

Article

Not peer-reviewed version

---

# Information Exchange Fluctuation Theorem Under Coarse-Graining

---

[Lee Jinwoo](#)\*

Posted Date: 21 July 2025

doi: 10.20944/preprints202507.1663.v1

Keywords: local non-equilibrium thermodynamics, fluctuation theorem, mutual information, entropy production, pointwise mutual information, thermodynamics of information, stochastic thermodynamics



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

*Article*

# Information Exchange Fluctuation Theorem Under Coarse-Graining

Lee Jinwoo 

Department of Mathematics, Kwangwoon University, 20 Kwangwoon-ro, Seoul 01897, Korea; jinwoolee@kw.ac.kr

## Abstract

The fluctuation theorem for information exchange, originally established by Sagawa and Ueda [Phys. Rev. Lett. 109, 180602 (2012)], provides a fundamental framework for understanding the role of correlations in coupled classical stochastic systems. Building upon this foundation, Jinwoo [Entropy 21, 477 (2019)] demonstrated that the pointwise mutual information between correlated subsystems captures entropy production as a state function during coupling processes. In this study, we investigate the robustness of this information-theoretic fluctuation theorem under coarse-graining in coupled classical fluctuating systems. We rigorously prove that the fluctuation theorem remains invariant under arbitrary coarse-graining transformations and derive hierarchical relationships between information measures across different scales, thereby establishing its fundamental character as independent of the level of system description. Our results demonstrate that the relationship between information exchange and entropy production is preserved across different scales of observation, providing deeper insights into the thermodynamic foundations of information processing in classical stochastic systems.

**Keywords:** local non-equilibrium thermodynamics, fluctuation theorem, mutual information, entropy production, pointwise mutual information, thermodynamics of information, stochastic thermodynamics

## 1. Introduction

Stochastic thermal fluctuations constitute a cornerstone in the operation of molecular machinery and systems operating far from thermodynamic equilibrium. These fluctuations facilitate energy transfer between molecular components and their surroundings, permitting molecules to surmount energetic barriers and achieve stable low-energy configurations. The inherent randomness of molecular motion transforms thermodynamic quantities such as heat and work into stochastic variables. During the past twenty years, fluctuation theorems have emerged as a powerful class of exact relations revealing universal principles that constrain these random thermodynamic quantities in processes driving systems away from equilibrium.

The theoretical foundation was established through seminal contributions including the Evans-Searles transient fluctuation theorem (1994), followed by the groundbreaking Jarzynski equality [1] (1997) and the Crooks fluctuation theorem [2] (1998), which spawned numerous theoretical extensions and generalizations. Seifert's formulation extended thermodynamic laws to individual stochastic trajectories [3], while Hatano and Sasa developed frameworks for transitions between non-equilibrium steady states [4]. Jinwoo and Tanaka [5,6] demonstrated that fluctuation theorems hold even within an ensemble of trajectories conditioned at a final (micro/meso) state, establishing that local free energy captures fluctuating work within the path ensemble reaching a specific (micro/meso) state during an externally driven non-equilibrium process. Single-molecule experimental studies have provided rigorous validation of these theoretical predictions, yielding valuable insights into biomolecular dynamics [7–12].

Contemporary research in fluctuation theorems for information exchange within coupled classical stochastic systems has established a comprehensive theoretical framework linking information

theory with non-equilibrium statistical mechanics [13–17]. This rapidly advancing field, built upon the pioneering contributions of Sagawa and Ueda, has witnessed substantial theoretical developments and experimental validations in recent years [18,19]. The discipline investigates fundamental principles governing the energetic requirements of information manipulation, the influence of statistical correlations on entropy generation, and the persistence of thermodynamic laws across different observational scales.

This theoretical framework has become particularly relevant for understanding biological systems, as living organisms have evolved sophisticated information processing architectures essential for their survival and reproduction [20–22]. These biological systems demonstrate remarkable capabilities in detecting environmental chemical signals [23,24], propagating information across complex signaling pathways [25–27], and regulating genetic expression through molecular communication networks [28,29]. Cellular systems have developed the ability to perform temporal integration by encoding environmental conditions into internal molecular configurations, thereby minimizing sensory uncertainties [30,31]. Understanding the thermodynamic implications of information processing is therefore essential for comprehending the intricate mechanisms underlying biological computation and how fluctuation theorems govern the fundamental limits of these processes.

Sagawa and Ueda incorporated informational considerations into fluctuation theorem formalism [15]. Their work established a fluctuation theorem for information exchange processes, creating a unified description of measurement and feedback control phenomena in non-equilibrium settings [17]. Their theoretical framework analyzed scenarios where system  $X$  undergoes evolution dependent on the instantaneous state  $y$  of system  $Y$ , under the assumption that  $Y$  remains stationary during  $X$ 's temporal evolution. They rigorously demonstrated that correlation establishment between subsystems inevitably generates entropy production, providing definitive resolution to Maxwell's demon paradox and establishing mathematical foundations for understanding energy-information conversion mechanisms.

Subsequent theoretical advances have extended these foundational results. Particularly significant is Jinwoo's work that eliminated the static constraint imposed by Sagawa and Ueda, demonstrating that identical fluctuation theorem forms remain valid when both subsystems  $X$  and  $Y$  undergo simultaneous dynamic evolution [32]. Jinwoo also demonstrated that pointwise mutual information between a coupled state captures the entropy production within the ensemble of paths that reach the coupled state [33].

The present investigation examines the stability of information-theoretic fluctuation theorems under coarse-graining in coupled classical stochastic systems. Establishing how these fundamental relationships maintain their validity across multiple observational scales is critical for demonstrating the universal nature of information-thermodynamic principles and their applicability to practical systems where complete microscopic details may be experimentally inaccessible or computationally intractable.

In this study, we investigate the robustness of information-theoretic fluctuation theorems under coarse-graining in coupled classical stochastic systems. Building upon the foundational framework established by Sagawa and Ueda [17] and the subsequent generalization by Jinwoo [32,33], our research addresses a fundamental question: do the relationships between information exchange and entropy production remain invariant when the system description is coarse-grained to different levels of detail?

The motivation for this investigation stems from the practical reality that physical systems are typically observed and described at various scales, from microscopic molecular dynamics to mesoscopic collective behaviors. Understanding whether fluctuation theorems maintain their validity across these different levels of description has profound implications for both theoretical understanding and practical applications. If these theorems are indeed robust under coarse-graining, it would establish their fundamental character as scale-independent principles governing information processing in thermodynamic systems.

Our approach rigorously proves that the fluctuation theorem for information exchange remains invariant under coarse-graining operations, thereby establishing its fundamental character as inde-

pendent of the level of system description. The results demonstrate that the relationship between information exchange and entropy production is preserved across different scales of observation, providing deeper insights into the thermodynamic foundations of information processing in classical stochastic systems.

## 2. Information Exchange Fluctuation Theorem

We examine a system weakly coupled to a thermal reservoir characterized by inverse temperature  $\beta := 1/(k_B T)$ , where  $k_B$  denotes the Boltzmann constant and  $T$  represents the reservoir temperature. The external protocol  $\mu_t$  perturbs the system from its equilibrium state throughout the time interval  $0 \leq t \leq \tau$ . We denote by  $\mathcal{T}$  the complete ensemble of microscopic trajectories, while  $\mathcal{T}_{x_\tau}$  represents the subset of trajectories that terminate at the specific microstate  $x_\tau$  at the final time  $\tau$ .

We now examine the Sagawa-Ueda fluctuation theorem for information exchange [17], focusing particularly on its generalized formulation [32]. For this analysis, we study two subsystems  $X$  and  $Y$  immersed in a thermal reservoir characterized by inverse temperature  $\beta$ . Throughout the protocol  $\mu_t$ , these subsystems undergo mutual interaction and coupled evolution. Under these conditions, the information exchange fluctuation theorem takes the form:

$$\left\langle e^{-\Pi + \Delta M} \right\rangle_{\mathcal{T}} = 1, \quad (1)$$

where the angular brackets denote ensemble averaging across all trajectories  $\mathcal{T}$  of the coupled system,  $\Pi$  represents the total entropy production encompassing contributions from system  $X$ , system  $Y$ , and the thermal bath, while  $\Delta M$  quantifies the variation in mutual information linking  $X$  and  $Y$ . It is worth noting that the original Sagawa-Ueda formulation restricts the thermal contact to system  $X$  alone, with system  $Y$  remaining stationary throughout the evolution [17,32]. A trajectory-conditioned variant of Equation (1) reads:

$$I_\tau(x_\tau, y_\tau) = -\ln \left\langle e^{-(\Pi + I_0)} \right\rangle_{x_\tau, y_\tau}, \quad (2)$$

where the averaging is performed over the restricted ensemble of trajectories terminating at the coupled state  $(x_\tau, y_\tau)$  at the final time  $\tau$ , and  $I_t$  (for  $0 \leq t \leq \tau$ ) denotes the pointwise mutual information between the microstates of systems  $X$  and  $Y$  at time  $t$  [33]. When initial correlations are absent, i.e.,  $I_0 = 0$ , relation Equation (2) demonstrates that the pointwise mutual information  $I_\tau$ , viewed as a state function of the coupled microstates  $(x_\tau, y_\tau)$ , captures the entropy production  $\Pi$  within the trajectory ensemble constrained to reach these final states. By analogy, the initial correlation  $I_0$  can be understood as representing the entropy production associated with establishing the initial correlated state.

In this study, we consider coarse-graining transformations  $a$  and  $b$ , which are functions of microstates  $x$  and  $y$ , respectively. The functions  $a$  and  $b$  can be arbitrary generalized coordinates, for example,

(i) *Spatial binning*:  $a(x) = \text{bin}(x)$ , where molecular positions  $x$  are discretized into spatial bins of a certain width, commonly used in analyzing diffusion processes and spatial correlations in biomolecular systems;

(ii) *Energy-based grouping*:  $a(x) = E(x)$  where microstates  $x$  are classified according to their energy levels  $E$ , particularly relevant for studying conformational transitions in proteins and folding dynamics;

(iii) *Reaction coordinate projection*:  $a(x) = f(x)$  where  $f$  represents a collective variable such as end-to-end distance in polymers, radius of gyration in protein folding, or dihedral angles in molecular conformational changes;

(iv) *Cluster-based coarse-graining*:  $a(x) = c$  where  $c$  denotes the cluster index obtained from unsupervised clustering algorithms applied to molecular configurations, widely employed in Markov state model construction for biomolecular dynamics;

(v) *Order parameter discretization*:  $a(x) = \text{sign}(\psi(x))$  where  $\psi(x)$  is an order parameter such as local crystallinity in phase transitions or helical content in protein secondary structure formation.

These transformations effectively reduce the dimensionality of the state space while preserving essential thermodynamic information. We establish that such transformations maintain the information exchange thermodynamic behavior across different scales of observation by proving that

$$I_\tau(a, b) = -\ln \left\langle e^{-(\Pi + I_0)} \right\rangle_{a,b} \quad (3)$$

where the angular brackets denote the ensemble average over all conditioned trajectories that terminate at coupled microstates  $(x_\tau, y_\tau)$  satisfying  $a(x_\tau) = a$  and  $b(y_\tau) = b$ . In the course of establishing Equation (3), we simultaneously derive

$$\left\langle e^{-I_\tau(x_\tau, y_\tau)} \right\rangle_{a(x_\tau)=a, b(y_\tau)=b} = e^{-I_\tau(a,b)}, \quad (4)$$

where the angular brackets represent the ensemble average taken over the conditional probability distribution  $P_\tau(x_\tau, y_\tau | a, b)$ . Equation (4) constitutes a fluctuation relation that reveals the hierarchical organization of pointwise mutual information across different observational scales.

Moreover, we prove the work fluctuation theorem for information exchange as follows:

$$\left\langle e^{-\beta W} \right\rangle_{a,b} = e^{-\beta \Delta \Phi_X(a, \tau) - \beta \Delta \Phi_Y(b, \tau) - I_\tau(a, b)}, \quad (5)$$

where  $\Delta \Phi_X(a, \tau) = \Phi_X(a, \tau) - F_{\text{eq}}(\mu_0)$  is the difference between the non-equilibrium free energy  $\Phi_X(a, \tau)$  and the equilibrium free energy  $F_{\text{eq}}(\mu_0)$  for system  $X$ , and  $\Delta \Phi_Y$  is the corresponding quantity for system  $Y$ . Equation (5) holds provided that the initial condition corresponds to the equilibrium distribution.

### 3. Results

#### 3.1. Theoretical Framework

Consider two finite classical stochastic systems  $X$  and  $Y$  that are weakly coupled to a thermal reservoir characterized by inverse temperature  $\beta$ . Throughout the time interval  $0 \leq t \leq \tau$ , an external driving parameter  $\mu_t$  may force either or both subsystems into non-equilibrium states [34–36]. The temporal evolution of systems  $X$  and  $Y$  under the influence of the time-dependent protocol  $\mu_t$  is governed by classical stochastic dynamics, resulting in stochastic trajectories  $\{x_t\}$  and  $\{y_t\}$ , where  $x_t$  (respectively  $y_t$ ) represents a particular microstate of system  $X$  (respectively  $Y$ ) at time  $t$  within the interval  $0 \leq t \leq \tau$  for any given realization.

Due to the inherent stochasticity of the trajectories, we perform multiple realizations of the driving protocol  $\mu_t$  starting from an initial joint probability distribution  $P_0(x, y)$  defined over the complete set of microstates  $(x, y)$  for the combined system. Through this ensemble of realizations, the coupled subsystems evolve to produce a time-dependent joint probability distribution  $P_t(x, y)$  throughout the duration  $0 \leq t \leq \tau$ . The associated marginal probability distributions are obtained as  $P_t(x) := \int P_t(x, y) dy$  and  $P_t(y) := \int P_t(x, y) dx$ .

We assume

$$P_0(x, y) \neq 0 \text{ for all } (x, y) \quad (6)$$

ensuring that  $P_t(x, y) \neq 0$ ,  $P_t(x) \neq 0$ , and  $P_t(y) \neq 0$  hold for all microstates  $x$  and  $y$  throughout the interval  $0 \leq t \leq \tau$ .

The entropy production  $\Pi$  generated during the protocol  $\mu_t$  over the time period  $0 \leq t \leq \tau$  is defined as

$$\Pi := \Delta \xi_{XY} + \beta H, \quad (7)$$



in which  $\Delta\zeta_{XY}$  represents the stochastic entropy changes along trajectories  $\{x_t\}$  and  $\{y_t\}$ , while  $H$  denotes the heat transferred to the thermal bath (corresponding to entropy generation in the reservoir) [2,3]. Explicitly, these quantities are expressed as

$$\begin{aligned}\Delta\zeta_{XY} &:= \Delta\zeta_X + \Delta\zeta_Y, \\ \Delta\zeta_X &:= -\ln P_\tau(x_\tau) + \ln P_0(x_0), \\ \Delta\zeta_Y &:= -\ln P_\tau(y_\tau) + \ln P_0(y_0).\end{aligned}\tag{8}$$

It is important to recognize that the stochastic entropy  $\zeta[P_t(\circ)] := -\ln P_t(\circ)$  associated with microstate  $\circ$  at time  $t$  quantifies the uncertainty in the occurrence of  $\circ$  at that instant: higher uncertainty in microstate realization corresponds to larger stochastic entropy values. It should be emphasized that the framework in [17] considers only system  $X$  to be thermally coupled to the heat reservoir, while system  $Y$  remains isolated and stationary. Consequently, their entropy production expression becomes  $\Pi_{\text{su}} := \Delta\zeta_X + \beta H$ .

We consider that throughout the temporal evolution governed by process  $\mu_t$ , subsystem  $X$  engages in information transfer with subsystem  $Y$ . This implies that the dynamical path  $\{x_t\}$  followed by system  $X$  is influenced by the corresponding path  $\{y_t\}$  taken by system  $Y$ . The pointwise mutual information  $I_t$  at temporal instance  $t$  characterizing the correlation between microstates  $x_t$  and  $y_t$  quantifies how much the uncertainty in  $x_t$  decreases when  $y_t$  is known [17]:

$$\begin{aligned}I_t(x_t, y_t) &:= \zeta[P_t(x_t)] - \zeta[P_t(x_t|y_t)] \\ &= \ln \frac{P_t(x_t, y_t)}{P_t(x_t)P_t(y_t)},\end{aligned}\tag{9}$$

where  $P_t(x_t|y_t)$  represents the conditional probability density of microstate  $x_t$  conditioned on  $y_t$ . As the degree of statistical dependence between microstates  $x_t$  and  $y_t$  increases, the magnitude of  $I_t(x_t, y_t)$  correspondingly grows larger.

It is important to observe that when  $x_t$  and  $y_t$  exhibit statistical independence at time  $t$ , the quantity  $I_t(x_t, y_t)$  vanishes. The expectation value of  $I_t(x_t, y_t)$  computed with the joint probability distribution  $P_t(x_t, y_t)$  across the entire microstate space yields the standard mutual information between the coupled subsystems, which maintains non-negativity [37].

### 3.2. Proof of Equations (3) and (4)

Now we are ready to prove the information exchange fluctuation theorem conditioned on a coupled-state  $(a, b)$ . We define reverse process  $\mu'_t := \mu_{\tau-t}$  for  $0 \leq t \leq \tau$ , where the external parameter is time-reversed [13,14]. The initial probability distribution  $P'_0(x, y)$  for the reverse process should be the final probability distribution for the forward process  $P_\tau(x, y)$  so that we have

$$\begin{aligned}P'_0(x) &= \int P'_0(x, y) dy = \int P_\tau(x, y) dy = P_\tau(x), \\ P'_0(y) &= \int P'_0(x, y) dx = \int P_\tau(x, y) dx = P_\tau(y).\end{aligned}\tag{10}$$

Then, by Equation (6), we have  $P'_t(x, y) \neq 0$ ,  $P'_t(x) \neq 0$ , and  $P'_t(y) \neq 0$  for all  $x$  and  $y$  during  $0 \leq t \leq \tau$ . For each trajectories  $\{x_t\}$  and  $\{y_t\}$  for  $0 \leq t \leq \tau$ , we define the time-reversed conjugate as follows:

$$\begin{aligned}\{x'_t\} &:= \{x_{\tau-t}^*\}, \\ \{y'_t\} &:= \{y_{\tau-t}^*\},\end{aligned}\tag{11}$$

where  $*$  denotes momentum reversal. Let  $\mathcal{T}$  be the set of all trajectories  $\{x_t\}$  and  $\{y_t\}$ , and  $\mathcal{T}_{x_\tau, y_\tau}$  be that of trajectories conditioned at coupled-microstates  $(x_\tau, y_\tau)$  at time  $\tau$ . By the definition of time-reversed conjugate, the set  $\mathcal{T}'$  of all time-reversed trajectories is identical to  $\mathcal{T}$ , and the set  $\mathcal{T}'_{x'_0, y'_0}$  of

time-reversed trajectories conditioned at  $x'_0$  and  $y'_0$  is identical to  $\mathcal{T}_{x_\tau, y_\tau}$ . Thus we may use the same notation for both forward and backward pairs. We note that the path probabilities  $P_{\mathcal{T}}$  and  $P_{\mathcal{T}_{x_\tau, y_\tau}}$  are normalized over all paths in  $\mathcal{T}$  and  $\mathcal{T}_{x_\tau, y_\tau}$ , respectively. With this notation, the microscopic reversibility condition that enables us to connect the probability of forward and reverse paths to dissipated heat reads as follows [2,38–40]:

$$\frac{P_{\mathcal{T}}(\{x_t\}, \{y_t\} | x_0, y_0)}{P'_{\mathcal{T}}(\{x'_t\}, \{y'_t\} | x'_0, y'_0)} = e^{\beta H}, \quad (12)$$

where  $P_{\mathcal{T}}(\{x_t\}, \{y_t\} | x_0, y_0)$  is the conditional joint probability distribution of paths  $\{x_t\}$  and  $\{y_t\}$  conditioned at initial microstates  $x_0$  and  $y_0$ , and  $P'_{\mathcal{T}}(\{x'_t\}, \{y'_t\} | x'_0, y'_0)$  is that for the reverse process. Now we consider the ensemble of paths  $\mathcal{T}_{a,b}$  and normalized path probability  $P_{\mathcal{T}_{a,b}}$  such that

$$\mathcal{T}_{a,b} = \bigcup_{a(x_\tau)=a, b(y_\tau)=b} \mathcal{T}_{x_\tau, y_\tau} \quad \text{and} \quad \int_{\mathcal{T}_{a,b}} P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\}) d\{x_t\} d\{y_t\} = 1. \quad (13)$$

We restrict our attention to those paths that are in  $\mathcal{T}_{a,b}$ , and divide both numerator and denominator of the left-hand side of Equation (12) by  $P_{\mathcal{T}}(a, b)$ . Since  $P_{\mathcal{T}}(a, b)$  is identical to  $P'_0(a, b)$ , Equation (12) becomes as follows:

$$\frac{P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\} | x_0, y_0)}{P'_{\mathcal{T}_{a,b}}(\{x'_t\}, \{y'_t\} | x'_0, y'_0)} = e^{\beta H} \quad (14)$$

since the probability of paths is now normalized over  $\mathcal{T}_{a,b}$ . Then we have the following:

$$\frac{P'_{\mathcal{T}_{a,b}}(\{x'_t\}, \{y'_t\})}{P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\})} = \frac{P'_{\mathcal{T}_{a,b}}(\{x'_t\}, \{y'_t\} | x'_0, y'_0)}{P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\} | x_0, y_0)} \cdot \frac{P'_0(x'_0, y'_0)}{P_0(x_0, y_0)} \quad (15)$$

$$= \frac{P'_{\mathcal{T}_{a,b}}(\{x'_t\}, \{y'_t\} | x'_0, y'_0)}{P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\} | x_0, y_0)} \cdot \frac{P'_0(x'_0, y'_0)}{P'_0(x'_0)P'_0(y'_0)} \cdot \frac{P_0(x_0)P_0(y_0)}{P_0(x_0, y_0)} \quad (16)$$

$$\times \frac{P'_0(x'_0)}{P_0(x_0)} \cdot \frac{P'_0(y'_0)}{P_0(y_0)}$$

$$= \exp\{-\beta H + I_\tau(x_\tau, y_\tau) - I_0(x_0, y_0) - \Delta\xi_X - \Delta\xi_Y\} \quad (17)$$

$$= \exp\{-\Pi + I_\tau(x_\tau, y_\tau) - I_0(x_0, y_0)\}. \quad (18)$$

To obtain Equation (16) from Equation (15), we multiply Equation (15) by  $\frac{P'_0(x'_0)P'_0(y'_0)}{P'_0(x'_0)P'_0(y'_0)}$  and  $\frac{P_0(x_0)P_0(y_0)}{P_0(x_0)P_0(y_0)}$ , which are 1. We obtain Equation (17) by applying Equations (8)–(10) and (14) to Equation (16). Finally, we use Equation (7) to obtain Equation (18) from Equation (17).

Now we multiply both sides of Equation (18) by  $e^{-I_\tau(x_\tau, y_\tau)}$  and  $P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\})$ , and take integral over all paths in  $\mathcal{T}_{a,b}$ :

$$\begin{aligned} \left\langle e^{-(\Pi + I_0)} \right\rangle_{a,b} &:= \int_{\{x_t\}, \{y_t\} \in \mathcal{T}_{a,b}} e^{-(\Pi + I_0)} P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\}) d\{x_t\} d\{y_t\} \\ &= \int_{\{x_t\}, \{y_t\} \in \mathcal{T}_{a,b}} e^{-I_\tau(x_\tau, y_\tau)} P'_{\mathcal{T}_{a,b}}(\{x'_t\}, \{y'_t\}) d\{x'_t\} d\{y'_t\}. \end{aligned} \quad (19)$$

Noting that

$$P_\tau(a, b) P'_{\mathcal{T}_{a,b}} = P_\tau(x_\tau, y_\tau) P'_{\mathcal{T}_{x_\tau, y_\tau}} \quad (20)$$

we obtain Equation (3) and Equation (4) as follows

$$\begin{aligned} \left\langle e^{-(\Pi+I_0)} \right\rangle_{a,b} &= \int_{a(x_\tau)=a, b(y_\tau)=b} e^{-I_\tau(x_\tau, y_\tau)} \frac{P_\tau(x_\tau, y_\tau)}{P_\tau(a, b)} dx_\tau dy_\tau \\ &\quad \times \int_{\{x_t\}, \{y_t\} \in \mathcal{T}_{x_\tau, y_\tau}} P'_{\mathcal{T}_{x_\tau, y_\tau}}(\{x'_t\}, \{y'_t\}) d\{x'_t\} d\{y'_t\} \\ &= \int_{a(x_\tau)=a, b(y_\tau)=b} e^{-I_\tau(x_\tau, y_\tau)} \frac{P_\tau(x_\tau, y_\tau)}{P_\tau(a, b)} dx_\tau dy_\tau \end{aligned} \quad (21)$$

$$\begin{aligned} &= \int_{a(x_\tau)=a, b(y_\tau)=b} \frac{P_\tau(x_\tau) P_\tau(y_\tau)}{P_\tau(x_\tau, y_\tau)} \frac{P_\tau(x_\tau, y_\tau)}{P_\tau(a, b)} dx_\tau dy_\tau \\ &= \frac{P_\tau(a) P_\tau(b)}{P_\tau(a, b)} \\ &= e^{-I_\tau(a, b)} \end{aligned} \quad (22)$$

Here we use the fact that  $e^{-I_\tau(x_\tau, y_\tau)}$  is constant for all paths in  $\mathcal{T}_{x_\tau, y_\tau}$ , probability distribution  $P'_{\mathcal{T}_{x_\tau, y_\tau}}$  is normalized over all paths in  $\mathcal{T}_{x_\tau, y_\tau}$ , and  $d\{x_t\} = d\{x'_t\}$  and  $d\{y_t\} = d\{y'_t\}$  by the definition of the time-reversal conjugate.

The information exchange fluctuation theorems demonstrate unambiguously that, analogous to how local free energy captures work [5], the pointwise mutual information between coupled states  $(a, b)$  captures entropy production within the ensemble of trajectories terminating at each respective state. The subsequent corollary elucidates entropy production in greater detail from an energetic perspective.

### 3.3. Work Fluctuation Theorem for Information Exchange

Here, we briefly review Jinwoo-Tanaka's work fluctuation theorem. We assume that the initial probability distribution corresponds to the equilibrium distribution at macrostate  $\mu_0$ . The local non-equilibrium free energy  $\phi(x_\tau, \tau)$  captures the work  $W$  performed on the system within the trajectory ensemble  $\mathcal{T}_{x_\tau}$  according to [5,6]:

$$\phi(x_\tau, \tau) = F_{eq}(\mu_0) - \frac{1}{\beta} \ln \left\langle e^{-\beta W} \right\rangle_{\mathcal{T}_{x_\tau}}, \quad (23)$$

where the angular brackets  $\langle \cdot \rangle_{\mathcal{T}_{x_\tau}}$  denote the average over all trajectories terminating at  $x_\tau$  at time  $\tau$ , and  $F_{eq}(\mu_0)$  represents the equilibrium free energy at the control parameter  $\mu_0$ . The local non-equilibrium free energy is expressed as

$$\phi(x, t) := E(x; \mu_t) + k_B T \ln P_t(x), \quad (24)$$

where  $E(x; \mu_t)$  is the internal energy. Under coarse-graining transformations, the fluctuation theorem preserves its functional form [5,6]:

$$\Phi(a, \tau) = F_{eq}(\mu_0) - \frac{1}{\beta} \ln \left\langle e^{-\beta W} \right\rangle_{\mathcal{T}_a}, \quad (25)$$

where the angular brackets  $\langle \cdot \rangle_{\mathcal{T}_a}$  denote the ensemble average over all trajectories that terminate at microstates  $x_\tau$  satisfying the condition  $a(x_\tau) = a$ . The corresponding local non-equilibrium free energy  $\Phi$  takes the form

$$\Phi(a, t) := G(a; \mu_t) + k_B T \ln P_t(a), \quad (26)$$

where  $G(a; \mu_t) := -\beta^{-1} \ln \int_{a(x)=a} e^{-\beta E(x; \mu_t)} dx$  represents the conformational free energy associated with the coarse-grained state  $a$ . The multi-scale relationship is captured by the following identity:

$$\left\langle e^{-\beta \phi(x, t)} \right\rangle_{a(x)=a} = e^{-\beta \Phi(a, t)}, \quad (27)$$



where the angular brackets denote the conditional ensemble average taken with respect to the normalized probability distribution  $P_t(x, t) / P_t(a, t)$ .

We now prove the information exchange work fluctuation theorem. The local free energies  $\phi_X$  and  $\phi_Y$  for systems  $X$  and  $Y$ , respectively, at time  $t$  with macrostate  $\mu_t$  are defined as follows:

$$\begin{aligned}\phi_X(x_t, t) &:= E_X(x_t; \mu_t) - k_B T \xi_X[P_t(x_t)] \\ \phi_Y(y_t, t) &:= E_Y(y_t; \mu_t) - k_B T \xi_Y[P_t(y_t)],\end{aligned}\quad (28)$$

where  $T$  is the temperature of the heat bath,  $k_B$  is the Boltzmann constant,  $E_X$  and  $E_Y$  are the internal energies of systems  $X$  and  $Y$ , respectively, and  $\xi_X$  and  $\xi_Y$  are the stochastic entropies of  $X$  and  $Y$ , respectively [2,3]. Work done on either one or both systems through process  $\mu_t$  is expressed by the first law of thermodynamics as follows:

$$W := \Delta E + H, \quad (29)$$

where  $\Delta E$  is the change in internal energy of the total system composed of  $X$  and  $Y$ . If we assume that systems  $X$  and  $Y$  are weakly coupled, in that interaction energy between  $X$  and  $Y$  is negligible compared to the internal energy of  $X$  and  $Y$ , we may have

$$\Delta E := \Delta E_X + \Delta E_Y, \quad (30)$$

where  $\Delta E_X := E_X(x_\tau, \tau) - E_X(x_0, 0)$  and  $\Delta E_Y := E_Y(y_\tau, \tau) - E_Y(y_0, 0)$  [41]. We rewrite Equation (17) by adding and subtracting the change of internal energy  $\Delta E_X$  of  $X$  and  $\Delta E_Y$  of  $Y$  as follows:

$$\frac{P'_{\mathcal{T}_{a,b}}(\{x'_t\}, \{y'_t\})}{P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\})} = \exp\{-\beta(H + \Delta E_X + \Delta E_Y) + \beta\Delta E_X - \Delta\xi_X + \beta\Delta E_Y - \Delta\xi_Y\} \quad (31)$$

$$\begin{aligned}&\times \exp\{I_\tau(x_\tau, y_\tau) - I_0(x_0, y_0)\} \\ &= \exp\{-\beta W + \beta\phi_X(x_\tau, \tau) - \beta\phi_X(x_0, 0) + \beta\phi_Y(y_\tau, \tau) - \beta\phi_Y(y_0, 0) \\ &\quad \times \exp\{I_\tau(x_\tau, y_\tau) - I_0(x_0, y_0)\}\end{aligned}\quad (32)$$

where we have applied Equations (28)–(30) consecutively to Equation (31) to obtain Equation (32). Now we multiply both sides of Equation (32) by  $e^{-\beta\phi_X(x_\tau, \tau) - \beta\phi_Y(y_\tau, \tau) - I_\tau(x_\tau, y_\tau)}$  and  $P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\})$ , and take integral over all paths in  $\mathcal{T}_{a,b}$  to obtain the following:

$$\begin{aligned}\left\langle e^{-\beta W - \beta\phi_X(x_0, 0) - \beta\phi_Y(y_0, 0) - I_0} \right\rangle_{a,b} &:= \int_{\{x_t\}, \{y_t\} \in \mathcal{T}_{x_\tau, y_\tau}} e^{-\beta W - \beta\phi_X(x_0, 0) - \beta\phi_Y(y_0, 0) - I_0} P_{\mathcal{T}_{a,b}}(\{x_t\}, \{y_t\}) d\{x_t\} d\{y_t\} \\ &= \int_{\{x_t\}, \{y_t\} \in \mathcal{T}_{x_\tau, y_\tau}} e^{-\beta\phi_X(x_\tau, \tau) - \beta\phi_Y(y_\tau, \tau) - I_\tau(x_\tau, y_\tau)} P'_{\mathcal{T}_{a,b}}(\{x'_t\}, \{y'_t\}) d\{x'_t\} d\{y'_t\}\end{aligned}\quad (33)$$

Noting that

$$P_\tau(a, b) P'_{\mathcal{T}_{a,b}} = P_\tau(x_\tau, y_\tau) P'_{\mathcal{T}_{x_\tau, y_\tau}}$$

we obtain

$$\begin{aligned}\left\langle e^{-\beta W - \beta\phi_X(x_0, 0) - \beta\phi_Y(y_0, 0) - I_0} \right\rangle_{a,b} &= \int_{a(x_\tau)=a, b(y_\tau)=b} e^{-\beta\phi_X(x_\tau, \tau) - \beta\phi_Y(y_\tau, \tau) - I_\tau(x_\tau, y_\tau)} \frac{P_\tau(x_\tau, y_\tau)}{P_\tau(a, b)} dx_\tau dy_\tau \\ &\quad \times \int_{\{x_t\}, \{y_t\} \in \mathcal{T}_{x_\tau, y_\tau}} P'_{\mathcal{T}_{x_\tau, y_\tau}}(\{x'_t\}, \{y'_t\}) d\{x'_t\} d\{y'_t\} \\ &= \int_{a(x_\tau)=a, b(y_\tau)=b} e^{-\beta\phi_X(x_\tau, \tau) - \beta\phi_Y(y_\tau, \tau)} \frac{P_\tau(x_\tau) P_\tau(y_\tau)}{P_\tau(x_\tau, y_\tau)} \frac{P_\tau(x_\tau, y_\tau)}{P_\tau(a, b)} dx_\tau dy_\tau \\ &= \int_{a(x_\tau)=a, b(y_\tau)=b} e^{-\beta\phi_X(x_\tau, \tau) - \beta\phi_Y(y_\tau, \tau)} \frac{P_\tau(x_\tau) P_\tau(y_\tau)}{P_\tau(a) P_\tau(b)} \frac{P_\tau(a) P_\tau(b)}{P_\tau(a, b)} dx_\tau dy_\tau \\ &= e^{-I_\tau(a, b)} \int_{a(x_\tau)=a} e^{-\beta\phi_X(x_\tau, \tau)} \frac{P_\tau(x_\tau)}{P_\tau(a)} dx_\tau \int_{b(y_\tau)=b} e^{-\beta\phi_Y(y_\tau, \tau)} \frac{P_\tau(y_\tau)}{P_\tau(b)} dy_\tau \\ &= e^{-\beta\Phi_X(a, \tau) - \beta\Phi_Y(b, \tau) - I_\tau(a, b)},\end{aligned}\quad (34)$$

which generalizes known relations in the literature [15,41–45]. The last equality follows from Equation (27). Under the initial equilibrium condition Equation (34) reduces to Equation (5). We note that Equation (34) holds under the weak-coupling assumption between systems X and Y during process  $\mu_t$ , and  $\Phi_X$  and  $\Phi_Y$  are non-equilibrium free energies, which are different from the equilibrium free energy that appears in similar relations in the literature [15,42–45].

## 4. Conclusions

In this study, we have rigorously established that the fluctuation theorem for information exchange remains invariant under coarse-graining transformations in coupled classical stochastic systems. Our theoretical framework demonstrates that the fundamental relationship between information exchange and entropy production is preserved across different scales of observation.

The key findings of our work include: (i) the derivation of coarse-grained fluctuation theorems that maintain their functional form under arbitrary coarse-graining transformations, (ii) the proof that pointwise mutual information between coarse-grained states captures entropy production within trajectory ensembles, and (iii) the establishment of hierarchical relationships between information measures at different observational scales.

These results have profound implications for understanding information processing in thermodynamic systems. The scale-independence of these fluctuation theorems establishes their fundamental character as universal principles governing energy-information conversion mechanisms, regardless of the level of system description. This universality is particularly significant for biological systems, where information processing occurs across multiple organizational scales, from molecular interactions to cellular signaling networks.

The robust theoretical framework developed here provides a foundation for analyzing the thermodynamics of dynamic molecular information processes and dynamic allosteric transitions in complex biological systems. Furthermore, our findings offer new perspectives on the thermodynamic constraints that govern information processing in classical stochastic systems, with potential applications ranging from molecular machinery design to understanding the fundamental limits of biological computation.

**Funding:** L.J. was supported by the National Research Foundation of Korea Grant funded by the Korean Government (RS-2016-NR017140), and in part by Kwangwoon University Research Grant in 2023.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Jarzynski, C. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.
2. Crooks, G.E. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys. Rev. E* **1999**, *60*, 2721–2726.
3. Seifert, U. Entropy production along a stochastic trajectory and an integral fluctuation theorem. *Phys. Rev. Lett.* **2005**, *95*, 040602.
4. Hatano, T.; Sasa, S.I. Steady-state thermodynamics of Langevin systems. *Phys. Rev. Lett.* **2001**, *86*, 3463–3466.
5. Jinwoo, L.; Tanaka, H. Local non-equilibrium thermodynamics. *Sci.Rep.* **2015**, *5*, 7832.
6. Jinwoo, L. Roles of local nonequilibrium free energy in the description of biomolecules. *Phys. Rev. E* **2023**, *107*, 014402. <https://doi.org/10.1103/PhysRevE.107.014402>.
7. Hummer, G.; Szabo, A. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Nat. Acad. Sci. USA* **2001**, *98*, 3658–3661.
8. Liphardt, J.; Onoa, B.; Smith, S.B.; Tinoco, I.; Bustamante, C. Reversible unfolding of single RNA molecules by mechanical force. *Science* **2001**, *292*, 733–737.
9. Liphardt, J.; Dumont, S.; Smith, S.; Tinoco Jr, I.; Bustamante, C. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science* **2002**, *296*, 1832–1835.
10. Trepagnier, E.H.; Jarzynski, C.; Ritort, F.; Crooks, G.E.; Bustamante, C.J.; Liphardt, J. Experimental test of Hatano and Sasa's nonequilibrium steady-state equality. *Proc. Nat. Acad. Sci. USA* **2004**, *101*, 15038–15041.

11. Collin, D.; Ritort, F.; Jarzynski, C.; Smith, S.B.; Tinoco, I.; Bustamante, C. Verification of the Crooks fluctuation theorem and recovery of RNA folding free energies. *Nature* **2005**, *437*, 231–234.
12. Alemany, A.; Mossa, A.; Junier, I.; Ritort, F. Experimental free-energy measurements of kinetic molecular states using fluctuation theorems. *Nature Phys.* **2012**.
13. Ponmurugan, M. Generalized detailed fluctuation theorem under nonequilibrium feedback control. *Physical Review E* **2010**, *82*, 031129.
14. Horowitz, J.M.; Vaikuntanathan, S. Nonequilibrium detailed fluctuation theorem for repeated discrete feedback. *Physical Review E* **2010**, *82*, 061120.
15. Sagawa, T.; Ueda, M. Generalized Jarzynski equality under nonequilibrium feedback control. *Phys. Rev. Lett.* **2010**, *104*, 090602.
16. Horowitz, J.M.; Parrondo, J.M. Thermodynamic reversibility in feedback processes. *EPL (Europhysics Letters)* **2011**, *95*, 10005.
17. Sagawa, T.; Ueda, M. Fluctuation theorem with information exchange: Role of correlations in stochastic thermodynamics. *Phys. Rev. Lett.* **2012**, *109*, 180602.
18. Zeng, Q.; Wang, J. New fluctuation theorems on Maxwell's demon. *Science Advances* **2021**, *7*, eabf1807, [<https://www.science.org/doi/pdf/10.1126/sciadv.abf1807>]. <https://doi.org/10.1126/sciadv.abf1807>.
19. Yan, L.L.; Bu, J.T.; Zeng, Q.; Zhang, K.; Cui, K.F.; Zhou, F.; Su, S.L.; Chen, L.; Wang, J.; Chen, G.; et al. Experimental Verification of Demon-Involved Fluctuation Theorems. *Phys. Rev. Lett.* **2024**, *133*, 090402. <https://doi.org/10.1103/PhysRevLett.133.090402>.
20. Hartwell, L.H.; Hopfield, J.J.; Leibler, S.; Murray, A.W. From molecular to modular cell biology. *Nature* **1999**, *402*, C47.
21. Crofts, A.R. Life, information, entropy, and time: vehicles for semantic inheritance. *Complexity* **2007**, *13*, 14–50.
22. Cheong, R.; Rhee, A.; Wang, C.J.; Nemenman, I.; Levchenko, A. Information transduction capacity of noisy biochemical signaling networks. *science* **2011**, *334*, 354–358.
23. McGrath, T.; Jones, N.S.; ten Wolde, P.R.; Ouldridge, T.E. Biochemical Machines for the Interconversion of Mutual Information and Work. *Phys. Rev. Lett.* **2017**, *118*, 028101. <https://doi.org/10.1103/PhysRevLett.118.028101>.
24. Ouldridge, T.E.; Govern, C.C.; ten Wolde, P.R. Thermodynamics of Computational Copying in Biochemical Systems. *Phys. Rev. X* **2017**, *7*, 021004. <https://doi.org/10.1103/PhysRevX.7.021004>.
25. Becker, N.B.; Mugler, A.; ten Wolde, P.R. Optimal Prediction by Cellular Signaling Networks. *Phys. Rev. Lett.* **2015**, *115*, 258103. <https://doi.org/10.1103/PhysRevLett.115.258103>.
26. Cheng, F.; Liu, C.; Shen, B.; Zhao, Z. Investigating cellular network heterogeneity and modularity in cancer: a network entropy and unbalanced motif approach. *BMC Systems Biology* **2016**, *10*, 65.
27. Whitsett, J.A.; Guo, M.; Xu, Y.; Bao, E.L.; Wagner, M. SLICE: determining cell differentiation and lineage based on single cell entropy. *Nucleic Acids Research* **2016**, *45*, e54–e54.
28. Statistical Dynamics of Spatial-Order Formation by Communicating Cells. *iScience* **2018**, *2*, 27 – 40.
29. Maire, T.; Youk, H. Molecular-Level Tuning of Cellular Autonomy Controls the Collective Behaviors of Cell Populations. *Cell Systems* **2015**, *1*, 349 – 360.
30. Mehta, P.; Schwab, D.J. Energetic costs of cellular computation. *Proceedings of the National Academy of Sciences* **2012**, *109*, 17978–17982.
31. Govern, C.C.; ten Wolde, P.R. Energy dissipation and noise correlations in biochemical sensing. *Physical review letters* **2014**, *113*, 258102.
32. Jinwoo, L. Fluctuation Theorem of Information Exchange between Subsystems that Co-Evolve in Time. *Symmetry* **2019**, *11*, 433.
33. Jinwoo, L. Fluctuation Theorem of Information Exchange within an Ensemble of Paths Conditioned on Correlated-Microstates. *Entropy* **2019**, *21*.
34. Jarzynski, C. Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale. *Annu. Rev. Codens. Matter Phys.* **2011**, *2*, 329–51.
35. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **2012**, *75*, 126001.
36. Spinney, R.; Ford, I., Fluctuation Relations: A Pedagogical Overview. In *Nonequilibrium Statistical Physics of Small Systems*; Wiley-VCH Verlag GmbH & Co. KGaA, 2013; pp. 3–56.
37. Cover, T.M.; Thomas, J.A. *Elements of information theory*; John Wiley & Sons, 2012.

38. Kurchan, J. Fluctuation theorem for stochastic dynamics. *Journal of Physics A: Mathematical and General* **1998**, *31*, 3719.
39. Maes, C. The fluctuation theorem as a Gibbs property. *Journal of statistical physics* **1999**, *95*, 367–392.
40. Jarzynski, C. Hamiltonian derivation of a detailed fluctuation theorem. *J. Stat. Phys.* **2000**, *98*, 77–102.
41. Parrondo, J.M.; Horowitz, J.M.; Sagawa, T. Thermodynamics of information. *Nature physics* **2015**, *11*, 131–139.
42. Kawai, R.; Parrondo, J.M.R.; den Broeck, C.V. Dissipation: The phase-space perspective. *Phys. Rev. Lett.* **2007**, *98*, 080602.
43. Generalization of the second law for a transition between nonequilibrium states. *Physics Letters A* **2010**, *375*, 88 – 92.
44. Generalization of the second law for a nonequilibrium initial state. *Physics Letters A* **2010**, *374*, 1001 – 1004.
45. Esposito, M.; Van den Broeck, C. Second law and Landauer principle far from equilibrium. *Europhys. Lett.* **2011**, *95*, 40004.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.