Quantification of correlations in quantum many-particle systems

Dieter Vollhardt

Supported by DFG
TRR 80
FOR 1346

Institut für Physik, Universität Mainz; 25. Januar 2013
Outline:

• Conventional measures of electronic correlations
• Classical information theory
  - information entropy
  - relative entropy
  - mutual information, correlations
• Excursion: Relative entropy and susceptibilities in thermodynamics
• Generalization to quantum information theory
• Applications:
  - Hubbard model (para, antiferro)
  - Transition-metal oxide series MnO, FeO, CoO, NiO

In collaboration with
K. Byczuk (Warsaw), J. Kuneš (Prague), W. Hofstetter (Frankfurt)
PRL 108, 087004 (2012)
arXiv:1110.3214
Correlation strength

E.g.: Weakly, moderately, highly, strongly, extremely correlated electrons

Conventional quantitative measures of correlation strength:

\[
\frac{U}{t}, \quad \frac{U}{W}, \quad \frac{m^*}{m}, \quad \frac{E_{\text{corr}}}{E - E_{\text{HF}}}, \quad \frac{E_{\text{HF}}}{E_{\text{corr}}}, \quad \frac{\langle n_{i\uparrow}n_{i\downarrow} \rangle}{\langle n_{i\uparrow} \rangle \langle n_{i\downarrow} \rangle}
\]

Very specific quantities

Related to the uncorrelated state: Correlations are relative

Is there an objective way to
- quantify correlations of a many-body system?
- compare correlation strength of different systems?

(Quantum) information theory e.g., Ziesche, Smith, Ho, Rudin, Gersdorf, Taut (1999)
Gottlieb, Mauser (2005, 2007)
Classical Information Theory

\( a = \{a_1, a_2, \ldots, a_M\} \) set of random variables (RV), e.g., events, signals, coin tosses, spin directions

\( p(a) \) probability distribution of \( a \)

\( p(a_i) \) probability of event \( a_i \)

\[ I(a_i) := -\ln p(a_i) \geq 0 \]

gain in information (“surprise”) due to \( a_i \)

\[ S(p(a)) := \langle I(a_i) \rangle_{p(a)} = -\sum_i p(a_i) \ln p(a_i) \]

average gain in information

Information (“Shannon”) entropy

Usually \( \ln \rightarrow \log_2 \) (bit)

Postulate of Statistical Physics: \( S(p(a)) \) maximal \( \rightarrow p(a) \)
Distinction between two probability distributions

Two sets of RVs (events, signals, ...):

\[ a = \{a_1, a_2, \ldots, a_M\} \]

\[ p(a) \quad \text{prob. distrib. of } a \]

Example: True/exact/known distrib., or input data

\[ b = \{b_1, b_2, \ldots, b_M\} \]

\[ p(b) \quad \text{prob. distrib. of } b \]

Example: Model/approximate distrib., or output data

How to quantify their (dis)similarity?
Relative entropy ("Kullbach-Leibler distance/divergence/measure") of the prob. distrib. $p(a)$ w.r.t. $p(b)$:

$$\Delta S(p(a)\|p(b)) := \langle I(b_i) - I(a_i) \rangle_{p(a)}$$

$$= \sum_i p(a_i)[\ln p(a_i) - \ln p(b_i)] \begin{cases} > 0, & p(a) \neq p(b) \\ = 0, & p(a) = p(b) \end{cases}$$

Quantifies the “dissimilarity” between the distributions $p(a)$ and $p(b)$.

$$\Delta S(p(a)\|p(b)) \neq \Delta S(p(b)\|p(a)) \quad \rightarrow \text{not a metric or distance}$$

**Example**

- Ising spin without field $H=0$/fair coin: $p(a) = \left(\frac{1}{2}, \frac{1}{2}\right)$

- Ising spin with $H=\infty$/completely unfair coin: $p(b) = (1,0)$

$$\Rightarrow \Delta S(H = 0\|H = \infty) = \frac{1}{2}(\ln \frac{1}{2} - \ln 0) + \frac{1}{2}(\ln \frac{1}{2} - \ln 1) = \infty$$

$$\Delta S(H = \infty\|H = 0) = 1(\ln 1 - \ln \frac{1}{2}) + 0(\ln 0 - \ln \frac{1}{2}) = \ln 2$$
Application: Sanov’s theorem (1957)

How to distinguish between two probability distributions $p(a), p(b)$?

Probability for wrongly identifying $p(b)$ with $p(a)$ after $N$ experiments on one of them:

$$P_N(p(a)|p(b)) \xrightarrow{N \to \infty} e^{-N \Delta S(p(a)\|p(b))}$$

$\neq P_N(p(b)|p(a))$

Example

- Ising spin without field $H=0$/fair coin: $p(a) = \left(\frac{1}{2}, \frac{1}{2}\right)$
- Ising spin with $H=\infty$/completely unfair coin: $p(b) = (1, 0)$

$$\Rightarrow \Delta S(H=0\|H=\infty) = \infty$$

$$\Delta S(H=\infty\|H=0) = \ln 2$$

$$\Rightarrow P_N(H=0|H=\infty) = e^{-N\infty} = 0$$

$$P_N(H=\infty|H=0) = e^{-N\ln 2} = \left(\frac{1}{2}\right)^N$$
Relative entropy as a measure of correlations

Two sets of RVs (events, signals, input/output data):

\[ a = \{a_1, a_2, \ldots, a_M\} \] but only the joint prob. distrib. \( p(a, b) \) is known

\[ b = \{b_1, b_2, \ldots, b_M\} \]

\[ \sum_{i,j} p(a_i, b_j) = 1 \]

Marginal distrib. \( p(a_i) := \sum_j p(a_i, b_j) \)

\[ p(b_j) := \sum_i p(a_i, b_j) \]

Central question of information theory:

How much information does one RV tell about another one ("mutual information")?
Relative entropy as a measure of correlations

Two sets of RVs (events, signals, input/output data):

\[ a = \{a_1, a_2, \ldots, a_M\} \quad \text{but only the joint prob. distrib. } p(a, b) \text{ is known} \]

\[ b = \{b_1, b_2, \ldots, b_M\} \]

Marginal distrib. \( p(a_i) := \sum_j p(a_i, b_j) \)

\[ p(b_j) := \sum_i p(a_i, b_j) \]

In general \( p(a_i, b_j) \neq p(a_i)p(b_j) \) → the RVs are correlated

Example: Roll of a dice

Let \( a=1 \) if number is even (2,4,6), \( a=0 \) otherwise;

\( b=1 \) if number is prime (2,3,5), \( b=0 \) otherwise.

→ Joint distribution of \( a \) and \( b \):

\[
\begin{align*}
P(a=0, b=0) &= P\{1\} = 1/6 \\
P(a=1, b=0) &= P\{4, 6\} = 1/3 \\
P(a=0, b=1) &= P\{3, 5\} = 1/3, \\
P(a=1, b=1) &= P\{2\} = 1/6
\end{align*}
\]

\[
\begin{align*}
p(a_i) := \sum_j p(a_i, b_j) &= 1/6 + 1/3 = 1/2 \\
p(b_j) := \sum_i p(a_i, b_j) &= 1/6 + 1/3 = 1/2
\end{align*}
\]
Relative entropy as a measure of correlations

Two sets of RVs (events, signals, input/output data):

\[ a = \{a_1, a_2, \ldots, a_M\} \] \hspace{1cm} but only the joint prob. distrib. \( p(a,b) \) is known

\[ b = \{b_1, b_2, \ldots, b_M\} \]

Marginal distrib. \( p(a_i) := \sum_j p(a_i,b_j) \)

\[ p(b_j) := \sum_i p(a_i,b_j) \]

In general \( p(a_i,b_j) \neq p(a_i)p(b_j) \) \( \rightarrow \) the RVs are correlated

Mutual information \( I_S(a : b) := S(p(a)) + S(p(b)) - S(p(a,b)) \)

\[ = \Delta S(p(a,b) \| p(a)p(b)) \]

= relative entropy ("dissimilarity")

between \( p(a,b) \) \( \text{(correlated)} \)
and \( p(a)p(b) \) \( \text{(uncorrelated)} \)
Excursion: 
Relation between the relative entropy and susceptibilities in classical equilibrium thermodynamics

Relative entropy \( \Delta S(p||q) = \sum_n p_n(\ln p_n - \ln q_n) \)

\( p_n, q_n \): equilibrium Gibbs probability distributions

\( p_n(x_1, \ldots, x_k) = \frac{1}{Z(x_1, \ldots, x_k)} e^{-\beta E_n(x_1, \ldots, x_k)} = e^{-\beta E_n(x_1, \ldots, x_k) + \beta F(x_1, \ldots, x_k)} \)

\( \Rightarrow \ln p_n = \beta E_n - \beta F \)

\( E_n \): eigen energies, \( F \): free energy depend on external variables \((x_1, \ldots, x_k)\)

Assume: \( q_n \) differs from \( p_n \) only infinitesimally in the \( x_i \): \( q_n(x_i) = p_n(x_i + \delta x_i) \)

Application: \( \Delta S(T|T + \delta T) = \frac{1}{2} C(T) \left( \frac{\delta T}{T} \right)^2 \) Schmalian (private comm., 2012)
More general: Expansion of relative entropy

Byczuk, DV (2012)

$$\Delta S(p(x_1, ..., x_k) || p(x_1 + \delta x_1, ..., x_k + \delta x_k) = -\frac{1}{2} \sum_{ij} g_{ij}(x_1, ..., x_k) \delta x_i \delta x_j$$

where

$$g_{ij}(x_1, ..., x_k) = \sum_n p_n \frac{\partial \ln p_n}{\partial x_i} \frac{\partial \ln p_n}{\partial x_j}$$

$$= \beta^2 \left[ \langle \frac{\partial E_n}{\partial x_i} \frac{\partial E_n}{\partial x_j} \rangle - \langle \frac{\partial E_n}{\partial x_i} \rangle \langle \frac{\partial E_n}{\partial x_j} \rangle \right]$$

thermodynamical averages are taken with respect to $$p_n(x_1, ..., x_k)$$

$$= \beta^2 \left[ \left( \frac{\partial E_n}{\partial x_i} - \langle \frac{\partial E_n}{\partial x_i} \rangle \right) \left( \frac{\partial E_n}{\partial x_j} - \langle \frac{\partial E_n}{\partial x_j} \rangle \right) \right]$$

provides connection between relative entropy and correlation functions of fluctuations

- is a symmetric matrix: “Fisher information matrix” (Fisher, 1922)
- defines a distance in the thermodynamic parameter space (Crooks, 2007)
Thermodynamic susceptibility: 
\[ \chi_{ij}(x_1, \ldots, x_k) := -\beta \frac{\partial^2 F(x_1, \ldots, x_k)}{\partial x_i \partial x_j} \]

\[ -\beta F = \ln p_n - \beta E_n \quad \text{and} \quad g_{ij}(x_1, \ldots, x_k) = \sum_n p_n \frac{\partial \ln p_n}{\partial x_i} \frac{\partial \ln p_n}{\partial x_j} \]

\[ \Rightarrow \quad g_{ij}(x_1, \ldots, x_k) = \chi_{ij}(x_1, \ldots, x_k) + \beta \left\langle \frac{\partial^2 E_n(x_1, \ldots, x_k)}{\partial x_i \partial x_j} \right\rangle \]

if linear response is applicable

\[ \Rightarrow \quad \Delta S(p(x_1, \ldots, x_k) \| p(x_1 + \delta x_1, \ldots, x_k + \delta x_k) = -\frac{1}{2} \sum_{i,j} \chi_{ij}(x_1, \ldots, x_k) \delta x_i \delta x_j \]

Applications: 
\[ \Delta S(T|T + \delta T) = \frac{1}{2} C(T) \left( \frac{\delta T}{T} \right)^2 \]

Schmalian (private comm., 2012)

- phase transitions
- many-body quantum field theory
- finite-time thermodynamics

Prokopenko (2001)

Toyoda (2001)

Niven and Andersen (2009)
From Classical to Quantum Information Theory

Classical probability distribution \( p(a) \)

\[ \Rightarrow \text{statistical operator/density matrix } \hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i| \]

\[ \text{probab. for } |\psi_i\rangle \]

- maximal information about quantum state
- includes classical and quantum correlations (entanglement)

In particular: product state \( \hat{\rho}^{PS} \) (uncorrelated)

Example: Two subsystems (e.g., 2 quantum particles)

\[ \hat{\rho}^{PS} = \hat{\rho}_1 \otimes \hat{\rho}_2 \]

How to compare \( \hat{\rho} \) with \( \hat{\rho}^{PS} \)?
### Information theory

<table>
<thead>
<tr>
<th><strong>classical</strong></th>
<th><strong>quantum</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>• prob. distrib.</td>
<td>• statistical operator ( \hat{\rho} = \sum_i p_i</td>
</tr>
<tr>
<td>• probability of event ( a_i ): ( p(a_i) )</td>
<td>• prob. for state (</td>
</tr>
<tr>
<td></td>
<td>• von Neumann entropy</td>
</tr>
<tr>
<td>Shannon entropy</td>
<td>( S(\hat{\rho}) = -\langle \ln \hat{\rho} \rangle_{\hat{\rho}} = -\text{Tr}(\hat{\rho} \ln \hat{\rho}) )</td>
</tr>
<tr>
<td>[ S(p(a)) := -\langle \ln p(a_i) \rangle_{p(a)} = -\sum_i p(a_i) \ln p(a_i) ]</td>
<td></td>
</tr>
<tr>
<td>Relative entropy</td>
<td>[ \Delta S(\hat{\rho} | \hat{\sigma}) = \text{Tr}\hat{\rho}(\ln \hat{\rho} - \ln \hat{\sigma}) ]</td>
</tr>
<tr>
<td>[ \Delta S(p(a) | p(b)) = \sum_i p(a_i)[\ln p(a_i) - \ln p(b_i)] ]</td>
<td></td>
</tr>
<tr>
<td>Quantifies dissimilarity between ( \hat{\rho} ) and uncorrelated product state ( \hat{\rho}_1 \otimes \hat{\rho}_2 )</td>
<td></td>
</tr>
</tbody>
</table>
Application: Quantum version of Sanov’s theorem (1957)

Hiai, Petz (1991)
Bjelković (2005)

How to distinguish between two states $\hat{\rho}, \hat{\sigma}$ of a quantum system $Q$?

Given $N$ identically prepared copies of $Q$

Probability for wrongly identifying $\hat{\sigma}$ with $\hat{\rho}$ after $N$ experiments:

$$P_N(\hat{\rho} | \hat{\sigma}) \xrightarrow{N \to \infty} e^{-N \Delta S(\hat{\rho} || \hat{\sigma})}$$

$\neq P_N(\hat{\sigma} | \hat{\rho})$

Schumacher, Westmoreland (2000)
Vedral (2002)
Quantification of correlations in quantum many-particle systems

Calculate relative entropy \( \Delta S(\hat{\rho} \parallel \hat{\rho}^{PS}) \)

- Accounts for all possible correlation functions generated by \( \hat{\rho} \)
- Provides unbiased quantification of correlations
- Enables comparison of \( \Delta S \) for different systems
Application to correlated many-electron systems

I. Hubbard model

\[ H = -t \sum_{\langle i,j \rangle, \sigma} c_{i \sigma}^\dagger c_{j \sigma} + U \sum_i n_{i \uparrow} n_{i \downarrow} \]

d=3: Cubic lattice (W=1), NN hopping
Application to correlated many-electron systems

I. Hubbard model

\[ H = -t \sum_{\langle i,j \rangle, \sigma} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \]

\( d=3 \): Cubic lattice \((W=1)\), NN hopping

Not exactly solvable
\( \rightarrow \) employ DMFT to compute relative entropy
\( \rightarrow \) only the local correlations of the exact solution are included

DMFT \( \rightarrow \) reduced statistical operator \( \hat{\rho}_i \)

Site \( i \):

\[ |a\rangle = \{ |0\rangle, |\uparrow\rangle, |\downarrow\rangle, |2\rangle \} \]

\[ \hat{\rho}_i = \text{Tr}_{j \neq i} \hat{\rho} = \sum_a p_a |a\rangle \langle a| \]

\[ p_0 = 1 - n + d \]

\[ p_\sigma = \frac{1}{2} (n + \sigma m) - d \]

\[ p_2 = d \]

Local von Neumann entropy

\[ S(\hat{\rho}_i) = \sum_a p_a \ln p_a \]

Relative entropy

\[ \Delta S(\hat{\rho}_i \parallel \hat{\rho}_i^{PS}) = \sum_a p_a (\ln p_a - \ln p_a^{PS}) \]

Drop index \( i \): \( \hat{\rho}_i \equiv \hat{\rho} \)

Rissler, Noack, White (2006)
Rycerz (2006)
Franca, Capelle (2006)
Larsson, Johannesson (2006)
Paramagnetic ground state $|PM\rangle$

Uncorrelated reference states:

$|HF\rangle = \prod_{k}^{k_F} a_{k\sigma}^{\dagger} |0\rangle$  product state in $k$ space ($U = 0 \Rightarrow d^{HF} = \frac{1}{4}$) (Hartree-Fock)

$|LM\rangle = \prod_{i}^{N} a_{i\sigma_i}^{\dagger} |0\rangle$  product state in position space ($U = \infty \Rightarrow d^{LM} = 0$) (Local Moment)

$n = 1, m = 0 \Rightarrow S(\hat{\rho}) = -2[dlnd + \frac{1}{2} + d)ln(\frac{1}{2} + d)]$ determined by $d(U)$

$S_{HF} = \ln 4$

$S_{LM} = \ln 2$

Correlated state $|PM\rangle$ can be distinguished from $|HF\rangle, |LM\rangle$ by:

$\Delta S(\hat{\rho} \parallel \hat{\rho}^{HF})$

$\Delta S(\hat{\rho}^{HF} \parallel \hat{\rho})$

$\Delta S(\hat{\rho} \parallel \hat{\rho}^{LM}) = \infty$ for $U < \infty$ since $d^{LM} = 0 \Rightarrow$ perfectly distinguishable

$\Delta S(\hat{\rho}^{LM} \parallel \hat{\rho})$
Paramagnetic ground state $|PM\rangle$

Uncorrelated reference states:

$$|HF\rangle = \prod_{k, \sigma}^{k_F} a^\dagger_{k\sigma} |0\rangle$$  
product state in $k$ space  
($U = 0 \Rightarrow d^{HF} = \frac{1}{4}$)

($\text{Hartree-Fock}$)

$$|LM\rangle = \prod_{i} a^\dagger_{i\sigma} |0\rangle$$  
product state in position space  
($U = \infty \Rightarrow d^{LM} = 0$)

($\text{Local Moment}$)

Local von Neumann entropy $S$

$$S^{HF} = \ln 4$$

$$\ln 2 = S^{LM}$$

$$U_{MIT} \approx 1.225$$

Weak increase in Mott phase

$$\Delta S(\hat{\rho} \parallel \hat{\rho}^{HF}) = \ln 4 - S(\hat{\rho})$$
Antiferromagnetic ground state $|AF\rangle$

**Uncorrelated reference states:**

$|Slat\rangle = \prod_{k \in (A,B)} a_{kA}^\dagger a_{kB}^\dagger |0\rangle$  \text{ Slater spin-density product state in k space}

$|Heis\rangle = \prod_{i \in (A,B)} a_{iA}^\dagger a_{iB}^\dagger |0\rangle$  \text{ Neél-type product (“Heisenberg”) state in position space ($m_{stag} = 1, d^{Heis} = 0$)}

**Correlated state $|AF\rangle$** can be distinguished from $|Slat\rangle, |Heis\rangle$ by:

$\Delta S(\hat{\rho} \| \hat{\rho}^{Slat})$

$\Delta S(\hat{\rho}^{Slat} \| \hat{\rho})$

$\Delta S(\hat{\rho} \| \hat{\rho}^{Heis}) = \infty$ \text{ since $d^{Heis} = 0 \Rightarrow$ perfectly distinguishable}

$\Delta S(\hat{\rho}^{Heis} \| \hat{\rho})$
Antiferromagnetic ground state $|AF\rangle$

Uncorrelated reference states:

$$|Slat\rangle = \prod_{k \in (A,B)} a_{k_A}^{\dagger} a_{k_B}^{\dagger} |0\rangle$$  
Slater spin-density product state in $k$ space

$$|Heis\rangle = \prod_{i \in (A,B)} a_{i_A}^{\dagger} a_{i_B}^{\dagger} |0\rangle$$  
Neél-type product ("Heisenberg") state in position space ($m_{stag} = 1, d_{Heis} = 0$)

Correlationwise similar to $|AF\rangle$ for $U \to \infty$
II. Correlations in materials

Example:
Late transition-metal monoxides TMO: TM=Mn, Fe, Co, Ni
II. Correlations in materials

Example:
Late transition-metal monoxides TMO: TM=Mn, Fe, Co, Ni

- Correlated 3d electrons $\rightarrow$ 5x2=10 orbitals
- Projection onto Wannier functions $\rightarrow$
  TM-3d and O-2p Hamiltonian (8-bands)

Kuneš et al. (2007, 2008)

| Correlated state: | $|TMO\rangle$ |
| Uncorrelated reference state: | $|LDA\rangle$ |

$\Delta S(\hat{\rho} \parallel \hat{\rho}^{LDA})$ \quad \text{computed with LDA+DMFT(CT-QMC)}$
Correlations in TMO

Correlation strength decreases from MnO to NiO

$S(\hat{\rho}^{LDA})$: # possibilities to distribute 5,6,7,8 electrons among 10 orbitals → decreases

$S(\hat{\rho})$: strong reduction of # local many-body state due to correlations

Correlation strength decreases from MnO to NiO

Thunström, Di Marco, Eriksson (arXiv:1202.3975)

Electronic entanglement in late transition metal oxides

→ Quantum contribution to the relative entropy of TMO
Effect of hole doping in NiO

Relative entropy decreases with # holes

Hole doping increases # local states

$\nu^8(t^{6}_{2g}e^2_{g}) \rightarrow d^7(t^{6}_{2g}e^1_{g}) \leftrightarrow d^8(t^{6}_{2g}e^2_{g}) \leftrightarrow d^9(t^{6}_{2g}e^3_{g})$

$\rightarrow$ increases local entropy

Relative entropy decreases with # holes

Doping $\rightarrow$ substantial decrease of correlations
Summary

• Unbiased method to quantify correlations in quantum many-particle systems

Compute relative von Neumann entropy \( \Delta S(\hat{\rho} \parallel \hat{\rho}^{PS}) \)

• Applications:

1. Correlations in the Hubbard model

Use DMFT to compute reduced statistical operator \( \hat{\rho}_i \)

\[ \Rightarrow \Delta S(\hat{\rho}_i \parallel \hat{\rho}_i^{PS}) \]

paramagnetic sol.: \( |PS\rangle = |HF\rangle, |LM\rangle \)

antiferromagnetic sol.: \( |PS\rangle = |Slat\rangle, |Heis\rangle \)

2. Correlations in transition metal monoxides

\[ \Rightarrow \Delta S(\hat{\rho}_i \parallel \hat{\rho}_i^{LDA}) \]

Correlations decrease \( \text{MnO} \rightarrow \text{FeO} \rightarrow \text{CoO} \rightarrow \text{NiO} \)

• Perspective: include non-local spatial correlations