left(s_{1}, s_{2}\right) p_{S_{2}}\left(s_{2}\right)} \\
& =\mathbb{E}\left[\ln \frac{T\left(s_{2}, s_{1}\right) p_{S_{1}}\left(s_{1}\right)}{T\left(s_{1} \mid s_{2}\right) p_{S_{2}}\left(s_{2}\right)}\right] \\
& =\mathbb{E}\left[\ln \frac{p_{S_{1} \mid S_{2}}\left(s_{1} \mid s_{2}\right)}{\exp \left[\hat{\theta}^{2} s_{1}+\hat{\theta}^{12} s_{1} s_{2}-\phi_{S_{2} \mid S_{1}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)\right]}\right] \\
& =\mathbb{E}\left[\ln \frac{\exp \left[\hat{\theta}^{1} s_{1}+\hat{\theta}^{12} s_{1} s_{2}-\phi_{S_{1} \mid S_{2}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)\right]}{\exp \left[\hat{\theta}^{2} s_{1}+\hat{\theta}^{12} s_{1} s_{2}-\phi_{S_{2} \mid S_{1}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)\right]}\right] \\
& =\mathbb{E}\left[s_{1}\left(\hat{\theta}^{1}-\hat{\theta}^{2}\right)+\phi_{S_{1} \mid S_{2}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)-\phi_{S_{1} \mid S_{2}}\left(s_{2} \mid \hat{\theta}^{1}, \hat{\theta}^{12}\right)\right], \tag{B25}
\end{align*}
$$

where we used $\phi_{S_{2} \mid S_{1}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)=\phi_{S_{1} \mid S_{2}}\left(s_{2} \mid \hat{\theta}^{2}, \hat{\theta}^{12}\right)$.

## APPENDIX C: THE PROJECTION THEOREM FOR A TWO-SPIN MODEL

We start with the joint distribution

$$
\begin{align*}
p_{\boldsymbol{S}}^{\hat{\theta}}(\boldsymbol{s})= & \exp \left[\sum_{i} s_{i} \hat{\theta}^{i}+\sum_{i<j} s_{i} s_{j} \hat{\theta}^{i j}+\sum_{i<j<k} s_{i} s_{j} s_{k} \hat{\theta}^{i j k}\right. \\
& \left.+\sum_{i<j<k<l} s_{i} s_{j} s_{k} s_{l} \hat{\theta}^{i j k l}-\phi_{\boldsymbol{S}}(\hat{\boldsymbol{\theta}})\right] \tag{C1}
\end{align*}
$$

where $\boldsymbol{s}=\left(s_{1}, s_{2}, s_{3}, s_{4}\right)=\left(x, y, x^{\prime}, y^{\prime}\right)$ is the spin notation, with $s_{i} \in\{0,1\}$, and $\phi_{S}(\hat{\boldsymbol{\theta}})$ is the normalization constant that satisfies $\sum_{s} p_{S}^{\hat{\theta}}(\boldsymbol{s})=1$.

We consider both bipartite conditions $\mathcal{C}_{\mathrm{BI}}$ and $\mathcal{C}_{\mathrm{BI}}^{*}$. We here compare $p_{X^{\prime} \mid \mathbf{Z}}^{\hat{\theta}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}\right)=\sum_{s_{4}} p_{S}^{\hat{\theta}}(\boldsymbol{s}) /\left[\sum_{s_{3}, s_{4}} p_{S}^{\hat{\theta}}(\boldsymbol{s})\right]$ with $p_{\boldsymbol{X}^{\prime} \mid \boldsymbol{Z}, \boldsymbol{Y}^{\prime}}^{\hat{\theta}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}, \boldsymbol{y}^{\prime}\right)=p_{S}^{\hat{\theta}}(\boldsymbol{s}) /\left[\sum_{s_{3}} p_{S}^{\hat{\theta}}(\boldsymbol{s})\right]$. The conditional probability $p_{X^{\prime} \mid Z^{\prime}}^{\hat{\theta}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}\right)$ is calculated as

$$
\begin{gather*}
p_{\boldsymbol{X}^{\prime} \mid \boldsymbol{Z}}^{\hat{\theta}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}\right)=e^{s_{3} \hat{\theta}^{3}+s_{1} s_{3} \hat{\theta}^{13}+s_{2} s_{3} \hat{\theta}^{23}+s_{1} s_{2} s_{3} \hat{\theta}^{123}-\phi_{\boldsymbol{X}^{\prime} \mid \mathbf{Z}}\left(s_{1}, s_{2} \mid \hat{\boldsymbol{\theta}}\right)},  \tag{C2}\\
\phi_{\boldsymbol{X}^{\prime} \mid \boldsymbol{Z}}\left(s_{1}, s_{2} \mid \hat{\boldsymbol{\theta}}\right):=\ln \left[e^{\hat{\theta}^{3}+s_{1} \hat{\theta}^{13}+s_{2} \hat{\theta}^{23}+s_{1} s_{2} \hat{\theta}^{23}}+1\right] . \tag{C3}
\end{gather*}
$$

The conditional probability $p_{\boldsymbol{X}^{\prime} \mid \boldsymbol{Z}, \boldsymbol{Y}^{\prime}}^{\hat{\theta}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}, \boldsymbol{y}^{\prime}\right)$ is calculated as

$$
\begin{align*}
& p_{\boldsymbol{X}^{\prime} \mid \mathbf{Z}, \boldsymbol{Y}^{\prime}}^{\hat{\theta}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}, \boldsymbol{y}^{\prime}\right) \\
& \quad=\exp \left[s_{3} \hat{\theta}^{3}+s_{1} s_{3} \hat{\theta}^{13}+s_{2} s_{3} \hat{\theta}^{23}+s_{3} s_{4} \hat{\theta}^{34}+s_{1} s_{2} s_{3} \hat{\theta}^{123}\right. \\
& \quad+s_{1} s_{3} s_{4} \hat{\theta}^{134}+s_{2} s_{3} s_{4} \hat{\theta}^{234}+s_{1} s_{2} s_{3} s_{4} \hat{\theta}^{1234} \\
& \left.\quad-\phi_{\boldsymbol{X}^{\prime} \mid \mathbf{Z}, \boldsymbol{Y}^{\prime}}\left(s_{1}, s_{2}, s_{4} \mid \hat{\boldsymbol{\theta}}\right)\right], \tag{C4}
\end{align*}
$$

$$
\begin{align*}
& \phi_{X^{\prime} \mid \mathbf{Z}, \boldsymbol{Y}^{\prime}}\left(s_{1}, s_{2}, s_{4} \mid \hat{\boldsymbol{\theta}}\right) \\
& \quad:=\ln \left[\operatorname { e x p } \left(\hat{\theta}^{3}+s_{1} \hat{\theta}^{13}+s_{2} \hat{\theta}^{23}+s_{4} \hat{\theta}^{34}+s_{1} s_{2} \hat{\theta}^{123}\right.\right. \\
& \left.\left.\quad+s_{1} s_{4} \hat{\theta}^{134}+s_{2} s_{4} \hat{\theta}^{234}+s_{1} s_{2} s_{4} \hat{\theta}^{1234}\right)+1\right] . \tag{C5}
\end{align*}
$$

From Eqs. (C3) and (C5), we obtain the condition of $\mathcal{C}_{\mathrm{BI}}$ : $p_{X^{\prime} \mid \mathbf{Z}, \boldsymbol{Y}^{\prime}}^{\hat{\theta}}=p_{\boldsymbol{X}^{\prime} \mid \mathbf{Z}}^{\hat{\theta}}$ as

$$
\begin{equation*}
\mathcal{C}_{\mathrm{BI}}: \hat{\theta}^{34}=\hat{\theta}^{134}=\hat{\theta}^{234}=\hat{\theta}^{1234}=0 . \tag{C6}
\end{equation*}
$$

In the same way, we also obtain the condition of $\mathcal{C}_{\mathrm{BI}}^{*}$ as

$$
\begin{equation*}
\mathcal{C}_{\mathrm{BI}}^{*}: \hat{\theta}^{12}=\hat{\theta}^{123}=\hat{\theta}^{124}=\hat{\theta}^{1234}=0 . \tag{C7}
\end{equation*}
$$

To clarify the relation between $\mathcal{C}_{\mathrm{BI}}$ and $\mathcal{C}_{\mathrm{BI}}^{*}$, we can consider the permutation $(\alpha(1), \alpha(2), \alpha(3), \alpha(4))=(3,4,1,2)$. The condition of $\mathcal{C}_{\mathrm{BI}}^{*}$ is given by the condition of $\mathcal{C}_{\mathrm{BI}}$ with the permutation $\alpha$,

$$
\begin{align*}
\mathcal{C}_{\mathrm{BI}}^{*} & : \hat{\theta}^{\alpha(3) \alpha(4)}=\hat{\theta}^{\alpha(3) \alpha(4) \alpha(1)}=\hat{\theta}^{\alpha(3) \alpha(4) \alpha(2)} \\
& =\hat{\theta}^{\alpha(3) \alpha(4) \alpha(1) \alpha(2)}=0 . \tag{C8}
\end{align*}
$$

Next, we discuss the backward manifold $\mathcal{M}_{\mathrm{B}}$. The transition probability $T\left(z^{\prime}, \boldsymbol{z}\right)=p_{\boldsymbol{Z}^{\prime} \mid \boldsymbol{Z}}^{\hat{\theta}}\left(z^{\prime} \mid \boldsymbol{z}\right)=p_{S}^{\hat{\theta}}(\boldsymbol{s}) /\left[\sum_{s_{3}, s_{4}} p_{S}^{\hat{\theta}}(\boldsymbol{s})\right]$ is calculated as

$$
\begin{align*}
T\left(\boldsymbol{z}^{\prime}, \boldsymbol{z}\right)= & \exp \left[s_{3} \hat{\theta}^{3}+s_{4} \hat{\theta}^{4}+\sum_{i<4} s_{i} s_{4} \hat{\theta}^{i 4}+\sum_{i<3} s_{i} s_{3} \hat{\theta}^{i 3}\right. \\
& +\sum_{i<j<k} s_{i} s_{j} s_{k} \hat{\theta}^{i j k}+\sum_{i<j<k<l} s_{i} s_{j} s_{k} s_{l} \hat{\theta}^{i j k l} \\
& \left.-\phi_{\mathbf{Z}^{\prime} \mid \mathbf{Z}}\left(s_{1}, s_{2} \mid \hat{\boldsymbol{\theta}}\right)\right], \tag{C9}
\end{align*}
$$

$$
\begin{align*}
\phi_{Z^{\prime} \mid \boldsymbol{Z}} & \left(s_{1}, s_{2} \mid \hat{\boldsymbol{\theta}}\right) \\
:= & \ln \left[\operatorname { e x p } \left(\hat{\theta}^{3}+\hat{\theta}^{4}+s_{1} \hat{\theta}^{14}+s_{2} \hat{\theta}^{24}+\hat{\theta}^{34}+s_{1} \hat{\theta}^{13}+s_{2} \hat{\theta}^{23}\right.\right. \\
\quad & \left.+s_{1} s_{2} \hat{\theta}^{123}+s_{1} s_{2} \hat{\theta}^{124}+s_{1} \hat{\theta}^{134}+s_{2} \hat{\theta}^{234}+s_{1} s_{2} \hat{\theta}^{1234}\right) \\
& \quad+\exp \left(\hat{\theta}^{3}+s_{1} \hat{\theta}^{13}+s_{2} \hat{\theta}^{23}+s_{1} s_{2} \hat{\theta}^{123}\right) \\
& \left.+\exp \left(\hat{\theta}^{4}+s_{1} \hat{\theta}^{14}+s_{2} \hat{\theta}^{24}+s_{1} s_{2} \hat{\theta}^{124}\right)+1\right] . \tag{C10}
\end{align*}
$$

The conditional probability $p_{\boldsymbol{Z} \mid Z^{\prime}}^{\hat{\theta}}\left(\boldsymbol{z} \mid \boldsymbol{z}^{\prime}\right)=p_{S}^{\hat{\theta}}(\boldsymbol{s}) /\left[\sum_{s_{1}, s_{2}} p_{S}^{\hat{\theta}}(\boldsymbol{s})\right]$ is also calculated as
$p_{Z \mid Z^{\prime}}^{\hat{\theta}}\left(z \mid z^{\prime}\right)$

$$
\begin{align*}
= & \exp \left[s_{1} \hat{\theta}^{1}+s_{2} \hat{\theta}^{2}+\sum_{1<i} s_{1} s_{i} \hat{\theta}^{1 i}+\sum_{2<i} s_{2} s_{i} \hat{\theta}^{2 i}\right. \\
& \left.+\sum_{i<j<k} s_{i} s_{j} s_{k} \hat{\theta}^{i j k}+\sum_{i<j<k<l} s_{i} s_{j} s_{k} s_{l} \hat{\theta}^{i j k l}-\phi_{Z \mid Z^{\prime}}\left(s_{3}, s_{4} \mid \hat{\boldsymbol{\theta}}\right)\right], \\
& \times \phi_{Z \mid Z^{\prime}}\left(s_{3}, s_{4} \mid \hat{\boldsymbol{\theta}}\right) \\
:= & \ln \left[\operatorname { e x p } \left(\hat{\theta}^{1}+\hat{\theta}^{2}+s_{3} \hat{\theta}^{23}+s_{4} \hat{\theta}^{24}+\hat{\theta}^{12}+s_{3} \hat{\theta}^{13}+s_{4} \hat{\theta}^{14}\right.\right. \\
& \left.+s_{3} s_{4} \hat{\theta}^{134}+s_{3} s_{4} \hat{\theta}^{234}+s_{3} \hat{\theta}^{123}+s_{4} \hat{\theta}^{124}+s_{3} s_{4} \hat{\theta}^{1234}\right) \\
& +\exp \left(\hat{\theta}^{1}+s_{3} \hat{\theta}^{13}+s_{4} \hat{\theta}^{14}+s_{3} s_{4} \hat{\theta}^{134}\right) \\
& \left.+\exp \left(\hat{\theta}^{2}+s_{3} \hat{\theta}^{23}+s_{4} \hat{\theta}^{24}+s_{3} s_{4} \hat{\theta}^{234}+s_{3} \hat{\theta}^{123}\right)+1\right] . \tag{C11}
\end{align*}
$$

The backward manifold is defined as

$$
\begin{equation*}
\mathcal{M}_{\mathrm{B}}=\left\{p_{S}^{\theta}(\boldsymbol{s}) \mid p_{S}^{\theta}(\boldsymbol{s})=p_{Z^{\prime}}^{\theta}\left(\boldsymbol{z}^{\prime}\right) T\left(\boldsymbol{z}, \boldsymbol{z}^{\prime}\right)\right\}, \tag{C12}
\end{equation*}
$$

where $p_{Z^{\prime}}^{\theta}\left(z^{\prime}\right)=\sum_{z} p_{S}^{\theta}(\boldsymbol{s})$. Equations (C10) and (C11) yield

$$
\begin{align*}
\mathcal{M}_{\mathrm{B}} & =\left\{p_{S}^{\hat{\theta}} \mid \theta^{1}=\hat{\theta}^{3}, \theta^{2}=\hat{\theta}^{4}, \theta^{23}=\hat{\theta}^{14}, \theta^{24}=\hat{\theta}^{24}, \theta^{12}\right. \\
& =\hat{\theta}^{34}, \\
\theta^{13} & =\hat{\theta}^{13}, \theta^{14}=\hat{\theta}^{23}, \theta^{134}=\hat{\theta}^{123}, \theta^{234}=\hat{\theta}^{124}, \\
\theta^{123} & \left.=\hat{\theta}^{134}, \theta^{124}=\hat{\theta}^{234}, \theta^{1234}=\hat{\theta}^{1234}\right\} . \tag{C13}
\end{align*}
$$

Under both bipartite conditions $\mathcal{C}_{\mathrm{BI}}$ and $\mathcal{C}_{\mathrm{BI}}^{*}$, the joint probability is given by

$$
\begin{equation*}
p^{\mathrm{BI}}{ }_{S}^{\hat{\theta}}=\left.p_{S}^{\hat{\theta}}\right|_{\hat{\theta}^{34}=\hat{\theta}^{134}=\hat{\theta}^{234}=\hat{\theta}^{12}=\hat{\theta}^{123}=\hat{\theta}^{124}=\hat{\theta}^{1234}=0 .} . \tag{C14}
\end{equation*}
$$

For this distribution $p_{S}^{\mathrm{BI} \hat{\theta}}$, the condition of the backward manifold is given by

$$
\begin{align*}
\mathcal{M}_{\mathrm{B}} & =\left\{p^{\mathrm{BI} \theta}{ }_{S}^{\theta} \mid \theta^{1}=\hat{\theta}^{3}, \theta^{2}=\hat{\theta}^{4}, \theta^{23}=\hat{\theta}^{14}, \theta^{24}=\hat{\theta}^{24},\right. \\
\theta^{13} & \left.=\hat{\theta}^{13}, \theta^{14}=\hat{\theta}^{23}\right\} . \tag{C15}
\end{align*}
$$

Next, we discuss the local backward manifold $\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}$. Then the transition probability $T^{\mathcal{X}}\left(z^{\prime}, z\right)=p_{X^{\prime} \mid Z, Y^{\prime}}^{\theta}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{z}\right)$ is given by Eq. (C5). The conditional probability $p_{X \mid Z^{\prime}, Y}^{\hat{\theta}}\left(\boldsymbol{x} \mid \boldsymbol{z}^{\prime}, \boldsymbol{y}\right)=$ $p_{S}^{\hat{\theta}}(\boldsymbol{s}) /\left[\sum_{s_{1}} p_{S}^{\hat{\theta}}(\boldsymbol{s})\right]$ is calculated as

$$
\begin{align*}
p_{\boldsymbol{X} \mid \mathbf{Z}^{\prime}, \boldsymbol{Y}}^{\hat{\theta}} & \left(\boldsymbol{x} \mid z^{\prime}, \boldsymbol{y}\right) \\
\quad= & \exp \left[s_{1} \hat{\theta}^{1}+s_{1} s_{2} \hat{\theta}^{12}+s_{1} s_{3} \hat{\theta}^{13}+s_{1} s_{4} \hat{\theta}^{14}\right. \\
& +s_{1} s_{2} s_{3} \hat{\theta}^{123}+s_{1} s_{2} s_{4} \hat{\theta}^{124}+s_{1} s_{3} s_{4} \hat{\theta}^{134} \\
& \left.+s_{1} s_{2} s_{3} s_{4} \hat{\theta}^{1234}-\phi_{\boldsymbol{X} \mid \mathbf{Z}^{\prime}, \boldsymbol{Y}}\left(s_{2}, s_{3}, s_{4} \mid \hat{\boldsymbol{\theta}}\right)\right] \tag{C16}
\end{align*}
$$

$$
\begin{align*}
& \phi_{X \mid Z^{\prime}, \boldsymbol{Y}}\left(s_{2}, s_{3}, s_{4} \mid \hat{\boldsymbol{\theta}}\right) \\
& \quad:=\ln \left[\operatorname { e x p } \left(\hat{\theta}^{1}+s_{2} \hat{\theta}^{12}+s_{3} \hat{\theta}^{13}+s_{4} \hat{\theta}^{14}+s_{2} s_{3} \hat{\theta}^{123}\right.\right. \\
& \left.\left.\quad+s_{2} s_{4} \hat{\theta}^{124}+s_{3} s_{4} \hat{\theta}^{134}+s_{2} s_{3} s_{4} \hat{\theta}^{1234}\right)+1\right] . \tag{C17}
\end{align*}
$$

The local backward manifold is defined as

$$
\begin{equation*}
\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}=\left\{p_{S}^{\theta} \mid p_{S}^{\theta}(\boldsymbol{s})=p_{Z^{\prime}, \boldsymbol{Y}}^{\theta}\left(z^{\prime}, \boldsymbol{y}\right) T^{\mathcal{X}}\left(\boldsymbol{z}, \boldsymbol{z}^{\prime}\right)\right\} \tag{C18}
\end{equation*}
$$

where $p_{Z^{\prime}, \boldsymbol{Y}}^{\theta}\left(z^{\prime}, \boldsymbol{y}\right)=\sum_{s_{1}} p_{S}^{\theta}(\boldsymbol{s})$. Equations (C5) and (C17) yield

$$
\begin{align*}
\mathcal{M}_{\mathrm{LB}}^{\mathcal{X}} & =\left\{p_{S}^{\theta} \mid \theta^{1}=\hat{\theta}^{3}, \theta^{12}=\hat{\theta}^{34}, \theta^{13}=\hat{\theta}^{13}, \theta^{14}=\hat{\theta}^{23},\right. \\
\theta^{123} & =\hat{\theta}^{134}, \theta^{124}=\hat{\theta}^{234}, \theta^{134}=\hat{\theta}^{123}, \\
\theta^{234} & \left.=\hat{\theta}^{124}, \theta^{1234}=\hat{\theta}^{1234}\right\} . \tag{C19}
\end{align*}
$$

In the same way, we obtain the condition of $\mathcal{M}_{\mathrm{LB}}^{\mathcal{Y}}$,

$$
\begin{align*}
\mathcal{M}_{\mathrm{LB}}^{\mathcal{Y}} & =\left\{p_{S}^{\theta} \mid \theta^{2}=\hat{\theta}^{4}, \theta^{12}=\hat{\theta}^{34}, \theta^{23}=\hat{\theta}^{14},\right. \\
\theta^{24} & =\hat{\theta}^{24}, \theta^{123}=\hat{\theta}^{134}, \theta^{124}=\hat{\theta}^{234} \\
\theta^{134} & \left.=\hat{\theta}^{123}, \theta^{234}=\hat{\theta}^{124}, \theta^{1234}=\hat{\theta}^{1234}\right\} \tag{C20}
\end{align*}
$$

To clarify the relation between $\mathcal{M}_{\mathrm{LR}}^{\mathcal{X}}$ and $\mathcal{M}_{\mathrm{LR}}^{\mathcal{Y}}$, we can consider the permutation $\left(\alpha^{\prime}(1), \alpha^{\prime}(2), \alpha^{\prime}(3), \alpha^{\prime}(4)\right)=$ $(2,1,4,3)$. The condition of $\mathcal{M}_{\mathrm{LR}}^{\mathcal{Y}}$ is given by the condition of $\mathcal{M}_{\mathrm{LR}}^{\mathcal{X}}$ with the permutation $\alpha^{\prime}$,

$$
\begin{align*}
\mathcal{M}_{\mathrm{LB}}^{\mathcal{Y}} & =\left\{p_{S}^{\theta} \mid \theta^{\alpha^{\prime}(1)}=\hat{\theta}^{\alpha^{\prime}(3)}, \theta^{\alpha^{\prime}(1) \alpha^{\prime}(2)}=\hat{\theta}^{\alpha^{\prime}(3) \alpha^{\prime}(4)},\right. \\
\theta^{\alpha^{\prime}(1) \alpha^{\prime}(3)} & =\hat{\theta}^{\alpha^{\prime}(1) \alpha^{\prime}(3)}, \theta^{\alpha^{\prime}(1) \alpha^{\prime}(4)}=\hat{\theta}^{\alpha^{\prime}(2) \alpha^{\prime}(3)}, \\
\theta^{\alpha^{\prime}(1) \alpha^{\prime}(2) \alpha^{\prime}(3)} & =\hat{\theta}^{\alpha^{\prime}(1) \alpha^{\prime}(3) \alpha^{\prime}(4)}, \\
\theta^{\alpha^{\prime}(1) \alpha^{\prime}(2) \alpha^{\prime}(4)} & =\hat{\theta}^{\alpha^{\prime}(2) \alpha^{\prime}(3) \alpha^{\prime}(4)}, \\
\theta^{\alpha^{\prime}(1) \alpha^{\prime}(3) \alpha^{\prime}(4)} & =\hat{\theta}^{\alpha^{\prime}(1) \alpha^{\prime}(2) \alpha^{\prime}(3)}, \\
\theta^{\alpha^{\prime}(2) \alpha^{\prime}(3) \alpha^{\prime}(4)} & =\hat{\theta}^{\alpha^{\prime}(1) \alpha^{\prime}(2) \alpha^{\prime}(4)}, \\
\theta^{\alpha^{\prime}(1) \alpha^{\prime}(2) \alpha^{\prime}(3) \alpha^{\prime}(4)} & \left.=\hat{\theta}^{\alpha^{\prime}(1) \alpha^{\prime}(2) \alpha^{\prime}(3) \alpha^{\prime}(4)}\right\} . \tag{C21}
\end{align*}
$$

For this distribution $p^{\mathrm{BI}{ }_{S}^{\hat{\theta}}}$ under both bipartite conditions, the local backward manifolds are given by

$$
\begin{align*}
& \mathcal{M}_{\mathrm{LB}}^{\mathcal{X}}=\left\{p^{\mathrm{BI}}{ }_{S}^{\theta} \mid \theta^{1}=\hat{\theta}^{3}, \theta^{13}=\hat{\theta}^{13}, \theta^{14}=\hat{\theta}^{23}\right\},  \tag{C22}\\
& \mathcal{M}_{\mathrm{LB}}^{\mathcal{Y}}=\left\{p^{\mathrm{BII}}{ }_{S}^{\theta} \mid \theta^{2}=\hat{\theta}^{4}, \theta^{24}=\hat{\theta}^{24}, \theta^{23}=\hat{\theta}^{14}\right\} . \tag{C23}
\end{align*}
$$

## APPENDIX D: THE CASE OF FEEDBACK CONTROL

We consider the situation where the time evolution of system $\mathcal{X}$ depends on the fixed memory $\mathcal{M}$. This situation is well known as the problem of Maxwell's demon under feedback control. We show that the partial entropy production for this case can also be discussed in our unified framework.

Let $\boldsymbol{X}$ and $\boldsymbol{X}^{\prime}$ be random variables of system $\mathcal{X}$ at times $t$ and $t+d t$, respectively. Let $\boldsymbol{M}$ be a random variable of the memory $\mathcal{M}$. We denote the set of random variables as $\boldsymbol{S}=$ $\left\{\boldsymbol{X}, \boldsymbol{X}^{\prime}, \boldsymbol{M}\right\}$ and the set of states as $\boldsymbol{s}=\left\{\boldsymbol{x}, \boldsymbol{x}^{\prime}, \boldsymbol{m}\right\}$. The joint probability of $\boldsymbol{S}$ is given by $p_{\boldsymbol{S}}(\boldsymbol{s})$. We consider the situation
where the transition probability of $\mathcal{X}$ depends on the state of memory,

$$
\begin{equation*}
p_{X^{\prime} \mid X, M}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{x}, \boldsymbol{m}\right)=: T^{\mathcal{X} \mathcal{M}}\left(\boldsymbol{x}^{\prime}, \boldsymbol{m}, \boldsymbol{x}\right), \tag{D1}
\end{equation*}
$$

where $p_{X^{\prime} \mid \boldsymbol{X}, \boldsymbol{M}}\left(\boldsymbol{x}^{\prime} \mid \boldsymbol{x}, \boldsymbol{m}\right)=p_{\boldsymbol{S}}(\boldsymbol{s}) /\left[\sum_{\boldsymbol{x}^{\prime}} p_{S}(\boldsymbol{s})\right]$. We here introduce the feedback backward manifold such that

$$
\begin{equation*}
\mathcal{M}_{\mathrm{FB}}=\left\{q_{S} \mid q_{S}(\boldsymbol{s})=T^{\mathcal{X} \mathcal{M}}\left(\boldsymbol{x}, \boldsymbol{m}, \boldsymbol{x}^{\prime}\right) q_{X^{\prime} \boldsymbol{M}}\left(\boldsymbol{x}^{\prime}, \boldsymbol{m}\right)\right\}, \tag{D2}
\end{equation*}
$$

where $q_{X^{\prime} M}\left(\boldsymbol{x}^{\prime}, \boldsymbol{m}\right)=\sum_{x} q_{S}(\boldsymbol{s})$. The feedback reversible manifold is equivalent to the reversible manifold $\mathcal{M}_{\mathrm{B}}=\mathcal{M}_{\mathrm{FB}}$ if we consider the time evolution from $\boldsymbol{Z}=\{\boldsymbol{X}, \boldsymbol{M}\}$ to $\boldsymbol{Z}^{\prime}=$ $\left\{\boldsymbol{X}^{\prime}, \boldsymbol{M}\right\}$. If the joint probability $q_{S}$ is on this manifold $\mathcal{M}_{\mathrm{FB}}$, the dynamics of $\mathcal{X}$ are reversible in time under feedback control. If we introduce the joint probability $q_{S}^{\mathcal{X} \mathcal{M} *}(\boldsymbol{s})=$ $T^{\mathcal{X} \mathcal{M}}\left(\boldsymbol{x}, \boldsymbol{m}, \boldsymbol{x}^{\prime}\right) p_{\boldsymbol{X}^{\prime}, \boldsymbol{M}}\left(\boldsymbol{x}^{\prime}, \boldsymbol{m}\right)$, the following Pythagorean theorem is valid for any $q_{S} \in \mathcal{M}_{\mathrm{FR}}$ :

$$
\begin{equation*}
D\left(p_{S} \| q_{S}\right)=D\left(p_{S} \| q_{S}^{\mathcal{X} \mathcal{M} *}\right)+D\left(q_{S}^{\mathcal{X} \mathcal{M} *} \| q_{S}\right) \tag{D3}
\end{equation*}
$$

Thus, the feedback backward manifold is flat, and the solution of the optimization problem $D^{\text {opt }}\left(p_{S} \| \mathcal{M}_{\mathrm{FB}}\right)$ is given by

$$
\begin{align*}
D^{\mathrm{opt}}\left(p_{S} \| \mathcal{M}_{\mathrm{FB}}\right) & :=\min _{q_{S} \in \mathcal{M}_{\mathrm{FB}}} D\left(p_{S} \| q_{S}\right)  \tag{D4}\\
& =D\left(p_{S} \| q_{S}^{\mathcal{X} \mathcal{M}^{*}}\right) \tag{D5}
\end{align*}
$$

We here derive the result that the partial entropy production under feedback control $\sigma_{\text {feedback }}^{\mathcal{X}}$ is given by the optimization problem

$$
\begin{equation*}
\sigma_{\text {feedback }}^{\mathcal{X}}=D^{\mathrm{opt}}\left(p_{S} \| \mathcal{M}_{\mathrm{FB}}\right) \tag{D6}
\end{equation*}
$$

The partial entropy production under feedback control $\sigma_{\text {feedback }}^{\mathcal{X}}$ is defined as

$$
\begin{gather*}
\sigma_{\text {feedback }}^{\mathcal{X}}:=\sigma_{\text {sys }}^{\mathcal{X}}+\sigma_{\text {bath }}^{\mathcal{X}}-\Delta \mathcal{I},  \tag{D7}\\
\sigma_{\text {sys }}^{\mathcal{X}}:=H\left(\boldsymbol{X}^{\prime}\right)-H(\boldsymbol{X}),  \tag{D8}\\
\sigma_{\text {bath }}^{\mathcal{X}}:=\mathbb{E}\left[\ln \frac{T^{\mathcal{X} \mathcal{M}}\left(\boldsymbol{x}^{\prime}, \boldsymbol{m}, \boldsymbol{x}\right)}{T^{\mathcal{X} \mathcal{M}}\left(\boldsymbol{x}, \boldsymbol{m}, \boldsymbol{x}^{\prime}\right)}\right],  \tag{D9}\\
\Delta \mathcal{I}:=I\left(\boldsymbol{X}^{\prime} ; \boldsymbol{M}\right)-I(\boldsymbol{X} ; \boldsymbol{M}), \tag{D10}
\end{gather*}
$$

where $\sigma_{\text {sys }}^{\mathcal{X}}$ is the entropy change of system $\mathcal{X}, \sigma_{\text {bath }}^{\mathcal{X}}$ is the entropy change of the heat bath attached to system $\mathcal{X}$, and $\Delta \mathcal{I}$ is the mutual information change between system $\mathcal{X}$ and the memory $\mathcal{M}$. To show the relation

$$
\begin{equation*}
\sigma_{\text {feedback }}^{\mathcal{X}}=D\left(p_{S} \| q_{S}^{\mathcal{X}} \mathcal{M} *\right) \tag{D11}
\end{equation*}
$$

we obtain the result (D6). The second law of information thermodynamics under feedback control is given by the nonnegativity of $\sigma_{\text {feedback }}^{\mathcal{X}}$,

$$
\begin{equation*}
\sigma_{\text {sys }}^{\mathcal{X}}+\sigma_{\text {bath }}^{\mathcal{X}} \geqslant \Delta \mathcal{I} . \tag{D12}
\end{equation*}
$$

This inequality implies a trade-off relation between the entropy changes in system $\mathcal{X}$ and the information between system $\mathcal{X}$ and the memory $\mathcal{M}$.
[1] S. I. Amari and H. Nagaoka, Methods of Information Geometry (American Mathematical Society, 2007).
[2] S. I. Amari, Information Geometry and Its Applications (Springer, Japan, 2016).
[3] S. I. Amari, K. Kurata, and H. Nagaoka, Information geometry of Boltzmann machines, IEEE Trans. Neural Networks 3, 260 (1992).
[4] S. I. Amari, Information geometry of the EM and em algorithms for neural networks, Neural Networks 8, 1379 (1995).
[5] T. Tanaka, Information geometry of mean-field approximation, Neural Comput. 12, 1951 (2000).
[6] D. C. Brody and D. W. Hook, Information geometry in vapourliquid equilibrium, J. Phys. A 42, 023001 (2008).
[7] J. Uffink and J. van Lith, Thermodynamic uncertainty relations, Found. Phys. 29, 655 (1999).
[8] G. E. Crooks, Measuring Thermodynamic Length, Phys. Rev. Lett. 99, 100602 (2007).
[9] S. Ito, Stochastic Thermodynamic Interpretation of Information Geometry, Phys. Rev. Lett. 121, 030605 (2018).
[10] S. Ito and A. Dechant, Stochastic Time Evolution, Information Geometry, and the Cramér-Rao Bound, Phys. Rev. X 10, 021056 (2020).
[11] F. Weinhold, Metric geometry of equilibrium thermodynamics, J. Chem. Phys. 63, 2479 (1975).
[12] G. Ruppeiner, Thermodynamics: A Riemannian geometric model, Phys. Rev. A 20, 1608 (1979).
[13] P. Salamon and R. S. Berry, Thermodynamic Length and Dissipated Availability, Phys. Rev. Lett. 51, 1127 (1983).
[14] F. H. Edward and G. E. Crooks, Length of Time's Arrow, Phys. Rev. Lett. 101, 090602 (2008).
[15] D. A. Sivak and G. E. Crooks, Thermodynamic Metrics and Optimal Paths, Phys. Rev. Lett. 108, 190602 (2012).
[16] M. Polettini and M. Esposito, Nonconvexity of the relative entropy for Markov dynamics: A Fisher information approach, Phys. Rev. E 88, 012112 (2013).
[17] B. B. Machta, Dissipation Bound for Thermodynamic Control, Phys. Rev. Lett. 115, 260603 (2015).
[18] S. Lahiri, J. Sohl-Dickstein, and S. Ganguli, A universal tradeoff between power, precision and speed in physical communication, arXiv:1603.07758.
[19] N. Shiraishi and H. Tajima, Efficiency versus speed in quantum heat engines: Rigorous constraint from Lieb-Robinson bound, Phys. Rev. E 96, 022138 (2017).
[20] G. M. Rotskoff, G. E. Crooks, and E. Vanden-Eijnden, Geometric approach to optimal nonequilibrium control: Minimizing dissipation in nanomagnetic spin systems, Phys. Rev. E 95, 012148 (2017).
[21] K. Takahashi, Shortcuts to adiabaticity applied to nonequilibrium entropy production: An information geometry viewpoint, New J. Phys. 19, 115007 (2017).
[22] H. Shimazaki, Neural Engine Hypothesis, in Dynamic Neuroscience (Springer, Cham, 2018), pp. 267-291.
[23] T. M. Cover and J. A. Thomas, Elements of Information Theory (Wiley, Hoboken, NJ, 2012).
[24] S. I. Amari, Information geometry on hierarchy of probability distributions, IEEE Trans. Inf. Theory 47, 1701 (2001).
[25] M. Oizumi, N. Tsuchiya, and S. I. Amari, Unified framework for information integration based on information geometry, Proc. Natl. Acad. Sci. U.S.A. 113, 14817 (2016).
[26] S. I. Amari, N. Tsuchiya, and M. Oizumi, Geometry of information integration, in Information Geometry and its Applications IV (Springer, Cham, 2016), pp. 3-17.
[27] G. Tononi, An information integration theory of consciousness, BMC Neurosci. 5, 42 (2004).
[28] D. Balduzzi and G. Tononi, Integrated information in discrete dynamical systems: Motivation and theoretical framework, PLoS Comput. Biol. 4, e1000091 (2008).
[29] A. B. Barrett and A. K. Seth, Practical measures of integrated information for time-series data, PLoS Comput. Biol. 7, e1001052 (2011).
[30] M. Oizumi, L. Albantakis, and G. Tononi, From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0, PLoS Comput. Biol. 10, e1003588 (2014).
[31] N. Ay, Information geometry on complexity and stochastic interaction, Entropy 17, 2432 (2015).
[32] G. Tononi, M. Boly, M. Massimini, and C. Koch, Integrated information theory: From consciousness to its physical substrate, Nat. Rev. Neurosci. 17, 450 (2016).
[33] M. Oizumi, S. I. Amari, T. Yanagawa, N. Fujii, and N. Tsuchiya, Measuring integrated information from the decoding perspective, PLoS Comput. Biol. 12, e1004654 (2016).
[34] M. Tegmark, Improved measures of integrated information, PLoS Comput. Biol. 12, e1005123 (2016).
[35] P. A. Mediano, F. Rosas, R. L. Carhart-Harris, A. K. Seth, and A. B. Barrett, Beyond integrated information: A taxonomy of information dynamics phenomena, arXiv:1909.02297.
[36] K. Sekimoto, Stochastic Energetics (Springer, 2010).
[37] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. 75, 126001 (2012).
[38] J. M. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, Nat. Phys. 11, 131 (2015).
[39] T. Sagawa and M. Ueda, Generalized Jarzynski Equality under Nonequilibrium Feedback Control, Phys. Rev. Lett. 104, 090602 (2010).
[40] S. Still, D. A. Sivak, A. J. Bell, and G. E. Crooks, Thermodynamics of Prediction, Phys. Rev. Lett. 109, 120604 (2012).
[41] T. Sagawa and M. Ueda, Fluctuation Theorem with Information Exchange: Role of Correlations in Stochastic Thermodynamics, Phys. Rev. Lett. 109, 180602 (2012).
[42] S. Ito and T. Sagawa, Information Thermodynamics on Causal Networks, Phys. Rev. Lett. 111, 180603 (2013).
[43] D. Hartich, A. C. Barato, and U. Seifert, Stochastic thermodynamics of bipartite systems: Transfer entropy inequalities and a Maxwell's demon interpretation, J. Stat. Mech. (2014) P02016.
[44] J. M. Horowitz and M. Esposito, Thermodynamics with Continuous Information Flow, Phys. Rev. X 4, 031015 (2014).
[45] S. Ito and T. Sagawa, Maxwell's demon in biochemical signal transduction with feedback loop, Nat. Commun. 6, 7498 (2015).
[46] R. E. Spinney, J. T. Lizier, and M. Prokopenko, Transfer entropy in physical systems and the arrow of time, Phys. Rev. E 94, 022135 (2016).
[47] S. Ito, Backward transfer entropy: Informational measure for detecting hidden Markov models and its interpretations in thermodynamics, gambling and causality, Sci. Rep. 6, 36831 (2016).
[48] G. E. Crooks and S. E. Still, Marginal and conditional second laws of thermodynamics, Europhys. Lett. 125, 40005 (2019).
[49] A. Auconi, A. Giansanti, and E. Klipp, Information Thermodynamics for Time Series of Signal-Response Models, Entropy 21, 177 (2019).
[50] S. Ito and T. Sagawa, Information flow and entropy production on Bayesian networks, in Mathematical Foundations and Applications of Graph Entropy, edited by M. Dehmer, F. Emmert-Streib, Z. Chen, X. Li, and Y. Shi (Wiley-VCH, 2016), Chap. 3.
[51] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, Dissipation: The Phase-Space Perspective, Phys. Rev. Lett. 98, 080602 (2007).
[52] T. Schreiber, Measuring Information Transfer, Phys. Rev. Lett. 85, 461 (2000).
[53] L. Barnett and A. K. Seth, The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference, J. Neurosci. Methods 223, 50 (2014).
[54] A. C. Barato, D. Hartich, and U. Seifert, Efficiency of cellular information processing, New J. Phys. 16, 103024 (2014).
[55] P. Sartori, L. Granger, C. F. Lee, and J. M. Horowitz, Thermodynamic costs of information processing in sensory adaptation, PLoS Comput. Biol. 10, e1003974 (2014).
[56] S. Bo, M. Del Giudice, and A. Celani, Thermodynamic limits to information harvesting by sensory systems, J. Stat. Mech. (2015) P01014.
[57] T. E. Ouldridge, C. C. Govern, and P. R. ten Wolde, Thermodynamics of Computational Copying in Biochemical Systems, Phys. Rev. X 7, 021004 (2017).
[58] T. McGrath, N. S. Jones, P. R. ten Wolde, and T. E. Ouldridge, Biochemical Machines for the Interconversion of Mutual Information and Work, Phys. Rev. Lett. 118, 028101 (2017).
[59] T. Matsumoto and T. Sagawa, Role of sufficient statistics in stochastic thermodynamics and its implication to sensory adaptation, Phys. Rev. E 97, 042103 (2018).
[60] C. Carathéodory, Untersuchungen über die Grundlagen der Thermodynamik, Math. Ann. 67, 355 (1909).
[61] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. 106, 620 (1957).
[62] E. T. Jaynes, Information theory and statistical mechanics. II, Phys. Rev. 108, 171 (1957).


[^0]:    Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

