Physical nature of higher-order mutual information: Intrinsic correlations and frustration

Hiroyuki Matsuda
Department of Materials Science, Faculty of Integrated Arts and Sciences, Osaka Prefecture University, Sakai, Osaka 599-8531, Japan
(Received 13 December 1999)

This paper studies some properties and implications of higher-order mutual information functions, which should serve for the analysis of general complex systems. We note that the higher-order mutual information can either be positive or negative depending on the correlation among ensembles. Two opposite types of correlations are discussed in connection with the concept of frustration. Simple examples are presented to demonstrate that our concepts are especially helpful in understanding the nature of correlations in frustrated systems. The higher-order mutual information provides an appropriate measure of the frustration effect.

PACS number(s): 05.90.+m, 02.50.−r, 05.20.−y

I. INTRODUCTION

A common approach for the analysis of complex systems is to use concepts from information theory. In particular, information entropy and the related concept of mutual information [1] are of fundamental importance; mutual information can serve as a general measure of correlation between two systems or ensembles. Mutual information as well as information entropy have found significance in various applications in diverse fields, e.g., in analyzing experimental time series [2–4], in characterizing symbol sequences such as DNA sequences [5–7], and in providing a theoretical basis for the notion of complexity [8–12].

Of particular interest in our work is the nature of correlations in systems with many degrees of freedom, which may contain features that are typical of complex systems. Complicating features may occur due to many-body correlation effects such as the frustration effect. In fact, there are many examples of complex systems that contain frustration as an essential ingredient: spin glasses, neural networks, real glasses, colloids, granular media, glass forming liquids, etc. In these systems, frustration (due to many competing interactions or geometrical constraints) causes various fascinating phenomena, such as complicated phase transitions, reentrance phenomena [13–16], partial disorder, nonexponential relaxation [17–20], etc. Appropriate measures of such many-body correlations could, therefore, be expected to contribute to a deeper understanding of complex systems.

In this context, the present work considers higher-order mutual information functions (see, e.g., [21,22]). They allow us to disentangle intrinsic many-body correlations from the insignificant ones governed by lower-order statistics. We will study some aspects or characteristics of higher-order mutual information in order to obtain insight into its meaning. In particular, we focus on the properties of three-body mutual information by considering its relation to the usual mutual information. We will note that higher-order mutual information can either be positive or negative depending on the correlation among ensembles, while the usual mutual information is always non-negative. Then, we will realize two opposite types of correlations among ensembles and relate them to the concept of frustration. To demonstrate the importance of these correlations, we will apply higher-order mutual information to simple examples of frustrated spin systems. Then we will see that the two types of correlations are typical of frustrated and unfrustrated systems, and higher-order mutual information provides an appropriate measure of the frustration effect.

Let us review the mutual information $I(A,B)$ between two ensembles A and B. It is defined in terms of entropies as

$$I(A,B)=S(A)+S(B)-S(AB), \tag{1.1}$$

or, equivalently, in terms of the joint probability distribution $p(a,b)$ as

$$I(A,B)=\sum_{a,b} p(a,b)\ln\frac{p(a,b)}{p(a)p(b)}. \tag{1.2}$$

Here, $S(A)$ and $S(B)$ are the entropies of A and B and $S(AB)$ is the joint entropy of AB. The probability distributions for A and B are given by $p(a)=\Sigma_b p(a,b)$ and $p(b)=\Sigma_a p(a,b)$. [In the case of continuous variables, the summation in Eq. (1.2) may be replaced by integration with respect to a and b.] The function $I(A,B)$ measures the amount of information about A that would be gained from a measurement of B, and vice versa, i.e., the amount of information shared between A and B. Equation (1.1) satisfies

$$0\leqslant I(A,B)\leqslant\min\{S(A),S(B)\}. \tag{1.3}$$

The equality $I(A,B)=0$ holds if A and B are independent, i.e., $p(a,b)=p(a)p(b)$, and the equality $I(A,B)=S(A)$ if A is completely determined by B. The mutual information $I(A,B)$ is smaller when A and B are more independent, and $I(A,B)$ characterizes the degree of correlation between A and B. For the difference between mutual information and the correlation function, one should note the following: (i) while the correlation function only measures linear correlation, mutual information characterizes a general dependence [5]; (ii) mutual information, defined for a joint probability distribution $p(a,b)$, is invariant for a transformation of a,b in contrast to the correlation function; (iii) mutual information can be directly applied to symbolic systems, while the correlation function relies on an assignment of numerical values.

Section II discusses the higher-order mutual information measure. Then some fundamental properties of the measure are derived as direct consequences of the definition; we obtain recursion relations and use them to derive some basic inequalities. Section III provides a novel concept for the nature of correlations in terms of higher-order mutual information. Simple examples are presented in Sec. IV, showing some important features of the measure. Section V contains discussion and conclusion.

II. HIGHER-ORDER MUTUAL INFORMATION AND ITS FUNDAMENTAL PROPERTIES

A. Higher-order mutual information

Let us consider a joint ensemble $A_1 \cdots A_n$, with A_i ($i = 1, \ldots, n$) being individual ensembles. To be specific, suppose that the ensemble $A_1 \cdots A_n$ is described by a discrete probability distribution $p(a_1, \ldots, a_n)$ [which is normalized such that $\Sigma_{\{a_i\}} p(a_1, \ldots, a_n) = 1$]. Then the entropy of $A_1 \cdots A_n$ is

$$S(A_1 \cdots A_n) = -\sum_{\{a_i\}} p(a_1, \ldots, a_n) \ln p(a_1, \ldots, a_n). \tag{2.1}$$

The reduced ensemble $A_1 \cdots A_{n-1}$, with its probability distribution $p(a_1, \ldots, a_{n-1})$, is related to $A_1 \cdots A_n$ by

$$p(a_1, \ldots, a_{n-1}) = \sum_{a_n} p(a_1, \ldots, a_n), \tag{2.2}$$

and its entropy is defined analogously. A similar relation holds for any reduced ensemble associated with $A_1 \cdots A_n$.

In a manner analogous to Eq. (1.1), the mutual information among three ensembles A_1, A_2, and A_3 can be defined as

$$\begin{aligned} I_3(A_1, A_2, A_3) = {} & S(A_1) + S(A_2) + S(A_3) - S(A_1 A_2) \\ & - S(A_1 A_3) - S(A_2 A_3) + S(A_1 A_2 A_3). \end{aligned} \tag{2.3}$$

Furthermore, Eqs. (1.1) and (2.3) can be generalized by

$$I_n(A_1, \ldots, A_n) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{i_1 < \cdots < i_k} S(A_{i_1} \cdots A_{i_k}), \tag{2.4}$$

where the sum $\Sigma S(A_{i_1} \cdots A_{i_k})$ runs over all possible combinations $\{i_1, \ldots, i_k\} \in \{1, \ldots, n\}$. I_2 is the usual mutual information. Notice that $I_n(A_1, \ldots, A_n)$ is symmetric under any permutation of $A_1, \ldots, A_n$. The generalized mutual information may be recognized as common information or entropy shared among n ensembles, in analogy with the usual mutual information. For example, I_3 may be viewed in Fig. 1 as the overlap of $S(A_1)$, $S(A_2)$, and $S(A_3)$. Note, however, that while $S(A_i)$ is a non-negative function, the quantity I_n for $n \geqslant 3$ can be not only positive but negative in contrast to the usual mutual information. Before focusing on this important fact, we consider some fundamental properties of the function (2.4).

If the n ensembles A_i are independent, i.e., the joint probability distribution is of the form

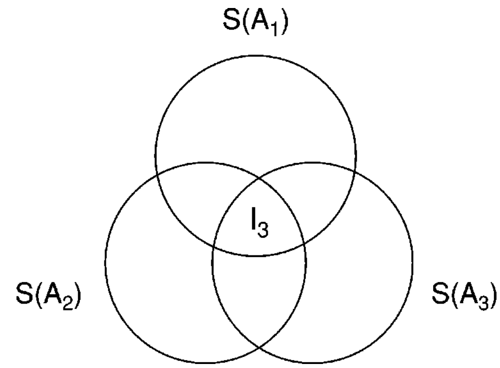

FIG. 1. An entropy diagram showing the three-body mutual information $I_3(A_1, A_2, A_3)$.

$$p(a_1, \ldots, a_n) = p(a_1) p(a_2) \cdots p(a_n), \tag{2.5}$$

then $S(A_{i_1} \cdots A_{i_k})$ is the sum of the individual entropies and $I_n(A_1, \ldots, A_n) = 0$. If one ensemble A_i is completely determined by any other A_j, then Eq. (2.4) results in $I_n(A_1, \ldots, A_n) = S(A_i)$, since $S(A_i A_{i_1} \cdots A_{i_k})$ reduces to $S(A_{i_1} \cdots A_{i_k})$. Notice that for $n \geqslant 3$, Eq. (2.5) is not a unique distribution that yields $I_n = 0$, while $I_2 = 0$ holds only if $p(a_1 a_2) = p(a_1) p(a_2)$. A more general condition for $I_n = 0$ is established in the following.

In analogy with the expression (1.2), the function (2.3) can be directly expressed in terms of the joint probability distribution $p(a_1, a_2, a_3)$ as

$$I_3(A_1, A_2, A_3) = -\sum_{a_1, a_2, a_3} p(a_1, a_2, a_3) \ln \frac{p(a_1, a_2, a_3)}{\hat{p}(a_1, a_2, a_3)}, \tag{2.6}$$

with

$$\hat{p}(a_1, a_2, a_3) = \frac{p(a_1, a_2) p(a_2, a_3) p(a_1, a_3)}{p(a_1) p(a_2) p(a_3)}. \tag{2.7}$$

The function (2.7), which corresponds to the Kirkwood superposition approximation [23], provides an estimate for the three-body distribution $p(a_1, a_2, a_3)$, given the two-body distributions. It follows immediately that $I_3(A_1, A_2, A_3) = 0$ for $p(a_1, a_2, a_3) = \hat{p}(a_1, a_2, a_3)$; this implies that I_3 measures the intrinsic three-body correlation, which is not governed by two-body statistics. In general, we have the expression

$$\begin{aligned} I_n(A_1, \ldots, A_n) = {} & (-1)^n \sum_{a_1, \ldots, a_n} p(a_1, \ldots, a_n) \\ & \times \ln \frac{p(a_1, \ldots, a_n)}{\hat{p}(a_1, \ldots, a_n)}, \end{aligned} \tag{2.8}$$

with

$$\hat{p}(a_1,\ldots,a_n)=\prod_{i_1<\cdots<i_{n-1}} p(a_{i_1},\ldots,a_{i_{n-1}})/ \times \prod_{i_1<\cdots<i_{n-2}} p(a_{i_1},\ldots,a_{i_{n-2}})/ \cdots/\prod_{i} p(a_i). \tag{2.9}$$

Equation (2.9) can be recognized as the generalized Kirkwood superposition approximation [24]. Again, note that $I_n(A_1,\ldots,A_n)=0$ for $p(a_1,\ldots,a_n)=\hat{p}(a_1,\ldots,a_n)$.

It should be mentioned that the generalized mutual information functions give the following contributions to the entropy:

$$S(A_1\cdots A_n)=\sum_{i=1}^{n} S(A_i)-\sum_{i<j} I_2(A_i,A_j) + \sum_{i<j<k} I_3(A_i,A_j,A_k)-\cdots. \tag{2.10}$$

That is, the relative magnitude of the functions is a measure of the contribution to the global behavior of the system $A_1 \cdots A_n$. An approximation for the entropy can be made by neglecting higher-order contributions. This approximation may yield good results for weakly correlated ensembles, with the first-order approximation corresponding exactly with the entropy for the case (2.5).

B. Some general properties of I_n

Now we consider general relationships between the functions (2.4). First, we can write $I_3(A_1,A_2,A_3)$ in terms of the usual mutual information as

$$I_3(A_1,A_2,A_3)=I_2(A_1,A_2)+I_2(A_1,A_3)-I_2(A_1,A_2A_3), \tag{2.11}$$

where the quantity

$$I_2(A_1,A_2A_3)=S(A_1)+S(A_2A_3)-S(A_1A_2A_3) \tag{2.12}$$

is the mutual information between A_1 and A_2A_3. In contrast to the sum $I_2(A_1,A_2)+I_2(A_1,A_3)$, $I_2(A_1,A_2A_3)$ measures the amount of information about A_1 that would be gained from simultaneous measurements of A_2 and A_3. If A_1 and A_2A_3 are independent, each term on the right-hand side of Eq. (2.11) is zero and $I_3(A_1,A_2,A_3)=0$. However, it is not necessary that if $I_2(A_1,A_2)$ and $I_2(A_1,A_3)$ are zero, A_1 should be independent of A_2A_3; thus it is possible that $I_3<0$. Notice that the relation (2.11) is symmetric under the permutations $A_1\leftrightarrow A_2$ and $A_1\leftrightarrow A_3$ (because of the symmetry in the definition (2.3)].

It follows that the higher-order mutual information functions I_n can be related to the lower-order ones I_{n-1} as

$$\begin{aligned} I_n(A_1,\ldots,A_n)=&I_{n-1}(A_1,\ldots,A_{n-2},A_{n-1}) \\ &+I_{n-1}(A_1,\ldots,A_{n-2},A_n) \\ &-I_{n-1}(A_1,\ldots,A_{n-2},A_{n-1}A_n). \end{aligned} \tag{2.13}$$

Again, notice that the permutation symmetry gives a set of recursion relations similar to Eq. (2.13). These relations allow us to express $I_n(A_1,\ldots,A_n)$ in terms of the mutual information functions of lower order than n. For example, we can write

$$\begin{aligned} I_4(A_1,A_2,A_3,A_4)=&I_2(A_1,A_2)+I_2(A_1,A_3)+I_2(A_1,A_4) \\ &-I_2(A_1,A_2A_3)-I_2(A_1,A_2A_4) \\ &-I_2(A_1,A_3A_4)+I_2(A_1,A_2A_3A_4). \end{aligned} \tag{2.14}$$

From expressions such as Eqs. (2.11) and (2.14), we recognize $I_n(A_1,\ldots,A_n)$ as the common information shared by $I_2(A_i,A_j)$. More generally, from the hierarchy of the function (2.4) it is natural to recognize I_n as the common information shared by lower-order functions $I_{n'}$ $(n'<n)$ when I_n is recognized as the common information shared by $S(A_i)$.

In analogy with the concept of conditional entropy, we can consider conditional quantities of the mutual information functions I_n. As a consequence, the function $I_2(A_1,A_2A_3)$ can be decomposed into the sum

$$I_2(A_1,A_2A_3)=I_2(A_1,A_3)+I_2(A_1,A_2|A_3). \tag{2.15}$$

The quantity $I_2(A_1,A_2|A_3)$ is the mutual information between A_1 and A_2 that is conditional on a measurement of A_3:

$$\begin{aligned} I_2(A_1,A_2|A_3)=&\sum_{a_3} p(a_3)\sum_{a_1,a_2} p(a_1,a_2|a_3) \\ &\times\ln\frac{p(a_1,a_2|a_3)}{p(a_1|a_3)p(a_2|a_3)} \\ =&S(A_1|A_3)+S(A_2|A_3)-S(A_1A_2|A_3). \end{aligned} \tag{2.16}$$

Here $p(a_1|a_3)=p(a_1,a_3)/p(a_3)$ is the conditional probability distribution for the variable a_1 given a measurement a_3, and $S(A_1|A_3)=S(A_1A_3)-S(A_3)$ is the conditional entropy of A_1 with respect to A_3. Equation (2.16) satisfies

$$0\leqslant I_2(A_1,A_2|A_3)\leqslant\min\{S(A_1|A_3),S(A_2|A_3)\}. \tag{2.17}$$

The equality $I_2(A_1,A_2|A_3)=0$ holds if A_1 and A_2 are statistically independent when A_3 is specified, i.e., if $p(a_1,a_2|a_3)=p(a_1|a_3)p(a_2|a_3)$. The equality $I_2(A_1,A_2|A_3)=S(A_1|A_3)$ holds if A_1 is completely determined by A_2 when A_3 is specified. In general, we can write

$$\begin{aligned} I_{n-1}(A_1,\ldots,A_{n-1}A_n)=&I_{n-1}(A_1,\ldots,A_{n-2},A_n) \\ &+I_{n-1}(A_1,\ldots,A_{n-1}|A_n), \end{aligned} \tag{2.18}$$

where

$$I_{n-1}(A_1,\ldots,A_{n-1}|A_n)=\sum_{k=1}^{n-1}(-1)^{k+1}\times\sum_{i_1<\cdots<i_k}S(A_{i_1}\cdots A_{i_k}|A_n), \tag{2.19}$$

$\{i_1,\ldots,i_k\}\in\{1,\ldots,n-1\}$, is the conditional mutual information among $A_1,\ldots,A_{n-1}$ with respect to A_n.

Inserting Eq. (2.15) into Eq. (2.11), we have the relation

$$I_3(A_1,A_2,A_3)=I_2(A_1,A_2)-I_2(A_1,A_2|A_3). \tag{2.20}$$

Combining Eq. (2.20) and the non-negativity of the function $I_2(A_1,A_2|A_3)$, we see that

$$I_3(A_1,A_2,A_3)\leqslant\min\{I_2(A_1,A_2),I_2(A_1,A_3),I_2(A_2,A_3)\}\leqslant\min\{S(A_1),S(A_2),S(A_3)\}. \tag{2.21}$$

On the other hand, the non-negativity of the function $I_2(A_1,A_2)$ yields

$$I_3(A_1,A_2,A_3)\geqslant-\min\{I_2(A_1,A_2|A_3),I_2(A_1,A_3|A_2),I_2(A_2,A_3|A_1)\}, \tag{2.22}$$

and then using the inequality (2.17) we see that the possible lower limit of I_3 is

$$I_3(A_1,A_2,A_3)\geqslant-\min_{i,j}S(A_i|A_j)\geqslant-\min_{i}S(A_i), \tag{2.23}$$

where $\{i,j\}\in\{1,2,3\}$. While the condition (2.21) is straightforward to explain, the condition (2.23) is complicated: the equality $I_3(A_1,A_2,A_3)=-S(A_1|A_3)$ is attained if A_1 is statistically independent of A_2 but is completely determined by A_2 when A_3 is specified.

In general, inserting Eq. (2.18) into Eq. (2.13), we have the relation

$$I_n(A_1,\ldots,A_n)=I_{n-1}(A_1,\ldots,A_{n-1})-I_{n-1}(A_1,\ldots,A_{n-1}|A_n). \tag{2.24}$$

Therefore, the functions $I_n(A_1,\ldots,A_n)$ and $I_{n-1}(A_1,\ldots,A_{n-1})$ are associated as follows:

$$I_n(A_1,\ldots,A_n)\leqslant I_{n-1}(A_1,\ldots,A_{n-1})\Leftrightarrow I_{n-1}(A_1,\ldots,A_{n-1}|A_n)\geqslant 0 \tag{2.25}$$

and

$$I_n(A_1,\ldots,A_n)>I_{n-1}(A_1,\ldots,A_{n-1})\Leftrightarrow I_{n-1}(A_1,\ldots,A_{n-1}|A_n)<0, \tag{2.26}$$

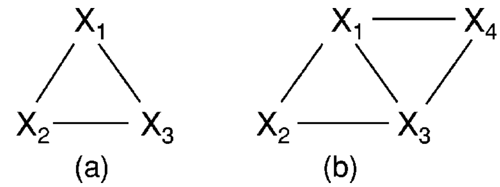

FIG. 2. Spin systems with couplings J, (a) consisting of three binary spins and (b) consisting of four binary spins, with no direct interaction between X_1 and X_4. Frustration arises when $J<0$. Note that in the case $J<0$, despite the presence of frustration, the system (b) has a stabilizing effect as a whole.

which is possible for $n\geqslant 4$. Whether $I_n<I_{n-1}$ or $I_n>I_{n-1}$, together with the sign of the measure, will be considered as an indication of the structure of correlation among ensembles. In the following section we clarify the differences between correlations characterized by positive and negative values of I_n.

III. CONCEPT OF FRUSTRATED CORRELATION

Correlation between A_1 and A_2 implies that A_1A_2 has certain preferred combinations of a_1 and a_2: two-body preference. When considering the ensembles A_1A_2, A_1A_3, and A_2A_3, the two-body preferences may be simultaneously satisfied or not. In this regard, let us consider the meaning of the three-body mutual information $I_3(A_1,A_2,A_3)$, focusing on the fact that $I_3(A_1,A_2,A_3)$ can either be positive or negative depending on the correlation among A_1,A_2,A_3. When given a measurement of A_1, one obtains information about A_2 and simultaneously information about A_3 as well. As mentioned in the preceding section, $I_3(A_1,A_2,A_3)$ can be recognized as the common information shared by $I_2(A_1,A_2)$ and $I_2(A_1,A_3)$. If the two-body preferences are simultaneously satisfied, then the information $I_2(A_1,A_2)$ should contain part of the information $I_2(A_1,A_3)$, which implies that $I_3(A_1,A_2,A_3)>0$. Thus, if the three-body mutual information is negative, we can recognize that the two-body preferences are simultaneously unsatisfied; we call such correlations frustrated. Similar considerations can apply to the higher-order functions I_n with the recognition of I_n in terms of the usual mutual information. Consequently, the correlations among n ensembles with negative I_n can be considered frustrated. In the following section, we will see that these correlations are especially important in frustrated statistical systems such as spin glasses.

IV. EXAMPLES

Now we apply the generalized mutual information to simple examples of spin systems to illustrate some important features of the measure. Let us first consider the system $X_1X_2X_3$ composed of three binary spins X_1, X_2, X_3, described in Fig. 2(a), with the Hamiltonian

$$H=-J(x_1x_2+x_2x_3+x_1x_3), \tag{4.1}$$

where the spin variable x_i takes values ±1 and the coupling J is set equal to 1 or -1. The interactions $-Jx_ix_j$ give rise to frustration when choosing $J=-1$. The probability distribution for the system is given by $p(x_1,x_2,x_3)=e^{-\beta H}/Z$,

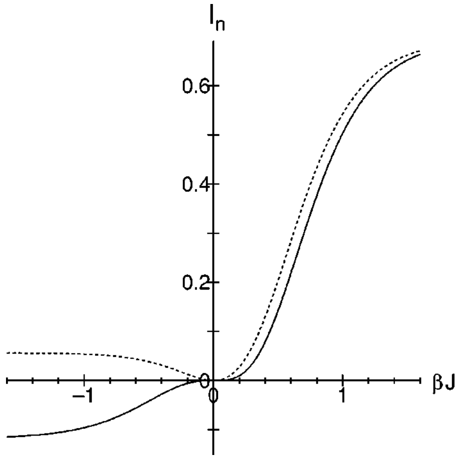

FIG. 3. Plots of the functions I_2 and I_3 given by Eqs. (4.4) and (4.5) as functions of βJ. Dotted line: $I_2(X_1,X_2)$. Full line: $I_2(X_1,X_2,X_3)$. While in the limit $\beta J \to \infty$ each function goes to $\ln 2$, in the limit $\beta J \to -\infty$, I_2 goes to 0.0566 and I_3 goes to -0.1177.

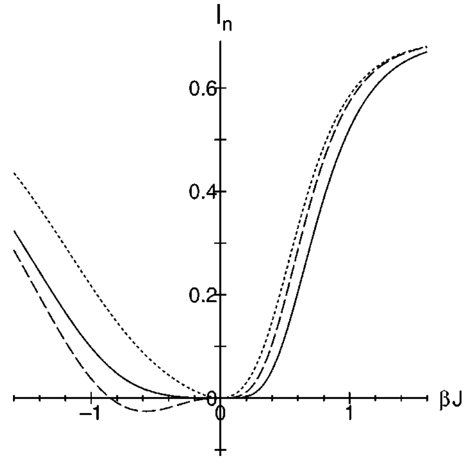

FIG. 4. Plots of the functions I_2, I_3, I_4 calculated for the system described in Fig. 2(b). Dotted line: $I_2(X_1,X_2)$. Dashed line: $I_3(X_1,X_2,X_3)$. Full line: $I_4(X_1,X_2,X_3,X_4)$. In the limit $\beta J \to \pm\infty$ each function goes to $\ln 2$.

where $Z = 2e^{3\beta J} + 6e^{-\beta J}$ and $\beta = 1/T$ is the inverse temparature. The mutual information between X_1 and X_2 is

$$I_2(X_1,X_2) = 2S(X_1) - S(X_1X_2), \tag{4.2}$$

and the three-body mutual information among X_1, X_2, and X_3 is

$$I_3(X_1,X_2,X_3) = 3S(X_1) - 3S(X_1X_2) + S(X_1X_2X_3). \tag{4.3}$$

Elementary calculation gives

$$I_2(X_1,X_2) = 3\ln 2 + \ln\frac{e^{-\beta J}}{Z} + 2\frac{e^{3\beta J}+e^{-\beta J}}{Z}\ln\frac{e^{3\beta J}+e^{-\beta J}}{2e^{-\beta J}} \tag{4.4}$$

and

$$I_3(X_1,X_2,X_3) = 6\ln 2 - 2\ln Z + 6\frac{e^{3\beta J}+e^{-\beta J}}{Z}\ln\frac{e^{3\beta J}+e^{-\beta J}}{2e^{\beta J}}. \tag{4.5}$$

These are plotted as functions of βJ in Fig. 3, which draws a comparison between the two cases $J=1$ and $J=-1$. In the limit $T\to\infty$ ($\beta J = 0$), they vanish since the spins X_1,X_2,X_3 become independent. In the case $J=1$ ($\beta J \geqslant 0$), both functions monotonically increase as the temperature is lowered, and in the limit $T\to 0$ ($\beta J \to \infty$) they go to $\ln 2$ since the spins become completely dependent. In contrast, in the case $J=-1$ ($\beta J<0$) the difference between the two functions is remarkable; the three-body mutual information is negative due to the frustration effect. In this case, the function I_3 monotonically decreases with increasing the frustration effect, and in the limit $T\to 0$ ($\beta J\to -\infty$), one has $I_2 = \frac{3}{5}\ln 2 - \ln 3 = 0.0566$ and $I_3 = 3\ln 2 - 2\ln 3 = -0.1177$. It should be pointed out that the three-point correlation function, defined by $\langle x_1x_2x_3\rangle$, cannot indicate any correlation in $X_1X_2X_3$ since it trivially results in $\langle x_1x_2x_3\rangle = 0$, while I_3 characterizes correlations that are typical of the considered frustrated system.

Next we consider the system described in Fig. 2(b) (a bond defined along the line that connects two adjacent spins X_i). Each bond is characterized by the same coupling constant J, which can be either $+1$ or -1. The three mutual information functions $I_2(X_1,X_2)$, $I_3(X_1,X_2,X_3)$, and $I_4(X_1,X_2,X_3,X_4)$ are shown in Fig. 4. In the case of ferromagnetic interactions $J=1$ ($\beta J>0$), they are all similar in behavior. However, an interesting feature occurs in the case of antiferromagnetic interactions $J=-1$ ($\beta J \leqslant 0$). Note that in this case, although frustration arises in two triangles, the spins X_i simultaneously have a stabilizing effect as a whole; in fact, in the limit $T\to 0$ ($\beta J\to -\infty$) the spins become completely dependent (while the interaction between X_1 and X_3 is unsatisfied) and all the functions go to $\ln 2$, as in the case $J=1$. Thus the behavior of the system at finite temperatures is a consequence of the frustration and the stabilizing effects as well as thermal fluctuation. These competing effects lead to the minimum in the three-body mutual information. The stabilizing effect overbalances the frustration effect below the temperature at which $I_3=0$. It should be noted that the case $J=-1$ gives $I_3<I_4$, in contrast to the case $J=1$; this is because frustration is contained in the triangle $X_1X_2X_3$ and the conditional mutual information $I_3(X_1,X_2,X_3|X_4)$ is still negative [compare the relationship (2.26)]. The four-body

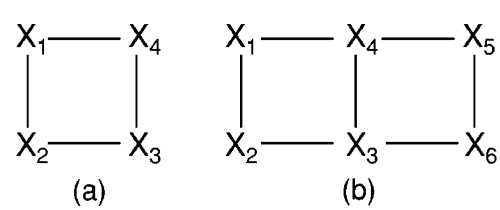

FIG. 5. Spin systems with nearest-neighbor interactions J_{ij}. The system (a) contains frustration when choosing J_{ij} such that $J \equiv J_{12}J_{23}J_{34}J_{14}<0$. The system (b) has couplings chosen such that $J\equiv J_{34}J_{45}J_{56}J_{36} = J_{12}J_{23}J_{34}J_{14}$, and in the case $J<0$, despite the presence of frustration, it has a stabilizing effect as a whole, similar to the system described in Fig. 2(b).

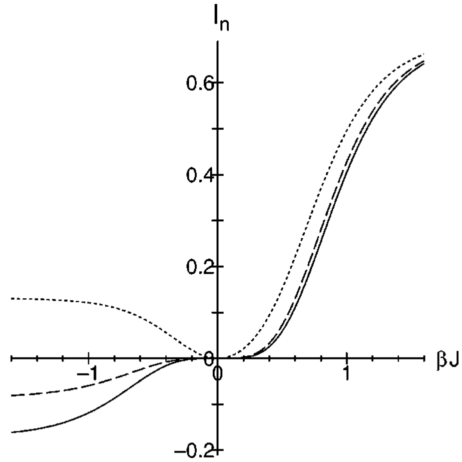

FIG. 6. Plots of the functions I_2, I_3, I_4 calculated for the system described in Fig. 5(a). Dotted line: $I_2(X_1,X_2)$. Dashed line: $I_3(X_1,X_2,X_3)$. Full line: $I_4(X_1,X_2,X_3,X_4)$. While in the limit $\beta J\rightarrow\infty$ each function goes to ln 2, in the limit $\beta J\rightarrow-\infty$, I_2,I_3,I_4 go to 0.1308, -0.0849, and -0.1698, respectively.

mutual information implies that the stabilizing effect is responsibe for the behavior of the spins as a whole.

Similar situations occur in the following. First consider the system described in Fig. 5(a), with couplings J_{ij} between nearest-neighbor spins X_i and X_j. Here we set $J_{ij}=1$ or -1. Define $J\equiv J_{12}J_{23}J_{34}J_{14}$. The system is frustrated if it contains one or three negative interactions, i.e., when $J=-1$. As shown in Fig. 6, in the case $J=-1$, the three- and four-body mutual information functions are negative due to the frustration effect. Next consider the system described in Fig. 5(b) with couplings J_{ij} ($J_{ij}=\pm 1$). Let us choose J_{ij} such that $J\equiv J_{12}J_{23}J_{34}J_{14}=J_{34}J_{45}J_{56}J_{36}$. Then the system has similar properties to the system described in Fig. 2(b); that is, in the case $J=-1$ although the system has frustration in each square, it has the stabilizing effect as a whole. The same mutual information functions as given in Fig. 6, calculated for this system, are shown in Fig. 7. In the case $J=-1$ the three- and four-body functions have minima at finite temperatures as a result of the two competing effects.

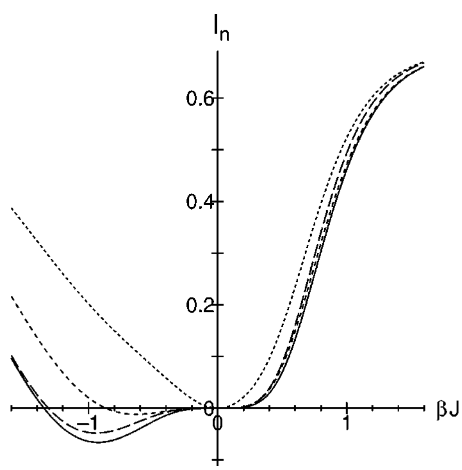

FIG. 7. Plots of the functions I_2, I_3, I_4 calculated for the system described in Fig. 5(b). Dotted line: $I_2(X_1,X_2)$. Short dashed line: $I_3(X_1,X_2,X_3)$. Dashed line: $I_3(X_1,X_3,X_4)$. Full line: $I_4(X_1,X_2,X_3,X_4)$. In the limit $\beta J\rightarrow\pm\infty$, each function goes to ln 2.

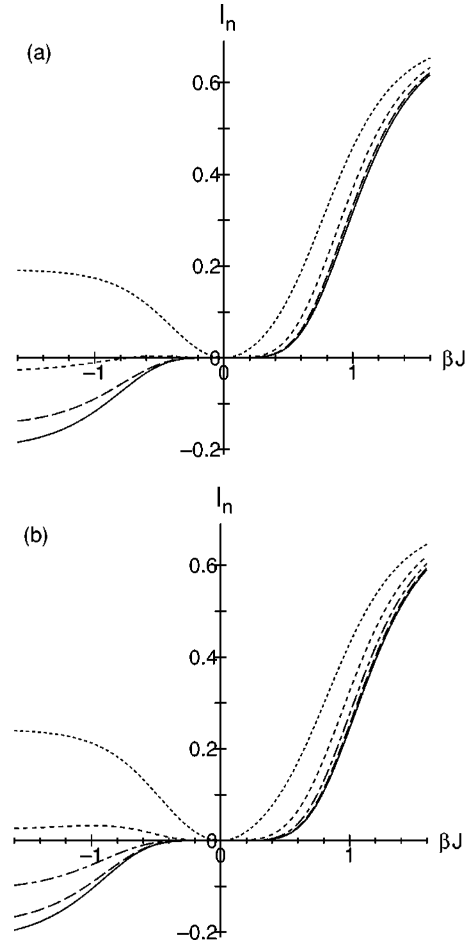

FIG. 8. Plots of the functions $I_2,I_3,\ldots,I_n$ calculated for the system $X_1X_2\cdots X_n$ with (a) $n=5$ and (b) $n=6$ spins forming a ring with nearest-neighbor interactions J_{ij}. The system contains frustration when choosing $J\equiv J_{12}J_{23}\cdots J_{1n}<0$. From upper to lower curves: $I_2(X_1,X_2),I_3(X_1,X_2,X_3),\ldots,I_n(X_1,\ldots,X_n)$.

Finally, consider systems in which the spins $X_1,X_2,\ldots,X_n$ form a ring with nearest-neighbor interactions $J_{12},J_{23},\ldots,J_{n1}$ ($J_{ij}=\pm 1$). Define $J\equiv J_{12}J_{23}\cdots J_{n1}$. Again, frustration arises when $J=-1$. Figure 8 shows the functions $I_2(X_1,X_2),I_3(X_1,X_2,X_3),\ldots,I_n(X_1,\ldots,X_n)$ calculated for the systems with $n=5$ and $n=6$ spins. If there is no frustration, i.e., in the case $J=1$, any mutual information functions including the conditional functions (2.19) are positive, and consequently, higher-order functions are smaller than lower-order ones following the relationship (2.25). In the case $J=-1$, the functions I_n of higher order than $n=3$ monotonically decrease with increasing the frustration effect. The function I_3 has a maximum at a finite temperature due to the competition between local ordering of spins and the frustration effect as a whole. The behaviors of these functions indicate that the frustration effect is more responsible for higher-order correlations.

V. DISCUSSION AND CONCLUSION

We have derived some properties and implications of higher-order mutual information (HMI) functions. The most important feature is that HMI can either be positive or negative depending on the correlation among ensembles, whereas the usual mutual information is always non-negative. Thus the HMI measure separates possible many-body correlations into two opposite types according to its sign. The type of correlation characterized by negative HMI is a remarkable phenomenon, and the sign of HMI serves as its own indicator. We called the phenomenon the frustrated correlation.

We have demonstrated the importance of HMI, together with the phenomenon of negative HMI, by applying the measure to simple examples of frustrated spin systems. We find that the frustration effect lowers HMI, while the thermal fluctuation effect decreases its absolute value; thus the phenomenon of negative HMI occurs due to the frustration effect. On the other hand, the stabilizing effect raises HMI naturally. In the presence of the frustration and the stabilizing effects, we have shown a characteristic behavior of HMI as a function of temperature; the minimum in HMI occurs at a finite temperature as a result of the competition between the two effects. It is important to notice that in the phenomenon of negative HMI, the frustration effect stands opposite to any other effect, in contrast to the case where HMI is positive. It is also remarkable that for negative HMI the thermal fluctuation effect is in the same direction as the stabilizing effect, while they are completely opposite when HMI is positive.

Our concepts should help to obtain deeper insights into complex systems that contain frustration as an essential ingredient. The HMI measure could clarify the complicated nature of correlations in such systems and then allows us to reveal the presence of frustration and the role played by it. An important feature is that the measure could characterize the competition between the frustration effect and some other effects, which may be responsible for complex behavior of frustrated systems.

ACKNOWLEDGMENTS

The author is grateful to Professor K. Kitaura for useful discussions and suggestions. The author also wishes to thank Professor K. Kudo, Professor R. Nakamura, and Dr. O. Yamakawa for helpful comments.

[1] C. E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).
[2] A. M. Fraser and H. L. Swinney, Phys. Rev. A **33**, 1134 (1986).
[3] J. A. Vastano and H. L. Swinney, Phys. Rev. Lett. **60**, 1773 (1988).
[4] A. M. Fraser, Physica D **34**, 391 (1989).
[5] W. Li, J. Stat. Phys. **60**, 823 (1990).
[6] H. Herzel, A. O. Schmitt, and W. Ebeling, Chaos Solitons Fractals **4**, 97 (1994).
[7] H. Herzel and I. Grosse, Phys. Rev. E **55**, 800 (1997).
[8] P. Grassberger, Int. J. Theor. Phys. **25**, 907 (1986).
[9] R. Wackerbauer, A. Witt, H. Atmanspacher, J. Kurths, and H. Scheingraber, Chaos Solitons Fractals **4**, 133 (1994).
[10] H. Matsuda, K. Kudo, R. Nakamura, O. Yamakawa, and T. Murata, Int. J. Theor. Phys. **35**, 839 (1996).
[11] T. Mori, K. Kudo, Y. Tamagawa, R. Nakamura, O. Yamakawa, H. Suzuki, and T. Uesugi, Physica D **116**, 275 (1998).
[12] D. P. Feldman and J. P. Crutchfield, Phys. Lett. A **238**, 244 (1998).
[13] W. M. Saslowk and G. Parker, Phys. Rev. Lett. **56**, 1074 (1986).
[14] P. Azaria, H. T. Diep, and H. Giacomini, Phys. Rev. Lett. **59**, 1629 (1987).
[15] M. Debauche, H. T. Diep, P. Azaria, and H. Giacomini, Phys. Rev. B **44**, 2369 (1991).
[16] Y. Ma and C. Gong, Phys. Rev. E **51**, 1573 (1995).
[17] I. A. Campbell, J. M. Flesselles, R. Jullien, and R. Botet, Phys. Rev. B **37**, 3825 (1988).
[18] S. Scarpetta, A. de Candia, and A. Coniglio, Phys. Rev. E **55**, 4943 (1997).
[19] A. Fierro, G. Franzese, A. de Candia, and A. Coniglio, Phys. Rev. E **59**, 60 (1999).
[20] G. Franzese and A. Coniglio, Phys. Rev. E **59**, 6409 (1999).
[21] N. J. Cerf and C. Adami, Physica D **120**, 62 (1998).
[22] N. J. Cerf, Phys. Rev. A **57**, 3330 (1998).
[23] J. G. Kirkwood, J. Chem. Phys. **10**, 394 (1942).
[24] P. Attard, O. G. Jepps, and S. Marcelja, Phys. Rev. E **56**, 4052 (1997).