Fisher information and density functional theory

Ágnes Nagy

Department of Theoretical Physics, University of Debrecen, Debrecen, Hungary

Correspondence
Ágnes Nagy, Department of Theoretical Physics, University of Debrecen, H-4002 Debrecen, Hungary.
Email: anagy@phys.unideb.hu

Funding information
National Research, Development and Innovation Fund of Hungary, Grant/Award Number: 123988

Abstract
Relationship of Fisher information and density functional theory is reviewed. Links between the Fisher and Shannon information, the local wave-vector and the relative information are displayed. Euler equations for the Fisher and relative Fisher information are presented. A combined information theoretical and thermodynamic view of density functional theory is analyzed. The extremum of the Shannon and Fisher information results a constant temperature and for Coulomb systems a simple relation between the total energy and phase-space Fisher information. Relations for phase-space fidelity, fidelity susceptibility and relative information are also presented.

KEYWORDS
Coulomb systems, density functional theory, Fisher information

1 INTRODUCTION

Fisher information [1] has proved to be very useful among others in physics and chemistry. It has turned out to be particularly valuable in density functional theory (DFT) [2]. Its suitability was first emphasized in the fundamental paper of Sears, Parr and Dinur [3] presenting a relationship between the Fisher information and the quantum mechanical kinetic energy functional. In the last four decades a great number of papers (e.g. [4–17]) analyzed this topic.

The Fisher information can be considered as a functional of the probability density distribution. In DFT the key quantity is the density. That is why Fisher information is so appropriate in DFT. The basic theorem of DFT has the consequence that the density determines the external potential, hence the Hamiltonian, therefore any property of the system. It is interesting to add that there are other quantities with this extraordinary characteristic. For example, the local Shannon information also determines every property of a finite Coulomb system both in the ground and excited states [18].

If the density is known, the Fisher information can also be obtained. It can be done for any state of the system. On the other hand, if the densities of two different states are known, relative Fisher information can be determined.

It has been shown that the Euler equation of DFT can be derived from the principle of extreme physical information. In case of Coulomb systems it can be done for excited states as well. Though there are several ways to treat excited states in DFT (see e.g. [19–27]), here the recent time-independent theory for a single excited state of a Coulomb system [28–30] is considered. Coulomb systems include a very important class of systems as all atoms, molecules or solids belong to it. The Coulomb density has the remarkable property that it determines both the Hamiltonian and the degree of excitation as well. There exists a universal excited-state variational functional for the sum of the kinetic and electron–electron repulsion energies. This functional is relevant for the ground state and all bound excited states. Furthermore, the Euler equation and Kohn-Sham equations have the same form as in the ground-state theory.
The Euler equation of the non-interacting system is

$$\left. \frac{\partial T_{\text{Coul}}}{\partial \mu} \right|_{\mu=\mu_k} + v_{KS}(\phi_k; r) = \mu_k,$$

where $T_{\text{Coul}}$, $v_{KS}$ and $\mu$ are the non-interacting kinetic energy functional, the Kohn-Sham potential and the chemical potential, respectively. $\phi_k$ is the density of the $k$th excited state. In orbital-free DFT, the Euler equation should be solved instead of the Kohn-Sham equations. As there is only one Euler equation, while there are several Kohn-Sham equations for a large system, it is worth using orbital-free DFT provided that adequate approximation for the kinetic energy functional is available. Information theoretical concepts turned to be useful to approximate kinetic energy functionals [13,31-34]. It is worth mentioning that there exists an orbital-free scheme in which the knowledge of the kinetic energy functional is not needed [35,36].

Information theoretical aspects can be combined with a thermodynamic transcription of DFT. Ghosh, Berkowitz and Parr (GBP) [37] proposed thermodynamic interpretation. According to it the ground-state DFT can be considered a local thermodynamics and local temperature can be defined. This fascinating method has found several extensions and applications [38–59]. For example, it is possible to define the local temperature by maximizing the phase-space Shannon information [60,61] or minimizing the phase-space Fisher information [62].

The review is organized as follows: Section 2 summarizes the definition and properties of the Fisher information, the relationship with the Shannon information, the local wave-vector, the relative information and the fidelity. In section 3 the Euler equation and the principle of extreme physical information are studied including the Euler equation for the Fisher information and the relative Fisher information. Section 4 presents a combined information theoretical and thermodynamic view of DFT. Utilizing the Ghosh - Berkowitz - Parr theory and its extension, phase-space fidelity, fidelity susceptibility and Fisher information for Coulomb systems are analyzed. The last section is devoted to discussion.

## 2 | FISHER INFORMATION

Fisher information [1] is a measure of intrinsic accuracy in statistical estimation theory. Consider a normalized probability density function $f(x|\theta)$ that contains a parameter $\theta$. The Fisher informational functional [1] is defined as

$$I(\theta) = \int f(x|\theta) \left[ \frac{\partial \ln f(x|\theta)}{\partial \theta} \right]^2 dx = \int \left[ \frac{f'(x|\theta)}{f(x|\theta)} \right]^2 dx.$$  \hspace{1cm} (2)

If we make a measurement of $x$ we can try to estimate the parameter $\theta$. We look for the best estimate $\hat{\theta} = \hat{\theta}(x)$. Examine a large number of $x$-samples and denote $e^2$ the mean-square error. According to the estimation theory $ke^2 = 1$ for the best possible estimator. For any other estimator the mean-square error is larger. That is, Fisher information provides the lower bound. The Cramer-Rao inequality states

$$ke^2 \geq 1.$$  \hspace{1cm} (3)

There is a family of distribution functions of the form $f(x|\theta) = f(x - \theta)$. This shift invariance has the consequence that

$$I(\theta) = \int \left[ \frac{\partial f(x - \theta)}{\partial (x - \theta)} \right]^2 / f(x - \theta)dx$$  \hspace{1cm} (4)

does not depend on $\theta$. Therefore, the parameter $\theta$ is taken to be zero:

$$I_f = \int \left[ \frac{f'(x)}{f(x)} \right]^2 dx.$$  \hspace{1cm} (5)

For a three variable function $f(r)$ Equation (5) takes the form

$$I_f = \int \left[ \frac{\nabla f(r)}{f(r)} \right]^2 dr.$$  \hspace{1cm} (6)
2.1 | Relationship with Shannon information and the local wave-vector

Two decades after Fisher [1] defined information Equation (2) Shannon [63] introduced a new kind of information

\[ S_f = - \int f(r) \ln f(r) dr. \]  \hspace{1cm} (7)

\( S \) is often expressed with the density \( \varphi \) because \( \varphi \) is the fundamental quantity in DFT.

\[ S = - \int \varphi(r) \ln \varphi(r) dr. \]  \hspace{1cm} (8)

Introducing the shape function \( \sigma(r) = \varphi(r)/N \) the corresponding Shannon information is

\[ S_\sigma = - \int \sigma(r) \ln \sigma(r) dr. \]  \hspace{1cm} (9)

Obviously,

\[ S = NS_\sigma - N \ln N \]  \hspace{1cm} (10)

because \( \varphi \) is normalized to the number of electrons \( N \). Frequently, the Fisher information Equation (6) is also expressed with the density

\[ I = \int \frac{|\nabla \varphi(r)|^2}{\varphi(r)} dr = NL_\sigma. \]  \hspace{1cm} (11)

It is worth introducing the quantities: the specific Shannon information (Shannon information per particle) \( s(r) = - \ln \varphi(r) \) and the specific Fisher information (Fisher information per particle) \( i(r) = (|\nabla \varphi(r)/\varphi(r)|)^2 \) the specific Shannon information (Shannon information per particle) [64].

To establish a relationship between the Fisher and the Shannon information [11,12] it is worth to introduce the local wave-number [65] as the ratio of the density gradient to the electron density

\[ q(r) = - \frac{\nabla \varphi(r)}{\varphi(r)}. \]  \hspace{1cm} (12)

Therefore [10,11,64].

\[ q(r) = \nabla s(r) \]  \hspace{1cm} (13)

and

\[ i(r) = |\nabla s(r)|^2 = q^2. \]  \hspace{1cm} (14)

That is, the local wave-number provides a link between the Shannon and Fisher information. Namely, \( q \) is the gradient of the specific Shannon information, while the square of \( q \) gives the specific Fisher information.

The relationship between the Shannon and Fisher information can be formalized in another way [9]. The local Shannon and Fisher information can be defined as \( \bar{\varphi}_{Shannon}(r) = - \ln \varphi(r) \), \( \varphi_{Fisher}(r) = |\nabla \varphi(r)|^2/\varphi(r) \) and \( i_{Fisher}(r) = - \nabla^2 \varphi(r) \ln \varphi(r) \). Note that \( \bar{\varphi}_{Shannon}(r) = \varphi(r)s(r) \) and \( \varphi_{Fisher}(r) = \varphi(r)i(r) \). \( \varphi_{Fisher} \) and \( \bar{\varphi}_{Fisher} \) are different functions of \( r \), but both integrate to the same Fisher information. The Shannon information can be expressed as [9]:

\[ S = \int \bar{\varphi}_{Shannon}(r) dr = -N + \frac{1}{4\pi} \int \frac{\varphi_{Fisher}(r') - \bar{\varphi}_{Fisher}(r')}{|r-r'|} dr'dr'. \]  \hspace{1cm} (15)

2.2 | Relative information, relative local wave-vector

If we have another (reference) probability density function \( f_{ref}(r) \), we can define the relative Fisher information [66–71] corresponding Equation (6)
\[ J_f = \int f(r) \left| \nabla \left( \ln \frac{f(r)}{f_{ref}(r)} \right) \right|^2 dr \]  

(16)

and the relative Shannon or Kullback–Leibler information (also called cross-entropy) [72].

\[ G_f = \int f(r) \ln \frac{f(r)}{f_{ref}(r)} dr \]  

(17)

as an information gain achieved by using \( f(r) \) instead of \( f_{ref}(r) \). \( J_f \) and \( G_f \) serve to measure the deviation of \( f(r) \) from the reference density \( f_{ref}(r) \). Obviously, \( J_f \geq 0 \) and \( G_f \geq 0 \) with equality if and only if \( f = f_{ref} \) almost everywhere.

The relative Fisher information obtained from the density takes the form

\[ J = \int \rho(r) \left| \nabla \left( \ln \frac{\rho(r)}{\rho_{ref}(r)} \right) \right|^2 dr. \]  

(18)

\( J \) can also be rewritten as

\[ J = \int \rho(r) \left| \nabla \frac{\rho(r)}{\rho_{ref}(r)} - \nabla \frac{\rho_{ref}(r)}{\rho_{ref}(r)} \right|^2 dr. \]  

(19)

On the other hand, the relative Shannon or Kullback–Leibler information constructed from the density has the form

\[ G = \int \rho(r) \ln \frac{\rho(r)}{\rho_{ref}(r)} dr = N G_f. \]  

(20)

\( G \) has been applied in chemistry to quantify electrophilicity, nucleophilicity, regioselectivity [73].

The relative local wave-vector was defined as [74]

\[ \tilde{q}(r) = q(r) - q^{ref}(r), \]  

(21)

where

\[ q^{ref}(r) = - \frac{\nabla \rho^{ref}(r)}{\rho^{ref}(r)}. \]  

(22)

From Equation (13) follows

\[ \tilde{q}(r) = \nabla [s(r) - s^{ref}(r)] = - \nabla \ln \left( \frac{\rho(r)}{\rho^{ref}(r)} \right). \]  

(23)

The relative specific Shannon information can be defined as

\[ \tilde{s}(r) = s(r) - s^{ref}(r) = - \ln \left( \frac{\rho(r)}{\rho^{ref}(r)} \right). \]  

(24)

It is related to the relative local wave-vector

\[ \tilde{q}(r) = q(r) - q^{ref}(r) = \nabla s(r). \]  

(25)

2.3 Fisher information, fidelity, fidelity susceptibility and Kullback–Leibler information

For pure states fidelity is defined by the overlap between the states \( \Psi \) and \( \Phi \) as
In DFT the density is the fundamental quantity, therefore DFT fidelity [75] is defined as the distance between two normalized densities $f$ and $g$:

$$ F(f, g) = \int f^{1/2} g^{1/2}. $$

Fidelity is an important concept of information theory to measure the “closeness” of two quantum states. In quantum chemistry quantum similarity measures have been used for a long time (See e.g. [76]). One of the most popular similarity measures is the Carbó index [77].

$$ R_{AB} = \frac{\int d\rho_A(r) \rho_B(r)}{\sqrt{\int d\rho_A^2(r) \int d\rho_B^2(r)}}. $$

where $\rho_A(r)$ and $\rho_B(r)$ are the molecular electron densities.

Later generalized quantum similarity index (QSI) was defined [78] as

$$ \text{QSI}_q = \frac{\int d\tilde{g}_1 \tilde{g}_2}{\sqrt{\int d\tilde{g}_1^2 \int d\tilde{g}_2^2}}. $$

where $\tilde{g}_1$ and $\tilde{g}_2$ are the distribution functions and $q$ is a real number. Obviously we get back the Carbó index for $q = 2$. The case $q = 1$, on the other hand, gives the fidelity. QSI between Krypton and all neutral atoms with nuclear charge 1–103 for different $q$ in position and momentum spaces were presented [78].

Consider now continuous density distributions depending on some parameter $\theta$. An expansion around $\theta$ leads to the definition of fidelity susceptibility $\chi$ [75]

$$ F(\theta, \theta + \delta\theta) = 1 - \frac{1}{2} (\delta\theta)^2 \chi + \cdots $$

Expanding $f(\theta + \delta\theta)$ around $f(\theta)$ we arrive at

$$ \chi = \frac{1}{2} \left( \frac{f(\theta)}{f(\theta + \delta\theta)} \right)^2. $$

(For convenience only the dependence on the parameter is written in the argument of $f$.) That is, the DFT fidelity susceptibility $\chi$ is proportional to the Fisher information (Equation (2)). In case of shift invariance $\chi$ is proportional to Fisher information of Equation (5) (or Equation (6)).

Following the derivation of Frieden [79] one can find a link between fidelity susceptibility and the relative or Kullback-Leibler information [72]. Rewrite Fisher information as

$$ I = \lim_{\delta\theta \to 0} \left[ \frac{1}{f(\theta)} \left( \frac{f(\theta + \delta\theta) - f(\theta)}{\delta\theta} \right)^2 \right] 
= \lim_{\delta\theta \to 0} \left[ \frac{1}{f(\theta)} \left( \frac{f(\theta + \delta\theta)}{f(\theta)} - 1 \right)^2 \right]. $$

Introducing the small quantity $\nu = f(\theta + \delta\theta)/f(\theta) - 1$ and using the expansion $\ln(1 + \nu) = \nu - \nu^2/2...$, we arrive at

$$ I \approx \frac{2}{(\delta\theta)^2} f(\theta) \ln \frac{f(\theta + \delta\theta)}{f(\theta)}. $$

Therefore, Equation (32) has the form

$$ \chi = -\frac{1}{2(\delta\theta)^2} f(\theta) \ln \frac{f(\theta + \delta\theta)}{f(\theta)}. $$
or

$$\chi = \frac{1}{2(\delta \theta)^2} G( f(\theta), f(\theta + \delta \theta)).$$  \hspace{1cm} (35)$$

That is, the DFT fidelity susceptibility is proportional to the relative information constructed from densities $f(\theta)$ and $f(\theta + \delta \theta)$.

In case of shift invariance [79],

$$I \approx \frac{2}{(\delta x)^2} G( f(x), f(x + \delta x)).$$  \hspace{1cm} (36)$$

From the definition Equation (30) follows a linear relationship between the relative information and DFT fidelity

$$G( f(\theta), f(\theta + \delta \theta)) = 1 - F(\theta, \theta + \delta \theta).$$  \hspace{1cm} (37)$$

### 2.4 Local values of observables and variance

Inequality Equation (3) is related to the variance of observables. To reveal this link it is worth introducing local values of quantum observables. This problem was investigated by several authors [31–33,80–82]. Here the $N$-electron generalization of the formalism of Luo [80] is summarized [12].

Consider a quantum mechanical observable $A$ acting on the $N$-electron wave function $\Psi$. Define a local quantity as

$$A(r) = \bar{A}(r) + i \tilde{A}(r).$$  \hspace{1cm} (38)$$

The real part $\bar{A}(r)$ gives the expectation value

$$\langle A \rangle_\Psi = \int \bar{A}(r) \varrho(r) dr,$$  \hspace{1cm} (39)$$

where $\varrho(r)$ is the electron density. The imaginary part $\tilde{A}(r)$ is related to the variance [12]:

$$\text{Var}_\Psi (A - \bar{A}) = \int \left( \tilde{A}(r) \right)^2 \varrho(r) dr.$$  \hspace{1cm} (40)$$

That is, the real part of the local value of a quantum observable is associated with the expectation value of the quantum observable, while the imaginary part comprises the fluctuation.

Consider the total momentum operator $P$ as a sum of one-electron operators $p_j$

$$P = \sum_{j=1}^{N} p_j.$$  \hspace{1cm} (41)$$

It was shown [12] that the imaginary part is the half of the local wave vector:

$$\tilde{P}(r) = \frac{1}{2} q(r)$$  \hspace{1cm} (42)$$

or the half of the gradient of the local Shannon information per particle:

$$\tilde{P}(r) = \frac{1}{2} \nabla s(r).$$  \hspace{1cm} (43)$$
Therefore the variance of the total momentum can be written as

\[
\text{Var}_{\Psi}(P) = \int \left( \frac{\partial P(r)}{\partial r} \right)^2 \psi(r) dr - \frac{1}{4} \int \psi(r) \left[ \nabla^2 \psi(r) \right]^2 dr = \frac{1}{4} I. 
\]  

(44)

That is, the imaginary part comprises the fluctuation which is proportional to the Fisher information.

3 | EULER EQUATION AND THE PRINCIPLE OF EXTREME PHYSICAL INFORMATION

3.1 | Kinetic energy, Euler equation

The relationship between the quantum mechanical kinetic energy and Fisher information was established by Sears, Parr and Dinur [3].

\[
E_{\text{kin}} = \frac{1}{8} \int \left( \frac{\partial \psi(r)}{\partial r} \right)^2 \psi(r) dr + \frac{1}{8} \int \psi(r) \frac{\partial^2 \psi(r)}{\partial r^2} dr, 
\]  

(45)

where the first term is proportional to the Fisher information and

\[
\frac{\partial^2 \psi(r)}{\partial r^2} \equiv \frac{\partial}{\partial r}, \quad f(2, ..., N | 1) = \frac{\psi^2}{\psi(r)}. 
\]  

(46)

is a Fisher information density associated with the conditional density

\[
f(2, ..., N | 1) = \frac{\psi^2}{\psi(r)}. 
\]  

(47)

\( \Psi \) is the wave function. The first term in Equation (45) is the full Weizsäcker kinetic energy [83].

\[
T_w = \frac{1}{8} \int \left( \frac{\partial \psi(r)}{\partial r} \right)^2 \psi(r) dr. 
\]  

(48)

A comparison of Equations (48) and (11) leads to the relation that the Weizsäcker kinetic energy [83] is proportional to the Fisher information:

\[
T_w = \frac{1}{8} I = \frac{N}{8} I. 
\]  

(49)

The difference of the total and the Weizsäcker kinetic energies

\[
T_p = E_{\text{kin}} - T_w
\]  

(50)

is called Pauli energy [84–87]. As the density is the same in both the interacting and in the non-interacting systems, the Pauli energy can be defined by Equation (50) in both systems with \( E_{\text{kin}} \) denoting the kinetic energy of the given system.

The functional derivative of the Weizsäcker term is

\[
\frac{\delta T_w}{\delta \psi} = \frac{1}{8} \frac{\partial^2 \psi}{\partial r^2} - \frac{1}{4} \frac{\nabla^2 \psi}{\psi} = \psi^{-1/2} \left( \frac{1}{2} \nabla^2 \right) \psi^{1/2}. 
\]  

(51)

The functional derivative of the Pauli energy is called Pauli potential

\[
\nu_p = \frac{\delta T_p}{\delta \psi},
\]  

(52)
$T_p$ and $v_p$ embody all the effects of the Pauli principle (the antisymmetry requirement for the wave function). The Pauli energy has been used to identify strong covalent interactions [88] and appraise the quality of approximate kinetic energy functionals [89].

The Euler-equation of the non-interacting system has the form

$$\frac{\delta T_w}{\delta \varrho} + v_p + v_{KS} = \mu,$$  \hspace{1cm} (53)

where $v_{KS}$ is the Kohn-Sham potential and $\mu$ is the Lagrange multiplier arising from the condition that the number of electrons is kept fixed. The Euler-equation can also be written as a single Schrödinger-like equation for $\varrho^{1/2}$

$$\left[-\frac{1}{2} \nabla^2 + v_p + v_{KS}\right] \varrho^{1/2} = \mu \varrho^{1/2}.$$  \hspace{1cm} (54)

It should be emphasized that the Euler Equation (53) of DFT is also the Euler equation for the Fisher information because the Weizsäcker kinetic energy is proportional to the Fisher information. Equations (49) and (53) lead to [4]

$$\frac{\delta I}{\delta \varrho} + w_p + w_{KS} = \nu,$$  \hspace{1cm} (55)

where

$$w_p = 8v_p,$$  \hspace{1cm} (56)

$$w_{KS} = 8v_{KS},$$  \hspace{1cm} (57)

and

$$\nu = 8\mu.$$  \hspace{1cm} (58)

An Euler-like equation can be derived for the specific Shannon information. First rewrite the Euler Equation (53) using Equations (12)–(14). We can immediately obtain

$$\nabla q = -\nabla^2 (\ln \varrho) = \nabla^2 s = \frac{\nabla \varrho}{\varrho}^2 - \frac{\nabla^2 \varrho}{\varrho} = q^2 - \frac{\nabla^2 \varrho}{\varrho}.$$  \hspace{1cm} (59)

Therefore the functional derivative of the Weizsäcker term in Equation (53) takes the form

$$\frac{1}{8} \frac{\nabla \varrho}{\varrho}^2 - \frac{1}{4} \frac{\nabla^2 \varrho}{\varrho} = -\frac{1}{8} q^2 + \frac{1}{4} \nabla^2 s = -\frac{1}{8} q + \frac{1}{4} \nabla^2 s.$$  \hspace{1cm} (60)

Substituting Equation (60) into Equation (53) we arrive at the Euler-like equation of the specific Shannon information

$$\frac{1}{4} \nabla^2 s - \frac{1}{8} \nabla s^2 + v_p + v_{KS} = \mu.$$  \hspace{1cm} (61)

or

$$\frac{1}{4} \nabla^2 s - \frac{1}{8} l + v_p + v_{KS} = \mu.$$  \hspace{1cm} (62)

Equation (61) has the advantage that it can provide directly the specific Shannon information [64].

### 3.2 Euler equation for the relative information

Suppose we have another (reference) density function $\varrho^{ref}(\mathbf{r})$ and it also satisfies an Euler equation
\[
\frac{\delta T_w[q^{\text{ref}}]}{\delta q^{\text{ref}}} + \sqrt{\nu_{\text{ref}}} + \nu_{\text{ref}}^{KS} = \mu^{\text{ref}},
\]

(63)

where \(\nu_{\text{ref}}^{KS}\) and \(\mu^{\text{ref}}\) are the Kohn-Sham and chemical potentials of the reference system, respectively. Equation (63) is equivalent to the Euler equation of reference Fisher information

\[
\frac{\delta [q^{\text{ref}}]}{\delta q^{\text{ref}}} + \nu_{\text{ref}}^{P} + \nu_{\text{ref}}^{KS} = \nu^{\text{ref}},
\]

(64)

where \(\nu_{\text{ref}}^{P} = 8\nu_{\text{ref}}^{P}, \nu_{\text{ref}}^{KS} = 8\nu_{\text{ref}}^{KS}\) and \(\nu^{\text{ref}} = 8\mu^{\text{ref}}\).

The difference of the Euler Equations (55) and (64) gives [90].

\[
\frac{\delta [J]}{\delta q} + \frac{\delta [q^{\text{ref}}]}{\delta q^{\text{ref}}} + \nu_{\text{ref}}^{P} + \nu_{\text{ref}}^{KS} = \nu,
\]

(65)

where

\(\nu_{\text{ref}}^{P} = w_{\text{ref}}^{P} - w_{\text{ref}}^{P^{\text{ref}}},\)

(66)

\(\nu_{\text{ref}}^{KS} = w_{\text{ref}}^{KS} - w_{\text{ref}}^{KS^{\text{ref}}},\)

(67)

and

\(\nu = \nu - \nu^{\text{ref}}.\)

(68)

Consider now the relative Fisher information \(J\) (Equation (19)) and calculate its functional derivative with respect to \(q\)

\[
\frac{\delta J[q,q^{\text{ref}}]}{\delta q} + \frac{\delta [q]}{\delta q^{\text{ref}}} + \nu_{\text{ref}}^{P} + \nu_{\text{ref}}^{KS} = \nu,
\]

(69)

Comparing Equations (64) and (69) we arrive at Euler equation for the relative information

\[
\frac{\delta J[q,q^{\text{ref}}]}{\delta q} + \nu_{\text{ref}}^{P} + \nu_{\text{ref}}^{KS} = \nu.
\]

(70)

Euler-like equation of the reference specific Shannon information can be written as

\[
\frac{1}{2} \nabla^2 s^{\text{ref}} - \frac{1}{8} |\nabla s^{\text{ref}}|^2 + \nu_{\text{ref}}^{P} + \nu_{\text{ref}}^{KS} = \mu^{\text{ref}}.
\]

(71)

Comparing Equations (61) and (71) we are led to the Euler-like equation for the relative Shannon information

\[
2\nabla^2 \tilde{s}(r) - |\nabla \tilde{s}(r)|^2 - 2\nabla s^{\text{ref}}(r)\nabla \tilde{s}(r) + \nu_{\text{ref}}^{P}(r) + \nu_{\text{ref}}^{KS}(r) = \nu.
\]

(72)

### 3.3 DFT from the principle of extreme physical information

The principle of extreme physical information is a variational principle that can be used to derive major physical laws. According to the principle formalized by Frieden [79] the “physical information” \(K\) of the system is extremum

\[
K = I - J_b = \text{extrem}.
\]

(73)
$I$ is the fixed form of “intrinsic” information defined above (Equations (5) and (6)). $J_b$ is the bound Fisher information that incorporates all constraints imposed by the physical phenomenon under measurement.

The fundamental equations of DFT can be derived using this principle [4,91]. Here the derivation of the time-independent Euler equation is summarized. The minimization is done under the conditions: 1. The wave function is antisymmetric. It was pointed out [92,93] that this requirement can be taken into account by a local potential $u_P(r)$. 2. The density is kept fixed. This constraint can be insured by a local potential $u(r)$ as it is used in the Kohn-Sham system when the potential is determined by requiring that the density be equal to the density of the interacting system. 3. The electron density is normalized to $N$. A Lagrange multiplicator $\chi$ is introduced. Minimization of the Fisher information

$$
\frac{\delta}{\delta \rho(r)} \left[ I[\rho(r)] + \int \rho(r) u_P(r) dr + \int \rho(r) u(r) dr - \chi \left[ \int \rho(r) dr \right] \right] = 0
$$

leads to the Euler–Lagrange equation

$$
\frac{\delta J}{\delta \rho(r)} + u_P(r) + u(r) = \chi. \tag{75}
$$

Comparing Equation (75) with the Euler Equations (53) and (55) we see that they are equivalent and

$$
u_P(r) = w_P(r) = 8v_P(r), \tag{76}
$$

$$u(r) = w_{KS}(r) = 8v_{KS}(r), \tag{77}
$$

$$\chi = \nu = 8\mu. \tag{78}
$$

The principle of extreme physical information can be used to derive the Euler equation for the relative Fisher information, too [90]. Then we have to minimize $J$ with the constraints above.

$$
\frac{\delta}{\delta \rho(r)} \left[ I[\rho(r), \rho^{\text{ref}}(r)] + \int \rho(r) \tilde{u}_P(r) dr + \int \rho(r) \tilde{u}(r) dr - \tilde{\chi} \left[ \int \rho(r) dr \right] \right] = 0. \tag{79}
$$

The variation leads to the Euler Equation (70) with

$$
\tilde{u}_P(r) = \tilde{w}_P(r), \tag{80}
$$

$$\tilde{u}(r) = \tilde{w}_{KS}(r) \tag{81}
$$

and

$$\tilde{\chi} = \tilde{\nu}. \tag{82}
$$

From Equation (79) the Euler-like Equation (72) of the relative Shannon information can be immediately derived.

The time-independent Kohn-Sham Equations [91] and the time-dependent Euler equation of the density functional theory were also derived using the principle of extreme physical information [4]. Moreover, the time-dependent Euler equation for the pair density has also recently been derived utilizing this principle [94]. It is worth mentioning that maximizing the relative Shannon entropy has also turned to be very useful for several chemical application [95,96].

## 4 | COMBINED INFORMATION THEORETICAL AND THERMODYNAMIC VIEW OF DFT

### 4.1 | Ghosh–Berkowitz–Parr theory

A combined information theoretical and thermodynamic interpretation of DFT was proposed by Ghosh, Berkowitz and Parr (GBP) [37]. It is summarized and extended in this section. Consider an electronic system with density $\rho$ that integrates to the number of electrons $N$: 
We search for a phase-space distribution function $f(r,p)$ satisfying the conditions

$$\int dp f(r,p) = \varrho(r)$$  \hspace{1cm} (84)$$

and

$$\int dp \frac{p^2}{2} f(r,p) = t(r).$$  \hspace{1cm} (85)$$

The kinetic energy density $t(r)$ integrates to the kinetic energy

$$E_{\text{kin}} = \int dr t(r).$$  \hspace{1cm} (86)$$

The phase-space distribution function can be determined by maximizing the information

$$S = -k \int dr dp (\ln f - 1).$$  \hspace{1cm} (87)$$

with keeping the density (Equation (84)) and the kinetic energy density (Equation (85)) fixed. The variation leads to a Maxwell-Boltzmann-like distribution function

$$f(r,p) = e^{-\alpha(r)} e^{-\beta(r)p^2/2},$$  \hspace{1cm} (88)$$

where $k$ is the Boltzmann constant. $\alpha(r)$ and $\beta(r)$ are $r$-dependent Lagrange multipliers arising from satisfying conditions (84) and (85). The distribution function Equation (88) provides the kinetic energy density

$$t(r) = \frac{3}{2} \frac{\varrho(r)}{\beta(r)}$$  \hspace{1cm} (89)$$

reminiscing the familiar ideal gas expression. Note that Equation (89) is the direct consequence of the fact that the kinetic energy density is kept fixed. Because of this analogy the $r$-dependent $\beta(r)$ is called local inverse temperature. $f$ can be recast in the form

$$f(r,p) = \left[ \frac{2\pi}{\beta(r)} \right]^{-3/2} \varrho(r) e^{-\beta(r)p^2/2}.$$  \hspace{1cm} (90)$$

The phase-space information Equation (87) with the maximizing distribution function Equation (90) leads to the Sackur–Tetrode expression

$$S = -k \int \varrho(r) \ln \varrho(r) dr + \frac{1}{2} k \int \varrho(r) [3 \ln (2\pi/\beta) + 5] dr.$$  \hspace{1cm} (91)$$

4.2 | Extremum of Shannon and Fisher information

It is wellknown that the kinetic energy density $t(r)$ is not uniquely defined. The kinetic energy is unique and does not change if we add a term that integrates to zero, to the kinetic energy density. But, of course, $\beta(r)$ and consequently, $f$ will be different. There are several forms for the kinetic energy density (e. g. [97–99]) and one can apply the one which is more suitable for one's purpose. It has recently been proposed to take the kinetic energy density for which phase-space Fisher/Shannon information is minimum/maximum [60,62]. We seek that particular kinetic energy density that maximizes $S$:
with keeping the orbitals and the density fixed. We arrive at
\[ \frac{3k_B}{2T} = \zeta, \]  
(93)

that is,
\[ \zeta = \beta. \]  
(94)

The Lagrange multiplier \( \zeta \) equals the inverse temperature, i.e., \( \beta \) is constant:
\[ \beta = \frac{3N}{2E_{\text{kin}}}. \]  
(95)

The same result can be obtained if the phase-space Fisher information is minimized. Take the normalized phase-space distribution function
\[ g(r, p | \beta) = \frac{1}{Nf(r, p)} \]
and construct the phase-space Fisher information
\[ I_g(\beta) = \int \left( \frac{\partial g(r, p | \beta)}{\partial \beta} \right)^2 g(r, p | \beta) dr dp. \]  
(97)

Observe that the parameter \( \theta \) in Equation (2) is the \( r \)-dependent inverse temperature \( \beta(r) \). Substituting Equation (96) into Equation (97) and integrating for \( p \), we get
\[ I_g(\beta) = \frac{3}{2N} \int \frac{\phi(r)}{(\beta(r))^2} dr. \]  
(98)

The minimization of \( I_g \) with fixed kinetic energy
\[ \frac{3}{2N} \int \frac{\phi(r)}{(\beta(r))^2} dr + \zeta \left( E_{\text{kin}} - \frac{3}{2} \frac{\phi(r)}{\beta(r)} dr \right). \]  
(99)

leads to
\[ C\beta = \frac{2}{N_0}, \]  
(100)

that is, to a constant inverse temperature. It has the consequence that the kinetic energy density is proportional to the electron density and there is a simple relation between the Fisher information and the kinetic energy:
\[ I_g = \frac{2}{3} \left( \frac{E_{\text{kin}}}{N} \right)^2. \]  
(101)
4.3 Fisher information for Coulomb systems

The GBP transcription was proposed for the Kohn-Sham scheme and applied for the non-interacting kinetic energy. But we can observe that in the derivation there was no reference to the non-interacting system, and the derivation is valid also if the true interacting kinetic energy is employed [100]. The extremum of the Shannon and Fisher information can also be studied both for the non-interacting and the interacting kinetic energies. If we consider Coulomb systems we can obtain further important relations. In equilibrium nuclear geometry the virial theorem has the form

\[ E = -E^{\text{kin}}, \]  

(102)

where \( E \) and \( E^{\text{kin}} \) are the true (interacting) total and kinetic energies, respectively. Therefore, Equation (101) can be rewritten as

\[ I_s = \frac{2}{3} \left( \frac{E}{N} \right)^2 \]  

(103)

or we can express the total energy with the Fisher information

\[ E = -N \sqrt{\frac{3}{2} I_s}. \]  

(104)

Having a look again on the derivation above we can also observe that it is valid for excited states as well. One only needs the density and the kinetic energy density, they do not have to be ground-state density and kinetic energy density. Moreover, the virial theorem is also valid for excited states. Therefore, Equations (103) and (104) can be written for an excited state. Then the \( i \)th excitation energy can be given

\[ E_i - E_0 = N \sqrt{\frac{3}{2} \left( \sqrt{I_s^0} - \sqrt{I_s^i} \right)}. \]  

(105)

where \( E_0, I_s^0 \) and \( E_i, I_s^i \) are total energy and Fisher information of the ground and the \( i \)th excited states.

4.4 Phase-space fidelity, fidelity susceptibility and Fisher information for Coulomb systems

Based on Equation (27) the phase-spase fidelity can be written with the phase-spase density functions \( g(r, p) \) and \( g^\text{ref}(r, p) \) as

\[ F(g, g^\text{ref}) = \int g^{1/2}(g^\text{ref})^{1/2} drdp. \]  

(106)

Using Equation (96) we obtain [100]

\[ F(g, g^\text{ref}) = \frac{1}{N} \int dr \varrho^{1/2} \left( \varrho^\text{ref} \right)^{1/2} \left( \frac{2 \sqrt{\beta \varrho^\text{ref}}}{\sqrt{\beta + \varrho^\text{ref}}} \right)^{3/2} \]  

(107)

For constant inverse temperatures \( F \) has the form

\[ F(g, g^\text{ref}) = \left( \frac{2 \sqrt{\beta \varrho^\text{ref}}}{\sqrt{\beta + \varrho^\text{ref}}} \right)^{3/2} F(\varrho, \varrho^\text{ref}). \]  

(108)

where

\[ F(\varrho, \varrho^\text{ref}) = \frac{1}{N} \varrho^{1/2} (\varrho^\text{ref})^{1/2} dr \]  

(109)
is the position-space fidelity between the states with densities \( \varrho \) and \( \varrho_{\text{ref}} \). That is, the phase-space fidelity is proportional to the position-space fidelity, where the factor of proportionality depends on the inverse temperatures. Instead of the inverse temperatures one can use the total energies, too [100]

\[
F(g, g^\text{eff}) = \left( \frac{2 \sqrt{\|E\| E^\text{eff}}}{\|E\| + \|E^\text{eff}\|} \right)^{3/2} F(\varrho, \varrho^\text{eff}). \tag{110}
\]

Based on Equation (17) the relative or Kullback–Leibler information defined as

\[
G(g, g^\text{eff}) = \int g \ln \frac{g}{g^\text{eff}} \tag{111}
\]

measures the “distance” between the phase-space density functions \( g \) and \( g^\text{eff} \). From Equations (96) and (111) follows

\[
G(g, g^\text{eff}) = \frac{1}{N} \int d\varrho \ln \frac{\varrho}{\varrho^\text{eff}} + \frac{3}{2N} \int d\varrho \left[ \frac{\varrho^\text{eff}}{\beta} - \ln \frac{\varrho^\text{eff}}{\beta} - 1 \right]. \tag{112}
\]

For constant inverse temperatures

\[
G(g, g^\text{eff}) = G(\varrho, \varrho^\text{eff}) + c [\beta, \beta^\text{eff}], \tag{113}
\]

where

\[
G(\varrho, \varrho^\text{eff}) = \frac{1}{N} \int d\varrho \ln \frac{\varrho}{\varrho^\text{eff}} \tag{114}
\]

and

\[
c[\beta, \beta^\text{eff}] = \frac{3}{2} \left[ \frac{\beta}{\beta^\text{eff}} - \ln \frac{\beta}{\beta^\text{eff}} - 1 \right]. \tag{115}
\]

That is, the phase-space relative information \( G(g, g^\text{eff}) \) is equal to the position-space Kullback–Leibler information \( G(\varrho, \varrho^\text{eff}) \) plus a term depending on the inverse temperatures. \( c \) can also be expressed with the total energies [100].

The phase-space Fisher information can be rewritten as

\[
I_g = \lim_{\delta \beta \to 0} \frac{1}{g(\beta)} \left( \frac{g(\beta + \delta \beta) - g(\beta)}{\delta \beta} \right)^2 \approx -\frac{2}{(\delta \beta)^2} \left[ g(\beta) \ln \frac{g(\beta + \delta \beta)}{g(\beta)} \right] \tag{116}
\]

or

\[
\chi = \frac{1}{2(\delta \beta)^2} G(g(\beta), g(\beta + \delta \beta)). \tag{117}
\]

That is, the Fisher information, the fidelity susceptibility and the Kullback–Leibler information are proportional quantities. There is a linear relationship between the Kullback–Leibler information and the fidelity [100]

\[
G(g(\beta), g(\beta + \delta \beta)) = 1 - F(g(\beta), g(\beta + \delta \beta)) \tag{118}
\]

for “close” states.


5 | DISCUSSION

Instead of the relative information $J_f$ (Equation (16)) and $G_f$ (Equation (17)) symmetrized forms called Fisher divergence and Jensen-Shannon divergence can also be used [66,67]. These divergences were studied for different families of probability distributions. They provide relevant information on the atomic shell structure both in position and momentum spaces.

Fisher information has found several interesting applications in DFT. A stimulating utilization is related to the steric effect. The steric effect is a widely applied qualitative concept of chemistry: it indicates that an atom occupies a certain amount of space in a molecule. A precise definition came to light in DFT by Liu [84]. According to it the steric term $E_{\text{steric}}$ equals the Weizsäcker kinetic energy:

$$E_{\text{steric}} = T_w = \frac{1}{8} I_f.$$  \hspace{1cm} (119)

It means that the steric term is proportional to the Fisher information. The Weizsacker kinetic energy has also been employed to predict stereoselectivity [101,102].

The Euler equation presented in several forms in Section 3.1 and 3.3 is valid for the ground state and if Coulomb systems are considered even for excited states. The Euler equation for the relative information in Section 3.2 and 3.3 can be gained from any two density functions and Euler-like equation can be derived for relative Shannon information with any appropriate reference density. So relative Shannon information can be studied for example between an excited state and the ground state or two different excited states or two (excited or ground) states of different systems. Specific relative Shannon and Fisher information of Hydrogen atom was analyzed in [90].

In Section 4 phase-space distribution functions are considered. Phase-space Fisher information has interesting properties [103]. These distribution functions are supposed to be nonnegative and produce the correct marginal distribution functions [104–107]. It means that integrating $f$ with respect to the momentum (position) coordinates we are led to the position (momentum) density. We mention in passing that the Wigner function [108] has the correct marginals, but it is not everywhere nonnegative. Moreover, Wigner showed that bilinear distribution functions are not universally nonnegative. As Cohen and Zaparovanny [109] proved there are distribution functions that are nonnegative, provide the correct marginals, but they are not bilinear. In the Ghosh-Berkowitz-Parr approach only the position-space marginal of the distribution function is correct. The momentum-space marginal is different from the correct one as the correct kinetic energy density is constrained instead.

The choice of the constant temperature is especially appealing because in Coulomb systems the GBP phase-space distribution function is a function of the density. It is the consequence of the fact that the density decays as

$$\lim_{r \to \infty} \frac{\partial}{\partial r} \ln \varrho(r) = -\sqrt{8(E_0^{N-1} - E)},$$  \hspace{1cm} (120)

where $E_0^{N-1} - E$ is the vertical ionization potential of the $N$-electron system. $E$ is the energy of the state considered and $E_0^{N-1}$ is the ground-state energy of the $N - 1$ electron system [28–30]. Thus, the asymptotic decay of the density decides the energy $E$ and $\beta$ via the virial theorem. Therefore, the density determines the phase-space distribution function $g$ (Equation (96)). Obviously, the density dictates the phase-space Fisher information $I_g$ (Equation (103)), too.

In Section 3 the position-space Fisher information with translational invariance is applied. In Section 4, on the other hand, the phase-space Fisher information ($I_g$) with parameter of $\beta$ is utilized. It is easy to derive a simple relation between these quantities. From Equation (50) follows that

$$T_w \leq E_{\text{kin}}.$$  \hspace{1cm} (121)

Then using Equations (49) and (101) we arrive at the inequality between the two kinds of Fisher information

$$I_f \leq 96 I_g.$$  \hspace{1cm} (122)

There is equality for one-level systems. The simplest example for it is the Hydrogen atom. In the ground state the kinetic energy is $E_{\text{kin}} = 1/2$ (in atomic units). Equation (49) gives $I_\sigma = 4$, while Equation (101) provides $I_g = 1/6$. That is, we are led to equality in relation (122).

Another advantage of the GBP phase-space distribution function is that phase-space fidelity is a product of the position-space fidelity and a term depending on the energies (Equation (110)). It should be emphasized that the normalized phase-space distribution functions $g$ and $g^{\text{ref}}$ are not necessarily associated to the same system. These functions can belong to any Coulomb system and the phase-space fidelity can uncover similarity of different systems.
In Coulomb system the phase-space fidelity and Kullback–Leibler information are also determined by the density, that is, the (position-space) density comprises phase-space information. Earlier, position and momentum-space Fisher information were generally analyzed separately (e.g. [110–112]). The phase-space fidelity and Kullback–Leibler information include information of both spaces and can provide a combined measure of resemblance.

Fidelity and fidelity susceptibility have been applied in several fields, e.g. in quantum phase transitions [113], topological phase transitions [114,115]. The link between relative information and fidelity susceptibility [116,117] has also been revealed and illustrated with the quantum phase transition.

Chemical reactions have turned to be an important field of application of information theoretical concepts. Earlier reactivity, selectivity descriptors are mainly used in the ground-state theory, though these descriptors are often linked to excitability. (Even the first excitation energy can be used as a reactivity index [118].) Excited-state reactivity is now a new frontier [59,119,120]. A recent paper studies chemical selectivity through well selected excited states [121]. It is expected that Fisher information and related quantities presented above would turn to be a valuable tool to investigate excited-state reactivity descriptors.

ACKNOWLEDGMENTS
This research was supported by the National Research, Development and Innovation Fund of Hungary, financed under 123988 funding scheme.

AUTHOR CONTRIBUTIONS
Agnes Nagy: Conceptualization; formal analysis; investigation; writing-original draft; writing-review & editing.

DATA AVAILABILITY STATEMENT
Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

REFERENCES
[1] R. A. Fisher, Proc. Camb. Philos. Soc. 1925, 22, 700.
[2] P. Hohenberg, W. Kohn, Phys. Rev. 1964, 136, B864.
[3] S. B. Sears, R. G. Parr, U. Dinur, Israel J. Chem. 1980, 19, 165.
[4] Á. Nagy, J. Chem. Phys. 2003, 119, 9401.
[5] E. Romera, P. Sánchez-Moreno, J. S. Dehesa, Chem. Phys. Lett. 2005, 414, 468.
[6] Á. Nagy, Chem. Phys. Lett. 2006, 425, 154.
[7] E. Romera, J. S. Dehesa, J. Chem. Phys. 2004, 120, 8906.
[8] E. Romera, Mol. Phys. 2002, 100, 3325.
[9] S. B. Liu, J. Chem. Phys. 2007, 126, 191107.
[10] Á. Nagy, Chem. Phys. Lett. 2007, 449, 212.
[11] Á. Nagy, S. B. Liu, Phys. Lett. A 2008, 372, 1654.
[12] Á. Nagy, E. Romera, S. B. Liu, Phys. Lett. A 2013, 377, 286.
[13] L. M. Ghiringhelli, L. Delle Site, R. A. Mosna, L. P. Hamilton, J. Math. Chem. 2010, 48, 78.
[14] Á. Nagy, K. D. Sen, Acta Phys. Debrec. 2011, 45, 105.
[15] Á. Nagy, E. Romera, Chem. Phys. Lett. 2010, 490, 242.
[16] D. Chakraborty, P. W. Ayers, J. Math. Chem. 2011, 49, 1810.
[17] Á. Nagy, E. Romera, Chem. Phys. Lett. 2014, 597, 139.
[18] Á. Nagy, Chem. Phys. Lett. 2013, 556, 355.
[19] H. Appel, E. K. U. Gross, K. Burke, Phys. Rev. Lett. 2010, 90, 043005.
[20] M. E. Casida, J. Mol. Struct. Theochem. 2009, 914, 3.
[21] M. Petersilka, U. J. Gossmann, E. K. U. Gross, Phys. Rev. Lett 1996, 76, 1212.
[22] A. K. Theophilou, J. Phys. C 1998, 12, 5419.
[23] L. N. Oliveira, E. K. U. Gross, W. Kohn, Phys. Rev. A 1998, 37, 2805, 2809, 2821.
[24] M. Levy, Á. Nagy, Phys. Rev. Lett. 1999, 83, 4631.
[25] Á. Nagy, M. Levy, Phys. Rev. A 2001, 63, 052502.
[26] Á. Nagy, Int. J. Quantum Chem. 1998, 69, 247.
[27] Á. Nagy, Int. J. Quantum Chem. 1998, 70, 681.
[28] P. W. Ayers, M. Levy, Á. Nagy, Phys. Rev. A 2012, 85, 042518.
[29] P. W. Ayers, M. Levy, Á. Nagy, J. Chem. Phys 2015, 143, 191101.
[30] P. W. Ayers, M. Levy, Á. Nagy, Theor. Chim. Acc 2018, 137, 152.
[31] R. A. Mosna, I. P. Hamilton, L. Delle Site, J. Phys. A 2005, 38, 3869. https://doi.org/10.1088/0305-4470/38/17/011. Accessed 25 Apr. 2021
[32] I. P. Hamilton, R. A. Mosna, L. Delle Site, Theor. Chim. Acta 2007, 118, 407.
[33] I. P. Hamilton, R. A. Mosna, J. Comp. Appl. Math. 2010, 233, 1542.
AUTHOR BIOGRAPHY

Ágnes Nagy is Professor Emeritus at Department of Theoretical Physics of University Debrecen, Hungary. Her main research interest is density functional theory, especially extensions to excited states. She has also been working on information theoretical concepts, including Fisher and Rényi information and applications in density functional theory. Professor Nagy is member of the Editorial Board of International Journal of Quantum Chemistry, the Organizing Committee of the International Conferences on Density Functional Theory and its Applications and the Board of Directors of International Society of Theoretical Chemical Physics.

How to cite this article: Nagy Á. Fisher information and density functional theory. Int J Quantum Chem. 2021:e26679. https://doi.org/10.1002/qua.26679