Quantum information and phase transitions: fidelity and state distinguishability

V R Vieira
Centro de Física das Interacções Fundamentais
and Departamento de Física
Instituto Superior Técnico
Universidade Técnica de Lisboa
Lisboa, Portugal
E-mail: vrv@cfif.ist.utl.pt

Abstract. We review the definition and properties of fidelity as a measure of distinguishability of quantum states, presenting some of its geometric aspects.

We apply then fidelity to the study of phase transitions in condensed matter systems, both at zero and finite temperatures. It does not rely on any already known order parameter, being useful not only in the usual Landau type phase transitions but also in phase transitions displaying topological order.

The so called fidelity susceptibility and its relation to the usual thermodynamic susceptibilities and response functions are discussed.

Finally, we consider the fidelity between different reduced density matrices for a subsystem of a larger system.

1. Introduction

In the last years, there have been very important developments in an interdisciplinary approach involving different fields of physics and quantum information and computation [1].

Probably the most well known of these developments involves entanglement. Entanglement was considered by Schrödinger, already in 1935, at the time of the Einstein-Podolsky-Rosen paradox, not one, but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. Physicists became used to it, in particular in the study of the correlation functions of physical systems, and of their subsystems, but only recently entanglement was taken seriously as an important resource for quantum information science, in particular in quantum computation, communication and cryptography [2].

Important steps have also been given in quantum computation, both theoretically, with the development of quantum algorithms, and experimentally, with the development of new physical systems and tools, regarding the applications.

Quantum mechanics is a probabilistic theory, allowing us to make predictions on the possible outcomes of experiments performed on physical systems, which have been previously prepared. It is then natural to expect that the techniques of estimation and the methods of statistical inference used in information theory would be very valuable in addressing the questions and problems posed by quantum mechanics, not only at a fundamental level, but also in its applications to different areas in physics, namely condensed matter physics.
We will consider the application of fidelity, a measure of distinguishability between probability distributions, to the distinguishability between quantum states of a system and to the study of phase transitions in condensed matter physics, stressing some of its geometrical features.

2. Classical Fidelity

Classically, the fidelity between two probability distributions \( p_i \) and \( q_i \) is defined by

\[
F(p, q) = \sum_i \sqrt{p_i q_i}
\]  

(sometimes, in the literature, fidelity is defined as the square of this quantity). It can be understood as the inner product of the two unit vectors \( \sqrt{p} \) and \( \sqrt{q} \), of components \( \sqrt{p_i} \) and \( \sqrt{q_i} \) respectively. One has \( 0 \leq F(p, q) \leq 1 \), with \( F(p, q) = 1 \), only when the two probability distributions are the same.

Fidelity itself is not a distance, but can be used to define distances. The Bures angle is defined by

\[
D_A(p, q) = \arccos F(p, q) = \theta.
\]

Geometrically, it represents the distance between the two vectors, measured along the sphere of unit length. The Bures distance is defined by

\[
D_B(p, q) = \sqrt{2(1 - F(p, q))} = 2 \sin \frac{\theta}{2}
\]

and represents the length of the chord between the two points in the sphere. Finally, the so called sine distance is defined by

\[
S_B(p, q) = \sqrt{1 - F^2(p, q)} = \sin \theta.
\]

If the two probability distributions are infinitesimally close, these distances become the same

\[
ds^2 = \sum_i (d\sqrt{p_i})^2 = \frac{1}{4} \sum_i \frac{dp_i^2}{p_i},
\]

which is the Fisher metric (in the case of several parameters \( \lambda^\mu \) one has \( dp_i = \sum_\lambda d\lambda^\mu \frac{\partial p_i}{\partial \lambda^\mu} \)). This parameter dependence induces a metric in parameter space from the metric in distribution space which we are considering.

Sometimes, it is convenient to consider non-normalized distributions \( \vec{r} \) and then normalize them, according to \( \sqrt{p} = \frac{\vec{r}}{r} \). From \( dr = \frac{\vec{r}}{r} \cdot d\vec{r} \), one has

\[
d\sqrt{p} = \frac{d\vec{r}}{r} = \frac{d\vec{r}}{r} - \frac{\vec{r}}{r} \left( \frac{\vec{r}}{r} \cdot d\vec{r} \right),
\]

removing the component of \( d\vec{r} \) along \( \vec{r} \), such that \( \vec{r} \cdot d\vec{r}_\perp = 0 \). The inner product between two deviations \( d\sqrt{p}_1 \) and \( d\sqrt{p}_2 \), from a probability distribution \( \sqrt{p} \), is then given by

\[
d\sqrt{p}_2 \cdot d\sqrt{p}_1 = \frac{d\vec{r}_2}{r} \cdot \frac{d\vec{r}_1}{r} = \frac{d\vec{r}_2 \cdot d\vec{r}_1}{r^2} - \frac{(d\vec{r}_2 \cdot \vec{r})(d\vec{r}_1 \cdot \vec{r})}{r^4},
\]

which is the Fubini-Study metric. For a single variation, it reduces to

\[
D^2_{FS}(d\vec{r}) = d\sqrt{p} = \left( \frac{d\vec{r}}{r} \right)^2 = \left( \frac{d\vec{r}}{r} \right)^2 - \left( \frac{\vec{r} \cdot d\vec{r}}{r^2} \right)^2.
\]
One sees then the Fisher metric as the metric of the unit sphere, also described by the Fubini-Study metric.

The Fisher metric is associated to the Fisher information, i.e. the variance of ln $p$, given by

$$\langle (\delta \ln p)^2 \rangle = \sum_i p_i \langle (\delta \ln p_i)^2 \rangle = \sum_i \frac{(\delta p_i)^2}{p_i}$$

(9)

$$= \sum_i \delta p_i \delta \ln p_i = -\langle \delta^2 \ln p \rangle,$$

(10)

since its expectation $\langle \delta \ln p \rangle = \sum_i p_i \delta \ln p_i = \delta \sum_i p_i = 0$.

In information theory, the Fisher information is used in connection to the Cramér-Rao inequality to set a lower bound on the variance of estimators [3]. Using the Cauchy-Schwarz inequality $\sqrt{\text{var}(T)\text{var}(V)} \geq |\text{cov}(V,T)|$, with $V = \frac{d\ln p}{d\lambda}$, var$(V)$ is the Fisher information and

$$\text{cov}(V,T) = -\frac{d\ln p}{d\lambda} |V,T| = \sum_i \frac{dp_i}{d\lambda} T.$$

The fidelity between two probability distributions described by $\vec{r}_1 = r + \lambda d\vec{r}_1$ and $\vec{r}_2 = r + \lambda d\vec{r}_2$ is given by

$$F = 1 - \frac{1}{2} D^2_{FS}(d\vec{r}_1 - d\vec{r}_2) + O(\lambda^2),$$

(11)

i.e. the deviation of the fidelity from 1 is, in lowest order, the square of the Fubini-Study metric of the difference of the variations. One can vary either of them or both. If the variations are the same, the fidelity stays equal to 1.

The derivation of eq. 11 is the same as for the quantum case, treated in the next section.

3. Quantum Fidelity

We will define the fidelity between two states of a quantum system considering first the case of pure states, next the case where one of the states is a mixed state but the other is still a pure state, and finally, the case of two mixed states.

3.1. Pure States

If the two states of the quantum system are pure, being specified by the normalized wave functions $|\varphi_1\rangle$ and $|\varphi_2\rangle$ (with the states given by the respective projectors), the fidelity between them is defined by

$$F(|\varphi_1\rangle\langle\varphi_1|, |\varphi_2\rangle\langle\varphi_2|) = |\langle\varphi_2|\varphi_1\rangle|,$$

(12)

i.e. is given by the modulus of the overlap, in agreement to its physical interpretation as the probability of the transition between those two states.

In general, we prefer to work with a non normalized wave function $|\psi\rangle$. The normalized wave function is then given by $|\varphi\rangle = \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}}$. Its phase can be changed by a gauge transformation, one of the basic invariances of quantum mechanics, transforming a wave $|\psi\rangle$ into $|\psi'\rangle = e^{i\chi}|\psi\rangle$, giving $|d\psi'\rangle = e^{i\chi}(|d\psi\rangle + id\chi|\psi\rangle)$. Projectors are phase independent and therefore gauge invariant. Only relative phases are relevant and the physical results should be invariant under gauge transformations.

The variation of $|\varphi\rangle$ is given by

$$|d\varphi\rangle = \frac{|d\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} - \frac{1}{2} \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} \langle d\psi|\psi\rangle + \langle \psi|d\psi \rangle,$$

(13)

One does not have $\langle \psi|d\varphi \rangle = 0$, since the normalization condition of the wave function implies $\langle d\varphi|\psi\rangle + \langle \psi|d\varphi \rangle = 0$, instead. However, one can make the decomposition

$$|d\varphi\rangle = |d\varphi\rangle_\parallel + |d\varphi\rangle_\perp$$

(14)
with
\[ |d\varphi\rangle_\perp = \frac{|d\psi\rangle - |\psi\rangle \langle \psi|d\psi\rangle}{\sqrt{\langle \psi|\psi \rangle}} = \left( I - \frac{|\psi\rangle \langle \psi|}{\langle \psi|\psi \rangle} \right) \frac{|d\psi\rangle}{\sqrt{\langle \psi|\psi \rangle}} \] (15)
\[ |d\varphi\rangle_\parallel = - \frac{1}{2} \frac{\langle \psi|d\psi\rangle - \langle \psi|d\psi\rangle |\psi\rangle}{\sqrt{\langle \psi|\psi \rangle}} = i\alpha |\varphi\rangle, \] (16)

with \( \langle \psi|d\varphi\rangle = 0 \) and \( \langle \varphi|d\varphi\rangle = i\alpha \), where \( \alpha = \frac{1}{2} \frac{\langle \psi|d\psi\rangle - \langle \psi|d\psi\rangle}{\langle \psi|\psi \rangle} \) is real, in agreement with the normalization condition \( \langle d\varphi|\varphi \rangle + \langle \varphi|d\varphi \rangle = 0 \), for \( |\varphi\rangle \).

Due to the elimination of the component along \( |\psi\rangle \), the variation \( |d\varphi\rangle_\perp \) is gauge invariant.

The overlap between two deviations \( |d\varphi_1\rangle \) and \( |d\varphi_2\rangle \) from a normalized wave function \( |\varphi\rangle \) is given by
\[ \langle d\varphi_2|d\varphi_1 \rangle_\perp = \langle d\varphi_2|d\varphi_1 \rangle_\parallel + \|d\varphi_2|d\varphi_1 \rangle \] (17)
\[ \langle d\varphi_2|d\varphi_1 \rangle_\perp = \frac{\langle d\psi_2|d\psi_1 \rangle}{\langle \psi|\psi \rangle} - \frac{\langle d\psi_2|\psi \rangle \langle \psi|d\psi_1 \rangle}{\langle \psi|\psi \rangle} \] (18)
\[ \|d\varphi_2|d\varphi_1 \rangle_\parallel = \frac{1}{4} \left( \frac{\langle \psi|d\psi_2 \rangle - \langle d\psi_2|\psi \rangle}{\langle \psi|\psi \rangle} - \frac{\langle \psi|d\psi_1 \rangle - \langle d\psi_1|\psi \rangle}{\langle \psi|\psi \rangle} \right) \] (19)

with the second term real and symmetric in \( |d\varphi_1\rangle \) and \( |d\varphi_2\rangle \).

Separating the real and imaginary parts we have \( \langle d\varphi_2|d\varphi_1 \rangle = g(|d\varphi_1\rangle, |d\varphi_2\rangle) + ih(|d\varphi_1\rangle, |d\varphi_2\rangle) \). The real part is symmetric in its arguments and the imaginary part is antisymmetric. They are given by
\[ g(|d\varphi_1\rangle, |d\varphi_2\rangle) = \frac{1}{2} \frac{\langle d\psi_2|d\psi_1 \rangle + \langle d\psi_1|d\psi_2 \rangle}{\langle \psi|\psi \rangle} - \frac{1}{4} \frac{\langle \psi|d\psi_2 \rangle + \langle \psi|d\psi_1 \rangle - \langle d\psi_2|\psi \rangle - \langle d\psi_1|\psi \rangle}{\langle \psi|\psi \rangle} \] (20)
\[ h(|d\varphi_1\rangle, |d\varphi_2\rangle) = \frac{1}{2} \frac{\langle d\psi_2|d\psi_1 \rangle - \langle d\psi_1|d\psi_2 \rangle}{\langle \psi|\psi \rangle} - \frac{1}{2} \frac{\langle \psi|d\psi_2 \rangle - \langle d\psi_2|\psi \rangle - \langle \psi|d\psi_1 \rangle - \langle d\psi_1|\psi \rangle}{\langle \psi|\psi \rangle^2} \] (21)

In differential geometry, \( g(|d\varphi_1\rangle, |d\varphi_2\rangle) \) defines the Riemann metric and \( h(|d\varphi_1\rangle, |d\varphi_2\rangle) \) the connection, leading to the Berry phase.

For a single variation \( |d\varphi\rangle \), the overlap \( \langle d\varphi|d\varphi \rangle \) reduces to
\[ \langle d\varphi|d\varphi \rangle = \frac{\langle d\psi|d\psi \rangle}{\langle \psi|\psi \rangle} - \frac{\langle d\psi|\psi \rangle \langle \psi|d\psi \rangle}{\langle \psi|\psi \rangle^2} + \left( \frac{i}{2} \frac{\langle d\psi|\psi \rangle - \langle \psi|d\psi \rangle}{\langle \psi|\psi \rangle} \right)^2. \] (22)

The first term is the Fubini-Study metric
\[ D^2_{FS}(|d\psi\rangle) = \langle d\varphi|d\varphi \rangle_\perp = \frac{\langle d\psi|d\psi \rangle}{\langle \psi|\psi \rangle} - \frac{\langle d\psi|\psi \rangle \langle \psi|d\psi \rangle}{\langle \psi|\psi \rangle^2} \] (23)
and the second term is real and positive. The condition of parallel transportation is therefore the one which minimizes $\langle d\varphi | d\varphi \rangle$, rendering it also gauge invariant and leading to the Kubo-Study metric.

Some of the peculiarities and subtleties of quantum mechanics, namely those of the so called wave function renormalization, come from the fact that its fundamental metric is, in fact, this metric.

The overlap $\langle d\varphi | d\varphi \rangle$ can be also rewritten as

$$\langle d\varphi | d\varphi \rangle = \left( \frac{\langle d\psi | d\psi \rangle}{\langle \psi | \psi \rangle} - \left( \frac{1}{2} \frac{\langle d\psi | d\psi \rangle + \langle \psi | d\psi \rangle}{\langle \psi | \psi \rangle} \right) \right)^2.$$  \hspace{1cm} (24)

It is interesting to take $|\psi\rangle$ itself already normalized, i.e. that $\langle \psi | \psi \rangle = 1$, and compare what happens to eqs. 22 and 24.

Given that normalized wave functions $|\varphi\rangle$ and projectors $P_{\varphi} = |\varphi\rangle \langle \varphi |$, satisfying the condition $Tr P_{\varphi}^2 = Tr P_{\varphi} = 1$, are directly related, it is instructive to look also at the variation of a projector. Either directly or using eqs. 14 and 16 we find

$$dP_{\varphi} = \left( |d\varphi\rangle_\perp + i\alpha |\varphi\rangle \right)\langle \varphi | + c.c.$$  \hspace{1cm} (25)

involving only the perpendicular term, the one which is gauge invariant.

We establish now the relation between the fidelity and the Kubo-Study metric.

Taking $|\psi_1\rangle = |\psi\rangle + |d\psi_1\rangle \lambda$ and $|\psi_2\rangle = |\psi\rangle + |d\psi_2\rangle \lambda$, with $\lambda$ real, and defining $\alpha_i = \frac{\langle \psi_1 | d\psi_i \rangle}{\langle \psi_\psi \rangle}$ and $\beta_{ij} = \frac{\langle \psi_2 | d\psi_i \rangle}{\langle \psi_2 | \psi \rangle}$ one has

$$F^2 = \frac{\langle \psi_2 | \psi_1 \rangle \langle \psi_1 | \psi_2 \rangle}{\langle \psi_2 | \psi_2 \rangle \langle \psi_1 | \psi_1 \rangle}$$

$$= \frac{1 + (\alpha_1 + \alpha_2^*) \lambda + \beta_{21} \lambda^2}{1 + (\alpha_2 + \alpha_1^*) \lambda + \beta_{12} \lambda^2}$$

$$= 1 - \frac{(\beta_{11} - \beta_{12} - \beta_{21} + \beta_{22}) + (\alpha_1^* - \alpha_2^*)(\alpha_1 - \alpha_2)}{2} \lambda^2 + O(\lambda^3)$$

$$= 1 - D_{FS}^2 \lambda^2 + O(\lambda^3)$$  \hspace{1cm} (26)

or, equivalently, $F \simeq 1 - \frac{1}{2} D_{FS}^2 \lambda^2$, where

$$D_{FS}^2 = \frac{(\langle d\psi_1 | - \langle d\psi_2 \rangle | d\psi_1 \rangle - | d\psi_2 \rangle)}{\langle \psi | \psi \rangle}$$  \hspace{1cm} (27)

is the square of the Kubo-Study metric for $|d\psi_1\rangle - |d\psi_2\rangle$, i.e. $D_{FS}^2(|d\psi_1\rangle - |d\psi_2\rangle)$.

The deviation from 1 of the fidelity between those two pure states is, in lowest order, the square of the Kubo-Study metric of the difference of the variations.

The proof of eq. 11 is similar, but with all quantities real.

3.2. One Mixed State
In this section we motivate physically the definition of fidelity between a pure and a mixed state. In the next section, we will give the general definition, retrieving the cases where at least one of the states is pure, as particular cases.

Since the square of the fidelity between pure states is $F^2(|\varphi_1\rangle \langle \varphi_1 |, |\varphi_2\rangle \langle \varphi_2 |) = |\langle \varphi_2 | \varphi_1 \rangle|^2 = |\langle \varphi_1 | \varphi_2 \rangle|^2 = |\langle \varphi_1 | \varphi_2 \rangle \langle \varphi_2 | \varphi_1 \rangle|$ and recalling that the state of a system in a statistical mixture of pure states
with probabilities $p_i$, is $\rho_2 = \sum_i |\varphi_{2,i}\rangle p_i \langle \varphi_{2,i}|$, it is natural to expect that, averaging with the probabilities $p_i$, one should have $F^2(\langle \varphi |\varphi \rangle, \rho) = \langle \varphi |\varphi \rangle$, leading to
\[ F(\langle \varphi |\varphi \rangle, \rho) = \sqrt{\langle \varphi |\varphi \rangle}. \tag{28} \]

The fidelity between a pure and a mixed state is then the square root of the matrix element of the density matrix on the pure state.

### 3.3. Mixed States

The case of two mixed states can be treated using the concept of purification of a mixed state. Every mixed state $\rho$ in a Hilbert space $H$ can be viewed as the reduced state of some pure state in a larger Hilbert space. It is always possible to find a second Hilbert space $H'$ such that $\rho = Tr_{H'} |\varphi \rangle \langle \varphi |$. The fidelity between two mixed states $\rho_1$ and $\rho_2$ is defined using the fidelity between their purifications $|\varphi_1\rangle$ and $|\varphi_2\rangle$. Since the process of purification is not unique, due to arbitrariness in the choice of the phases, one looks for the maximum over all possible purifications. The fidelity is then defined as [6]
\[ F(\rho_1, \rho_2) = \max_{|\varphi_1\rangle, |\varphi_2\rangle} |\langle \varphi_2 |\varphi_1 \rangle|. \tag{29} \]

In general, it is sufficient to use for $H'$ a copy of the Hilbert space itself. Since $\rho$ is Hermitian the set of its eigenvectors $|e_i\rangle$, with eigenvalues $p_i$, is a basis of $H$. One chooses then a basis of vectors $|f_i\rangle$ of $H'$, not necessarily the same. The state $|\varphi\rangle = \sum_i \sqrt{p_i} |e_i\rangle \otimes |f_i\rangle = (\sqrt{\rho} \otimes I) |\Psi\rangle$, where $|\Psi\rangle = \sum_i |e_i\rangle \otimes |f_i\rangle$, is a purification of the state $\rho$. It is already in the Schmidt decomposition form, defined in the Appendix.

This approach, with its characteristic duplication of the degrees of freedom, has been applied also in other areas of physics [7, 8]. In particular, it is the basis of the so called Thermo Field Dynamics [8, 9] (where a “thermal state” is defined, to treat equilibrium quantum statistical mechanical systems by field theoretical methods using Wick’s theorem as in the zero temperature formalism), similar to the Keldish formalism.

When one or both states are pure, the general definition leads to those already given. From its definition, several properties of the fidelity between mixed states can be derived: it is symmetric in its arguments, $0 \leq F(\rho_1, \rho_2) \leq 1$, reaching its maximum value only when the two states are the same, etc.

In general [6, 10], given two mixed states $\rho_1$ and $\rho_2$, we can have the purifications $|\varphi_1\rangle = \sum_i \sqrt{p_1} |e_i\rangle \otimes |f_i\rangle$ and $|\varphi_2\rangle = \sum_i \sqrt{p_2} |\tilde{e}_i\rangle \otimes |g_i\rangle$. The different basis are related by arbitrary unitary transformations $|\tilde{e}_i\rangle = V |e_i\rangle$, $|f_i\rangle = U_1 |e_i\rangle$ and $|g_i\rangle = U_2 |\tilde{e}_i\rangle$. Using $\sum_i A |e_i\rangle \otimes |e_i\rangle = \sum_i |e_i\rangle \otimes A^T |e_i\rangle$, which can be verified evaluating the matrix elements, one has
\[ |\varphi_1\rangle = \sum_i \sqrt{p_1} |e_i\rangle \otimes U_1 |e_i\rangle = \sum_i \sqrt{p_1} U_1^T |e_i\rangle \otimes |e_i\rangle \]
\[ |\varphi_2\rangle = \sum_i \sqrt{p_2} U_2^T |\tilde{e}_i\rangle \otimes |\tilde{e}_i\rangle = \sum_i \sqrt{p_2} U_2^T V V^T |e_i\rangle \otimes |e_i\rangle. \]

The inner product of the two purifications is
\[ \langle \varphi_2 |\varphi_1 \rangle = \sum_i (e_i^T (U_2^T V V^T)^\dagger) \sqrt{p_2 \sqrt{p_1}} U_1^T |e_i\rangle. \tag{30} \]

Using the left polar decomposition (defined in the Appendix) of an operator to write $\sqrt{p_2 \sqrt{p_1}} = U \sqrt{p_2 \sqrt{p_1}}$, where the modulus of the operator is given by $|\sqrt{p_2 \sqrt{p_1}}|^2 =
We must solve
\[
\langle \varphi_2 | \varphi_1 \rangle = Tr \left( U_T^2 V V^T \right)^\dagger U |\sqrt{p_2} \sqrt{p_1}| U^T
\]
\[
= Tr |\sqrt{p_2} \sqrt{p_1}| U^T \left( U_T^2 V V^T \right)^\dagger U .
\]  (31)

Since \( |Tr |A|U| \leq Tr |A| \), for \( U \) unitary, with equality only when \( U = I \), as can be easily seen using the basis in which \( |A| \) is diagonal, we finally arrive at the general expression for the fidelity
\[
F(p_1, p_2) = Tr |\sqrt{p_2} \sqrt{p_1}| = Tr \sqrt{p_1 p_2} .
\]  (32)

As already stated, it reduces to eq. 12, in the case of two pure states, and to eq. 28, in the case of a pure and a mixed state.

Two different situations may occur: the commutative and the non commutative cases. In the commutative case, the two states commute, having therefore a common basis of eigenvectors \( |\varphi_i| \). Both states can be diagonalized simultaneously, being described by the probabilities \( p_{1,i} \) and \( p_{2,i} \). In this case, the fidelity reduces to the classical fidelity expression
\[
F(p_1, p_2) = \sum_i \sqrt{p_{1,i} p_{2,i}}
\]  (33)

for these probability distributions.

In the non commutative case not only the eigenvalues change but also the eigenvectors themselves change, rotating in the Hilbert space. In general, using eqs. 6 and 25, we have
\[
d\sqrt{\rho} = \sum_i |\varphi_i| \frac{d r_{1,i}}{r} \langle \varphi_i | + \sum_i \frac{r_{1,i}}{r} (|d \varphi_i\rangle_\perp \langle \varphi_i | + |\varphi_i\rangle \langle d \varphi_i|)
\]  (34)

\[
d\rho = \sum_i |\varphi_i| dp_i \langle \varphi_i | + \sum_i p_i (|d \varphi_i\rangle_\perp \langle \varphi_i | + |\varphi_i\rangle \langle d \varphi_i|)
\]  (35)

for the variation of a mixed state. The first terms are purely classical, whereas the second terms are a quantum feature.

In the case of mixed states it is possible to define different metrics, with different properties and drawbacks [11]. We will define the metric directly from the fidelity as \( F = 1 - \frac{1}{2} ds^2 + \cdots \), finding the so called Bures metric [12, 6]. The maximization condition in the definition of the fidelity is comparable to the transversality condition in the case of pure states.

For simplicity, we take \( p_1 = \rho \) fixed and \( p_2 = \rho + dp + \frac{1}{2} d^2 \rho + \cdots \), with \( Tr dp = Tr d^2 \rho = 0 \). We must solve \( \sqrt{\rho_1 \rho_2} \sqrt{p_1} = \rho + X + \frac{1}{2} Y + \cdots \), with \( X \) first order and \( Y \) second order, for \( X \) and \( Y \). One finds \( \rho X + X \rho = \sqrt{\rho} dp \sqrt{\rho} \) and \( \rho Y + 2(d \rho)^2 + Y \rho = \sqrt{\rho} (d^2 \rho) \sqrt{\rho} \), leading, on the basis in which \( \rho \) is diagonal, to \( X_{ij} = \frac{|d \rho_{ij}|}{p_{i} + p_{j}} \) and \( Y_{ii} = -\frac{1}{p_{i}} (dp_{i})^2 + d^2 \rho_{ii} \). Finally, one has \( Tr X = \frac{1}{4} Tr dp = 0 \) and
\[
ds^2 = -Tr Y = \frac{1}{2} \sum_{ij} \frac{|dp_{ij}|^2}{p_{i} + p_{j}}
\]  (36)

which is the Bures metric [13].

In the discussion of the Bures metric and of its parallel transport condition it is convenient [5] to rewrite \( ds^2 \) in terms of \( G \), defined by
\[
dp = \rho G + G \rho.
\]  (37)
as
\[
ds^2 = \frac{1}{2} Tr dG .
\]  (38)
4. Fidelity and Phase Transitions

Phase Transitions is one of the most challenging areas in Condensed Matter Physics.

Traditionally, the phase transitions considered were the so called classical or thermal phase transitions, i.e. at finite temperatures, driven by thermal fluctuations [14]. Typically, in the thermodynamic limit, the partition function or some of its derivatives become non analytic, displaying singularities at some points. Phase transitions and critical phenomena where one of the topics driving the development of the renormalization group [15].

More recently, there has been a great interest in the so called quantum phase transitions, i.e. at zero temperatures, driven by quantum fluctuations [16]. Typically, the ground state (or some low lying excited state) of the system changes qualitatively with the variation of some parameter in which the Hamiltonian of the system depends.

Most phase transitions are associated with a symmetry breaking order parameter and some type of Landau-Ginzburg-Wilson free energy is normally used. However, there are some phase transitions which do not fit into the Landau-Ginzburg-Wilson paradigm of phase transitions, as the Beresinskii-Kosterlitz-Thouless phase transition [17, 18, 19] and the so called topological phase transitions [20], with applications in topological quantum computation.

The physical properties of the system change drastically across the boundary between two different phases [21, 22]. The fidelity between states should then decrease noticeably when the states approach that boundary and in particular if those states belong to different sites of it. Fidelity simply compares different states of a system, without being necessary any knowledge about some predefined order parameter. It is then a very natural tool to use and can be considered as an order parameter by itself.

In the next sections we see how the ideas presented above can be immediately applied to the study of phase transitions and have been applied to the study of several models.

4.1. Zero Temperature

At zero temperature, the system is usually in its ground state, which depends on the parameters defining the system Hamiltonian, like interaction strength, magnetic or electric fields, etc, defining the phase space.

Since we are dealing with pure states, the fidelity is simply the modulus of the overlap between their wave functions.

The behavior of the ground state and of the lowest energy levels are fundamental for the understanding of the phase transitions and of the fidelity.

4.2. Finite Temperatures

In quantum statistical mechanics two ensembles are usually considered: the canonical and the grand canonical ensembles, where the system is coupled to an environment exchanging only energy or energy and particles with the system (in this context the zero temperature case, just considered, is called the micro canonical ensemble). In the Appendix we summarize the derivation of these ensembles from information theory, using the Boltzmann-Gibbs-Shannon-von Neumann entropy.

The density matrix for the canonical ensemble is

$$\rho = \frac{e^{-\beta H}}{Z}$$  \hspace{1cm} (39)

where $\beta = \frac{1}{k_B T}$ is the inverse of the absolute temperature $T$ ($k_B$ being the Boltzmann constant) and $Z = Tr e^{-\beta H}$ is the partition function. The density matrix for the grand canonical ensemble is

$$\rho = \frac{e^{-\beta (H - \mu N)}}{Z}$$  \hspace{1cm} (40)
with the partition function given by \( Z = Tr \, e^{-\beta (H - \mu N)} \).

In these cases, the general formula for the fidelity must be used. Temperature is an additional parameter to be considered in the definition of the phase space.

In the commutative case, the fidelity, expressed in terms of partition functions, is

\[
F(\rho_1, \rho_2) = \frac{Z}{\sqrt{Z_1Z_2}}
\]

where \( Z_1 \) and \( Z_2 \) are the partition functions for the two states and \( Z \) is the partition function obtained replacing \( \beta H \) by the average \( \frac{1}{Z} (\beta_1 (H_1 - \mu_1 N) + \beta_2 (H_2 - \mu_2 N)) \), reducing to \( \beta (H - \mu N) \) with \( \beta = \frac{1}{Z} (\beta_1 + \beta_2) \), for a variation of the temperature, or \( \beta \mathcal{H} \), with \( \mathcal{H} = \mathcal{H}_0 + \mathcal{H}' \), with \( \mathcal{H}_1 = \frac{1}{2} (\lambda_1 + \lambda_2) \), if the Hamiltonian can be divided into commuting parts \( [\mathcal{H}_0, \mathcal{H}'] = 0 \).

As in the zero temperature case level crossings are very important for the behavior of phase transitions and of the fidelity. These level crossings can occur already for the finite systems or only in the thermodynamic limit.

In some applications, it is interesting to consider a small part of a larger system. One is then led to consider the reduced density matrix of that subsystem and the fidelity between two such reduced density matrices [23]. Even if the large system is in the ground state in general the small system will be in a mixed state.

5. Fidelity Susceptibility

Fidelity depends on two density matrices, i.e. on two points of phase space. In practice, one fixes, at a small value, the difference \( \Delta \lambda = \lambda_2 - \lambda_1 \) of the parameters defining them, and varies their center of gravity \( \lambda = \frac{1}{2} (\lambda_2 + \lambda_1) \), exploring the phase space, in particular the boundaries between phases.

It is then natural to consider the so called fidelity susceptibility [24, 25] defined by \( F = 1 - \frac{1}{2} \chi_F (\Delta \lambda)^2 + \cdots \), for zero temperature, or by \( \ln F = -\frac{1}{2} \chi_F (\Delta \lambda)^2 + \cdots \), for finite temperatures, which depends on a single point in phase space only.

The derivation we have done for eq. 26, using independent variations of the non normalized wave functions cannot be used, and must be redone, using the normalized wave functions instead. In second order of the (variation) of the driving parameter we have \( |\varphi(\lambda)\rangle = |\varphi\rangle + |\varphi\rangle' \lambda + \frac{1}{2} (\varphi'' \lambda^2 + \cdots \rangle \). These derivatives must satisfy constraints required by the normalization condition \( \langle \varphi | \varphi \rangle = 1 \) given by \( \langle \varphi' | \varphi \rangle + \langle \varphi | \varphi' \rangle = 0 \) and \( \langle \varphi'' | \varphi \rangle + 2 \langle \varphi' | \varphi' \rangle + \langle \varphi | \varphi'' \rangle = 0 \). We have then \( \langle \varphi_2(\lambda) | \varphi_1(\lambda) \rangle = (1 + (\varphi_2' | \varphi_2') (1 + (\varphi_1' | \varphi_1') \lambda + \frac{1}{2} ((\varphi_2'' | \varphi_2') + 2 \langle \varphi_2' | \varphi_1' \rangle + \langle \varphi | \varphi_1'' \rangle) \lambda^2 + \cdots \rangle \). Using those constraints, the square of the fidelity is

\[
F^2 = \langle \varphi_2(\lambda) | \varphi_1(\lambda) \rangle \langle \varphi_1(\lambda) | \varphi_2(\lambda) \rangle
= 1 - (\langle \varphi_1' | - \langle \varphi_2' \rangle (1 - \langle \varphi | \varphi \rangle) (\langle \varphi_1' \rangle - \langle \varphi_2' \rangle) \rangle \lambda^2 + \cdots \rangle
\approx 1 - \rho \rho' \lambda^2,
\]

with \( \rho \rho' \) given by eq. 15. We find again the Fubini-Study metric, due to the presence of the perpendicular projector \( 1 - |\varphi \rangle \langle \varphi | \). This result can be more simply understood noticing that \( F^2 = Tr \, P_{\varphi_2} P_{\varphi_1} \) and using eq. 25 and the constraints resulting from \( Tr \, P_{\varphi}^2 = 1 \) by differentiation, i.e. \( Tr \, P_{\varphi} dP_{\varphi} = 0 \)

\[
Tr \, P_{\varphi} dP_{\varphi} = -Tr \, dP_{\varphi}^2 = -Tr \, (dP_{\varphi})^2.
\]

From the Schrödinger equation \( (H(\lambda) - E(\lambda)) |\varphi(\lambda)\rangle = 0 \), it follows immediately that \( (dH(\lambda) - dE(\lambda) |\varphi(\lambda)\rangle + (H(\lambda) - E(\lambda)) d|\varphi(\lambda)\rangle = 0 \). One finds then the Hellman-Feynman theorem \( dE = \langle \varphi | dH(\varphi) \rangle \) and that \( d|\varphi\rangle = \frac{1}{E - H} (dH - dE) |\varphi\rangle \), where the resolvent is the pseudo-inverse (obtained inverting the non-zero eigenvalues and doing nothing do those which are zero), or, equivalently introducing explicitly the orthogonal projector, as usually done in the
time independent perturbation theory. Applying this result to the ground state, the fidelity susceptibility is

\[ \chi_F = \frac{1}{\beta} \left( \langle d\varphi |d\varphi \rangle \right)_\perp = \langle 0 | (d\mathcal{H} - dE) \frac{1}{(E_0 - \mathcal{H})^2} (d\mathcal{H} - dE) | 0 \rangle. \] (42)

Inserting a decomposition of the identity (with the ground state excluded) one finally has

\[ \chi_F = \sum_{m \neq 0} \frac{d\mathcal{H}_{mn} d\mathcal{H}_{n0}}{(E_0 - E_n)^2}, \] (43)

where \( d\mathcal{H}_{mn} = \langle m |d\mathcal{H}|n \rangle \).

The matrix element structure resembles that of the second order perturbation theory expression for the energy, but with the energy difference denominator squared, suggesting that the fidelity should be a more sensitive quantity for the detection of phase transitions, diverging faster [26].

For finite temperatures, in the commutative case, the fidelity susceptibility coincides with the usual thermodynamic susceptibilities or response functions [24, 25]. Writing \( F = e^{-\frac{1}{2} \chi_F (\Delta \lambda)^2} \), the divergence of the susceptibility is clearly associated to a drop of the fidelity at the phase transitions, in agreement with our intuition.

In the Appendix we have showed that \( d \ln Z = -E d\beta + N d(\beta \mu) - \beta (\frac{d\mathcal{H}}{d\lambda}) d\lambda \). From \( \frac{\partial \ln Z}{\partial \beta} = -E \), we have \( \chi_F = \beta^2 \left( \langle \mathcal{H} - \langle \mathcal{H} \rangle \rangle^2 \right) = k_B T^2 \frac{\partial E}{\partial T} = k_B T^2 C_v \), proportional to the specific heat (in the canonical ensemble), for a variation in temperature. Similarly, we have \( \chi_F = \beta^2 \left( \langle \mathcal{H}' - \langle \mathcal{H}' \rangle \rangle^2 \right) = \beta^2 \chi_F \), proportional to the appropriate susceptibility, for a variation of some other parameter, in the commutative case. If \( \lambda \) is a magnetic field and \( \mathcal{H}' \) is a spin component, \( \chi_F \) is the usual longitudinal magnetic susceptibility.

From eq. 35 the matrix elements of \( dp \), in the basis where \( \rho \) is diagonal, are \( \langle \varphi_i |dp|\varphi_j \rangle = dp \delta_{ij} + \rho_i \langle \varphi_i |d\varphi_i \rangle + \rho_j \langle \varphi_j |d\varphi_j \rangle = dp \delta_{ij} + (p_i - p_j) \langle d\varphi_i |\varphi_j \rangle \). The Bures metric becomes then

\[ ds^2 = \frac{1}{4} \sum_i (dp_i)^2 + \frac{1}{2} \sum_{i \neq j} |\langle d\varphi_i |\varphi_j \rangle|^2 \frac{(p_i - p_j)^2}{p_i + p_j}, \] (44)

called, in the literature [27], the classical and the non classical terms, respectively. The non classical terms lead to the so called Uhlmann phase [28], as also numerically found [25].

In the commutative case, the Bures metric reduces to the Fisher metric, given by the first term. Since the second term is always non negative, the Fisher metric is a lower bound of the Bures metric, in general.

Alternatively to this fidelity susceptibility approach, one may also look for the singularities of the Riemann metric tensor [29].

### 6. Application to Specific Models

Fidelity has been applied with considerable success to the different types of thermal, quantum and topological phase transitions.

Applications have involved spin, electronic or bosonic systems, both for finite systems and in the thermodynamic limit, using the solutions of exactly solvable models (typically in one dimension models), exact diagonalization of finite systems, mean field approximations, renormalization group analysis, finite scaling, density matrix renormalization group, etc.

For a recent, complete and exhaustive review, see [30], and references therein.
7. Conclusions
We have reviewed the notion of fidelity as a measure of distinguishability of probability distributions in information theory and its extension to the quantum states, stressing some of its geometric features. We have also showed how to apply it to the study of phase transitions in condensed matter systems.

We have tried to keep the notation as simple as possible, but pointing out the connection to differential and information geometry, at the same time.

We have tried to organize the presentation in a pedagogical manner and we hope that this review will be useful as an introduction to more advanced topics such as the study of the geometric Berry and Uhlmann phases.

Acknowledgments
Valuable discussions with Nikola Paunković and Pedro Ribeiro are acknowledged. This work was supported by FCT Project PTDC/FIS/64926/2006, Portugal.

Some mathematical results
Matrix singular value decomposition
Given a matrix $A$ (of dimension $(m \times n)$, not necessarily square) the matrices $A^\dagger A$ and $AA^\dagger$ are Hermitian and therefore each of them have a complete set of orthogonal eigenvectors, with real non negative eigenvalues. Having obtained the eigenvectors $|\alpha_i\rangle$ with eigenvalues $\lambda_i$, of $A^\dagger A$, the vectors $|\beta_i\rangle$, defined by $\sqrt{\lambda_i}|\beta_i\rangle = A|\alpha_i\rangle$, for $\lambda_i \neq 0$, are eigenvectors of $AA^\dagger$, also with eigenvalue $\lambda_i$. If necessary, the other eigenvectors of $AA^\dagger$ can be found, completing the orthogonal basis. Some attention must be paid to the zero eigenvalues, in particular when the matrix is not square. Acting with $A$ on the decomposition of the identity, written as $I = \sum_i |\alpha_i\rangle\langle\alpha_i|$, we obtain $A = \sum_i |\beta_i\rangle\sqrt{\lambda_i}\langle\alpha_i|$. Its matrix elements can be written in the form $\langle b | A | a \rangle = \sum_i (\beta_i)_i \sqrt{\lambda_i} \langle \alpha_i | a \rangle$, i.e. we can write $A$ in the form $V^\dagger \Lambda U$, with $U$, $V$ unitary and $\Lambda$, of dimension $(m \times n)$, with diagonal elements only.

Matrix polar decompositions
Defining $J = \sum_i |\alpha_i\rangle\sqrt{\lambda_i}\langle\alpha_i| = \sqrt{A^\dagger A}$, $K = \sum_i |\beta_i\rangle\sqrt{\lambda_i}\langle\beta_i| = \sqrt{AA^\dagger}$ and $U = \sum_i |\beta_i\rangle\langle\alpha_i|$ we have $A = UJ = KU$. Note that $Tr \sqrt{A^\dagger A} = Tr \sqrt{AA^\dagger}$.

When $A$ is square, $U$ is unitary and $A = UJ$ and $A = KU$ are called the left and right polar decompositions of $A$, respectively. This is a generalization of the usual polar decomposition of a complex number into modulus and phase.

Schmidt decomposition
A state of a composite system $|\psi\rangle$ can be written using the basis states $|e_i\rangle$ $|f_j\rangle$ of the Hilbert spaces $H_A$ and $H_B$, as $|\psi\rangle = \sum_{ij} c_{ij} |e_i\rangle \otimes |f_j\rangle$. Using the singular value decomposition we can write $c_{ij} = \sum_k \alpha_{ik} l_k \beta_{kj}$, with $l_k$ positive. Defining $|a_k\rangle = \sum_i |e_i\rangle \alpha_{ik}$ and $|b_k\rangle = \sum_j |f_j\rangle \beta_{kj}$, we find $|\psi\rangle = \sum_k l_k |a_k\rangle |b_k\rangle$.

Information theory and thermodynamics
We derive now the Boltzmann factor for the canonical and grand canonical ensembles and obtain the combined expression of the first and second laws of thermodynamics from the Boltzmann-Gibbs-Shannon-von Neumann entropy.

Starting with a general treatment, we consider the problem of determining the least biased density matrix $\rho$, knowing the expectation values of the observables $F^\alpha$, given that $Tr \rho F^\alpha = f^\alpha$, $\alpha = 1, \cdots, N$, besides the probability normalization condition $Tr \rho = 1$. 

The answer is to maximize the Boltzmann-Gibbs-Shannon-von Neumann entropy $S = -k_B T \rho \ln \rho$, where $k_B$ is the Boltzmann’s constant, subject to those constraints, i.e. maximize $\frac{S}{k_B} = \lambda_0 (Tr \rho - 1) - \sum_{\alpha=1}^{N} \lambda_{\alpha} (Tr \rho F^{\alpha} - f^{\alpha})$, where $\lambda_{\alpha}$ are Lagrange multipliers. This approach is also known as the Jaynes principle [31, 32].

Using the cyclic property of the trace to write $d \operatorname{Tr} f(\rho) = \operatorname{Tr} f'(\rho) d\rho$, the solution is given by

$$\rho = \frac{e^{-\sum_{\alpha} \lambda_{\alpha} F^{\alpha}}}{Z},$$

where $Z = e^{\lambda_{0} + 1} = Tr e^{-\sum_{\alpha} \lambda_{\alpha} F^{\alpha}}$ is the partition function.

The problem under consideration can change, either because the expectation values $f^{\alpha}$ change or because the observables $F^{\alpha}$ themselves change with some external parameter $\theta$, as for example the eigen-energies of a system when its volume $V$ or magnetic field changes. Using again the cyclic property of the trace, one has then

$$d \ln Z = - \sum_{\alpha} d\lambda_{\alpha} f^{\alpha} - \sum_{\alpha} \lambda_{\alpha} \langle \frac{dF^{\alpha}}{d\theta} \rangle d\theta,$$

where $\langle \frac{dF^{\alpha}}{d\theta} \rangle$ denotes the expectation values, defined by $\rho$, of the variation of the observables. Using eq. B.1 the entropy becomes

$$\frac{S}{k_B} = \ln Z + \sum_{\alpha} \lambda_{\alpha} f^{\alpha}.$$  

Its variation is given by

$$d \frac{S}{k_B} = \sum_{\alpha} \lambda_{\alpha} \left( df^{\alpha} - \langle \frac{dF^{\alpha}}{d\theta} \rangle d\theta \right),$$

having used eq. B.2. Since one has $\frac{\partial \frac{S}{k_B}}{\partial f^{\alpha}} = \lambda_{\alpha}$, i.e. the Lagrange multipliers are the partial derivatives of the entropy with respect to the averages $f^{\alpha}$, and $\frac{\partial \frac{S}{k_B}}{\partial \theta} = - \sum_{\alpha} \lambda_{\alpha} \langle \frac{dF^{\alpha}}{d\theta} \rangle$, it follows that

$$\ln Z = \frac{S}{k_B} - \sum_{\alpha} \frac{\partial \frac{S}{k_B}}{\partial f^{\alpha}} f^{\alpha},$$

is the Lévy transform of the entropy with respect to the expectation values of the observables.

If the entropy is extensive, i.e. a homogeneous function of degree 1 of its variables, including the external parameter, then we will have $\frac{S}{k_B} = \sum_{\alpha} \lambda_{\alpha} \left( f^{\alpha} - \langle \frac{dF^{\alpha}}{d\theta} \rangle \theta \right)$ and $

\ln Z = - \sum_{\alpha} \lambda_{\alpha} \langle \frac{dF^{\alpha}}{d\theta} \rangle \theta$.

Turning now to thermodynamics, the observable to be considered in the canonical ensemble is the energy $E = \langle \mathcal{H} \rangle$ and in the grand canonical ensemble are the energy and the particle number. The Lagrange multiplier associated to the energy is $\beta = \frac{1}{k_B T}$, the inverse of the absolute temperature $T$. The Lagrange multiplier associated to the particle number $N$ is $-\beta \mu$, where $\mu$ is the chemical potential. The density matrices for these ensembles are then given by eqs. 39 and 40.

We find that $d \frac{S}{k_B} = \beta (dE - \langle d\mathcal{H} \rangle) - \beta \mu dN$. Defining $\langle \frac{d\mathcal{H}}{dV} \rangle = -p$, where $V$ is the volume, we have $dS = \frac{dE}{T} + \beta dV - \frac{d\mu}{dN}$, or equivalently,

$$dE = -pdV + TdS + \mu dN,$$

which is the combined expression of the first and second laws of thermodynamics.

We also have $\ln Z = \frac{S}{k_B} - \beta E + \beta \mu N = -\beta (E - TS - \mu N)$ and $d \ln Z = -Ed\beta + Nd(\beta \mu) + \beta pdV$.

Finally, we have $\ln Z = \beta pV$, for an extensive system.
References

[1] Nielsen M A and Chuang I L 2000 Quantum Computation and Quantum Information (Cambridge, England: Cambridge University Press)
[2] Amico L, Fazio R, Osterloh A and Vedral V 2008 Rev. Mod. Phys. 80 517
[3] Amari S and Nagaoka H 2000 Methods of information geometry (Transactions of mathematical monographs; v. 191, American Mathematical Society, Oxford University Press)
[4] Nakahara M 1990 Geometry, Topology and Physics (Bristol: Institute of Physics Publishing)
[5] Uhlmann A 1993 Rep. Math. Phys. 33 253
[6] Uhlmann A 1976 Rep. Math. Phys. 9 273
[7] Louisell W H 1973 Quantum Statistical Properties of Radiation (John Wiley & Sons)
[8] Takahashi Y and Umezawa H 1975 Collective Phenomena 2 55–80 [reprinted in Int. J. Mod. Physics B 10, 1755 (1996)]
[9] Umezawa H, Matsumoto H and Tachiki M 1982 Thermofield Dynamics and Condensed States (North-Holland)
[10] Jozsa R 1994 J. Mod. Opt. 41 2315
[11] Sommers H J and Życzkowski K 2003 J. Phys. A 36 10083
[12] Bures D J C 1969 Trans. Amer. Math. Soc. 135 199
[13] Hübner M 1993 Phys. Lett. A 179 226
[14] Domal C 1996 The Critical Point: A Historical Introduction to the Modern Theory of Critical Phenomena (CRC Press)
[15] Goldenfeld N 1992 Lectures on Phase Transitions and the Renormalization Group (Frontiers in Physics 85, Westview Press)
[16] Sachdev S 2000 Quantum Phase Transitions (Cambridge, England: Cambridge University Press)
[17] Beresinskii V L 1971 Sov. Phys. JETP 32 493
[18] Kosterlitz J M and Thouless D J 1973 J. Phys. C 6 1181
[19] Kosterlitz J M 1974 J. Phys. C 7 1046
[20] Wen X G 2004 Quantum Field Theory of Many-Body Systems (New York: Oxford University)
[21] Zanardi P and Paunkovic N 2006 Phys. Rev. E 74 031123
[22] Zhou H Q and Barjaktarevic J P 2008 J. Phys. A: Math. Theor. 41 412001
[23] Paunkovic N, Sacramento P D, Nogueira P, Vieira V R and Dugaev V K 2008 Phys. Rev. A 77 052302
[24] You W L, Li Y W and Gu S J 2007 Phys. Rev. E 76 022101
[25] Paunkovic N and Vieira V R 2008 Phys. Rev. E 77 011129
[26] Yu W C, Kwok H M, Cao J and Gu S J 2009 Phys. Rev. E 80 021108
[27] Zanardi P, Venuti L C and Giorda P 2007 Phys. Rev. A 76 062318
[28] Uhlmann A 1986 Rep. Math. Phys. 24 229
[29] Zanardi P, Giorda P and Cozzini M 2007 Phys. Rev. Lett. 99 100603
[30] Gu S J arXiv:0811.3127v1
[31] Jaynes E T 1957 Physical Review 106 620
[32] Jaynes E T 1957 Physical Review 108 171