



**HAL**  
open science

## On two properties of the Fisher information

Nicolas Rougerie

► **To cite this version:**

Nicolas Rougerie. On two properties of the Fisher information. *Kinetic and Related Models*, 2020, 14 (1), pp.77-88. hal-02397359v3

**HAL Id: hal-02397359**

**<https://hal.science/hal-02397359v3>**

Submitted on 20 Aug 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ON TWO PROPERTIES OF THE FISHER INFORMATION

NICOLAS ROUGERIE

ABSTRACT. Alternative proofs for the superadditivity and the affinity (in the large system limit) of the usual and some fractional Fisher informations of a probability density of many variables are provided. They are consequences of the fact that such informations can be interpreted as quantum kinetic energies.

## CONTENTS

1. Introduction	1
2. Definitions and Results	2
3. Proofs	6
3.1. Preliminaries	6
3.2. Superadditivity	8
3.3. Affinity	9
References	12

## 1. INTRODUCTION

The Fisher information of a symmetric probability measure of  $N$  variables (see Definition 2.1 below) is known to be:

**A.** Superadditive: the information of the full measure is not smaller than the sum of the informations of marginals with  $n$  and  $N - n$  variables.

**B.** Affine linear in the large  $N$  limit: for probability measures that have a limit  $N \rightarrow \infty$ , the natural limiting information (mean, or level-3, information) is an affine functional.

**A** has been proved first in [3, Theorem 3] and then, by another method, in [8, Lemma 3.7]. **B** is deduced in [12, Proposition 3] from the corresponding property for the mean entropy, originating in [18]. In [8, Section 5.3] a different proof is provided, based on a general abstract linearity lemma. The analogues for some fractional variants of Fisher's information are obtained in [21, 22].

This note provides alternative proofs of these properties. If the Fisher information of a probability density  $\mu_N \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{dN})$  is interpreted as a quantum kinetic energy for the quantum state  $|\Psi_N\rangle\langle\Psi_N|$ , orthogonal projector on the “quantum wave-function”  $\Psi_N = \sqrt{\mu_N} \in L^2_{\text{sym}}(\mathbb{R}^{dN})$ , both properties become quite natural. Roughly speaking **A** is a consequence of the convexity of the kinetic energy as a function of  $|\Psi_N|^2$  and **B** follows from its affinity as a function of  $|\Psi_N\rangle\langle\Psi_N|$ .

---

*Date:* August, 2020.

As for the motivations behind proving **A** and **B**, they mostly come from the study of mean-field limits of large systems of statistical mechanics, classical and quantum. Indeed, they do not seem to be of use in the theory of sufficient statistics, where the Fisher information originates. We however note that any quantity interpreted as an “information” should certainly satisfy **A**. **B** has been known to hold for the entropy for a long time [18], and it seems natural to ask its equivalent for the Fisher information.

Both properties are crucial to Kiessling’s approach [12] of the mean-field limit of bosonic ground states (see [19, 20, Appendix A] for review), which involves interpreting the quantum kinetic energy as a classical Fisher information. In fact we exploit here the reverse of Kiessling’s point of view.

In [5], Fisher information bounds are used to control the mean-field limit of a classical statistical mechanics system with stochastic diffusions (see [7] for review). The method is general and has been adapted to other models, e.g. in [21, 22]. Briefly, the entropy production<sup>1</sup> along the flow is controlled first. The Fisher information is naturally linked to variations of the entropy<sup>2</sup>, and one then deduces a control of the former quantity. This has several important applications, one of which relies on **A** and **B** above. Briefly, **A** allows to pass to the large  $N$  limit and control the mean information of limiting objects. Then **B** implies that the de Finetti-Hewitt-Savage mixing measure of the limit is concentrated on probability measures with finite Fisher information. Uniqueness theorems for the mean-field equation in the latter class can then be put to good use.

**Acknowledgments.** Thanks to Samir Salem, conversations with whom motivated the write-up of this note, and to Mathieu Lewin for discussions on related topics a while ago. Financial support was provided by the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation Programme (Grant agreement CORFRONMAT No 758620).

## 2. DEFINITIONS AND RESULTS

We are concerned with classical mechanics states, symmetric probability measures  $\mu_N \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{dN})$  for the distribution of  $N$  indistinguishable particles living say in  $\mathbb{R}^d$ . We will freely identify measures and their densities with respect to Lebesgue measure on  $\mathbb{R}^{dN}$ . Symmetric here means

$$\mu_N(x_1, \dots, x_N) = \mu_N(x_{\sigma(1)}, \dots, x_{\sigma(N)}) \quad (2.1)$$

for almost all  $X_N = (x_1, \dots, x_N) \in \mathbb{R}^{dN}$  and any permutation  $\sigma$  of the  $N$  indices. In the applications we have in mind,  $X_N$  can be a collection of spatial coordinates (this is the case for bosonic mean-field limits as considered in [12]) or a collection of velocity variables (this is the case for the applications in kinetic theory [8, 21, 22]).

---

<sup>1</sup>I follow the physicists’ convention that the entropy is produced, not dissipated.

<sup>2</sup>The Fisher information is the derivative of the entropy along the heat flow.

The quantities of our interest are as given in the

**Definition 2.1 (Fisher informations).**

For  $\mu_N \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{dN})$  we define

- the Fisher information

$$\mathcal{I}_1[\mu_N] = \left\langle \sqrt{\mu_N} \left| \sum_{j=1}^N -\Delta_{x_j} \right| \sqrt{\mu_N} \right\rangle_{L^2(\mathbb{R}^{dN})}. \quad (2.2)$$

- the fractional Fisher information of order  $0 < s < 1$

$$\mathcal{I}_s[\mu_N] = \left\langle \sqrt{\mu_N} \left| \sum_{j=1}^N (-\Delta_{x_j})^s \right| \sqrt{\mu_N} \right\rangle_{L^2(\mathbb{R}^{dN})}. \quad (2.3)$$

*Remarks.*

**1.** The following equivalent definitions are well-known. For the usual Fisher information (the first one is maybe the most commonly used) we have

$$\begin{aligned} \mathcal{I}_1[\mu_N] &= \frac{1}{4} \int_{\mathbb{R}^{dN}} |\nabla \log \mu_N|^2 \mu_N = \frac{1}{4} \int_{\mathbb{R}^{dN}} \frac{|\nabla \mu_N|^2}{\mu_N} = \int_{\mathbb{R}^{dN}} |\nabla \sqrt{\mu_N}|^2 \\ &= \sum_{j=1}^N \int_{\mathbb{R}^{dN}} |k_j|^2 \left| \widehat{\sqrt{\mu_N}}(k) \right|^2 dk \end{aligned} \quad (2.4)$$

and for the fractional Fisher information:

$$\begin{aligned} \mathcal{I}_s[\mu_N] &= \sum_{j=1}^N \int_{\mathbb{R}^{dN}} |k_j|^{2s} \left| \widehat{\sqrt{\mu_N}}(k) \right|^2 dk \\ &= C_{d,s} N \int_{\mathbb{R}^{d(N+1)}} \frac{\left| \sqrt{\mu_N}(x, x_1, \dots, x_{N-1}) - \sqrt{\mu_N}(y, x_1, \dots, x_{N-1}) \right|^2}{|x - y|^{d+s}} dx dy dx_1 \dots dx_{N-1}. \end{aligned} \quad (2.5)$$

Here hat-bearing functions stand for Fourier transforms and  $C_{d,s}$  is a constant only depending on  $d$  and  $s$ . That these various definitions are equivalent either follows from straightforward calculations or is proved in standard textbooks, such as [15]. The precise value of the constant  $C_{d,s}$  is of no concern to this note but can be found in [15, 21, 22].

**2.** It follows from results of [1, 16, 2, 17] that

$$\mathcal{I}_1[\mu_N] = C_d \lim_{s \uparrow 1} (1 - s) C_{d,s}^{-1} \mathcal{I}_s[\mu_N] \quad (2.6)$$

with

$$C_d = \left( \int_{S^{d-1}} \cos \theta d\sigma \right)^{-1}.$$

Here  $S^{d-1}$  is the euclidean sphere equipped with its Lebesgue measure  $d\sigma$  and  $\theta = \theta(\sigma)$  represents the angle of  $\sigma$  with respect to the vertical axis. This implies that  $\mathcal{I}_1 \equiv \mathcal{I}$  is a natural limit case of  $\mathcal{I}_s$  for  $s \rightarrow 1$ .

**3.** Other types of fractional Fisher informations are discussed in [25, 27, 26], in connection with statistics, and [21], in connection with the fractional heat flow.

**4.** In [22] the fractional Fisher information is defined with an extra “cut-off”

$$\mathcal{I}_{s,\gamma}[\mu_N] = \sum_{j=1}^N \langle \sqrt{\mu_N} | \chi(x_j) (-\Delta_{x_j})^s \chi(x_j) | \sqrt{\mu_N} \rangle_{L^2}$$

with  $\chi(x) = (1 + |x|^2)^{2\gamma}$ ,  $\gamma < 0$ . This does not significantly change the structure of the object, nor the proofs of the results below. Indeed the “kinetic energy”

$$\langle u | \chi(-\Delta)^s \chi | u \rangle$$

still enjoys the properties we need (see Properties 3.1), of which convexity as a function of  $|u|^2$  and affinity as a function of the orthogonal projector  $|u\rangle\langle u|$  are the most crucial. We leave the adaptations to the reader.  $\square$

Define now, for any integer  $n \leq N$ , the  $n$ -th marginal/reduced density of a measure  $\mu_N \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{dN})$  as the probability measure on  $\mathbb{R}^{dn}$  with density

$$\mu_N^{(n)}(x_1, \dots, x_n) := \int_{\mathbb{R}^{d(N-n)}} \mu_N(x_1, \dots, x_n, y_{n+1}, \dots, y_N) dy_{n+1} \dots dy_N. \quad (2.7)$$

The first result we provide an alternative proof for is the superadditivity of the functionals from Definition 2.1:

**Theorem 2.2 (Superadditivity of Fisher informations).**

Let  $n < N$  be two integers and  $\mu_N \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{dN})$ . We have that

$$\mathcal{I}_s[\mu_N] \geq \mathcal{I}_s[\mu_N^{(n)}] + \mathcal{I}_s[\mu_N^{(N-n)}] \quad (2.8)$$

for any  $0 < s \leq 1$ .

*Remarks.*

**1.** For  $s = 1$ , Carlen proved this via a Minkowski-like inequality. Hauray-Mischler use a dual formulation of the Fisher information (not mentioned in the remark after Definition 2.1) to recover the result. For  $s < 1$  the result is obtained in [22] using the last formulation in (2.5). As per (2.6), this also implies the result for  $s = 1$ .

**2.** The short proof we provide uses standard tools of quantum mechanics: reduced density matrices and convexity of  $\mu_N \mapsto \mathcal{I}_s[\mu_N]$ .

**3.** A useful consequence is that, if  $N$  is an integer times  $n$

$$\frac{1}{N} \mathcal{I}_s[\mu_N] \geq \frac{1}{n} \mathcal{I}_s[\mu_N^{(n)}]. \quad (2.9)$$

$\square$

In statistical mechanics one is often interested in the limit of large particle numbers,  $N \rightarrow \infty$ , maybe with other parameters of the model scaled appropriately. Then our classical states turn into symmetric probability measures over infinite sequences,  $\mu \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{d\mathbb{N}})$ . One can also be interested in the limits  $N \rightarrow \infty$  with fixed  $n$  of the marginals (2.7). It can then be useful to have a notion of “mean Fisher information” (sometimes also referred to as level-3 information):

**Definition 2.3 (Mean Fisher information).**

Let  $\mu \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{d\mathbb{N}})$  be a symmetric probability measure over sequences in  $\mathbb{R}^d$ . Equivalently<sup>3</sup>, let  $(\mu^{(n)})_n$  be a sequence of symmetric probability measures over  $\mathbb{R}^{dn}$  satisfying the consistency condition

$$\left(\mu^{(n+1)}\right)^{(n)} = \mu^{(n)}.$$

For  $0 < s \leq 1$  the mean (fractional) Fisher information of  $\mu$  is

$$\mathcal{I}_s[\mu] := \limsup_{n \rightarrow \infty} \frac{1}{n} \mathcal{I}_s[\mu^{(n)}]. \quad (2.10)$$

The existence of the lim sup follows from (2.9). It is in fact possible to see that the lim sup is both a sup and a lim, using Theorem 2.4.

The definition is in complete analogy with that of the mean entropy, originating in [18]. Perhaps surprisingly, this functional is affine, just as the mean entropy. We shall give an alternative proof of the

**Theorem 2.4 (The mean Fisher information is affine).**

Let  $\mu \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{d\mathbb{N}})$  and  $P \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$  be its unique de Finetti-Hewitt-Savage measure, i.e., for all  $n \geq 0$

$$\mu^{(n)} = \int_{\mathcal{P}(\mathbb{R}^d)} \rho^{\otimes n} dP(\rho). \quad (2.11)$$

Assume that there exists a locally bounded  $V : \mathbb{R}^d \mapsto \mathbb{R}$  with  $V(x) \xrightarrow{|x| \rightarrow \infty} +\infty$  such that

$$\int_{\mathbb{R}^d} V \mu^{(1)} < \infty. \quad (2.12)$$

Then, the mean Fisher informations from Definition 2.3 satisfy, for all  $0 < s \leq 1$

$$\mathcal{I}_s[\mu] = \int_{\mathcal{P}(\mathbb{R}^d)} \mathcal{I}_s[\rho] dP(\rho) \quad (2.13)$$

with  $\mathcal{I}_s[\rho]$  as in Definition 2.1 with  $N = 1$ .

*Remarks.*

**1.** In (2.12) I demand a bit more “confinement” than in previous versions of the statement [12, 8, 22]. This is harmless in applications, for the theorem is meant to be applied to limits  $N \rightarrow \infty$  of  $N$ -body classical states  $\mu_N$ . In order for a state with infinitely many particles to exist in the limit, a tightness argument of the type

$$\int_{\mathbb{R}^d} V \mu_N^{(1)} < \infty, \text{ independently of } N$$

is usually needed.

**2.** Kiessling [12] proved the  $s = 1$  case, using the better known [18] affinity of the mean entropy

$$\mathcal{S}[\mu] := - \limsup_{n \rightarrow \infty} \frac{1}{n} \int_{\mathbb{R}^{dn}} \mu^{(n)} \log \mu^{(n)}.$$

The Fisher information is the derivative of the entropy along the heat flow. Since the latter is linear (as is differentiation), the affinity of the Fisher information follows.

<sup>3</sup>By a theorem of Kolmogorov.

**3.** Hauray and Mischler [8] gave another proof for  $s = 1$ . A first sanity check is to convince oneself that

$$\mathcal{I} \left[ \frac{1}{2} \rho_1^{\otimes \infty} + \frac{1}{2} \rho_2^{\otimes \infty} \right] = \frac{1}{2} \mathcal{I} [\rho_1^{\otimes \infty}] + \frac{1}{2} \mathcal{I} [\rho_2^{\otimes \infty}] \quad (2.14)$$

where  $\rho^\infty$  is the measure over  $\mathbb{R}^{dN}$  with  $n$ -th marginal  $\rho^{\otimes n}$  for any  $n$ . The reason for this is that, if  $\rho_1 \neq \rho_2$ ,  $\rho_1^{\otimes N}$  becomes more and more alien (“orthogonal”) to  $\rho_2^{\otimes N}$  for large  $N$ .

Very briefly, the proof of [8] checks a more elaborate version of the “partial affinity” (2.14), and then applies a general abstract lemma implying full affinity.

**4.** The same strategy is applied to the  $s < 1$  case in [22]. The abstract lemma applies mutatis mutandis, but the argument giving the partial affinity is different. The sanity check (2.14) can also be found in [19, 20, Appendix A].

**5.** Salem proves in [21] that the variant mean fractional information<sup>4</sup> based on

$$\tilde{\mathcal{I}}_s[\mu_N] = N \int_{\mathbb{R}^{d(N+1)}} \frac{\Phi \left( \sqrt{\mu_N(x, x_1, \dots, x_{N-1})}, \sqrt{\mu_N(y, x_1, \dots, x_{N-1})} \right)}{|x - y|^{d+s}} dx dy dx_1 \dots dx_{N-1}$$

with

$$\Phi(x, y) = (x - y)(\log x - \log y)$$

enjoys similar properties as that we defined, in particular affinity. See [21, Remark 3.4] for more comments on the relation between  $\mathcal{I}_s$  and  $\tilde{\mathcal{I}}_s$  and their respective uses. In particular, for the applications of [21], results on  $\mathcal{I}_s$  could serve as alternatives to those on  $\tilde{\mathcal{I}}_s$ .  $\square$

### 3. PROOFS

**3.1. Preliminaries.** Our point of view in this note is to think quantum mechanically, that is, in terms of  $L^2$  functions and operators acting on them, rather than in terms of probability measures. Pick  $\mu_N \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{dN})$ . For it to have a finite Fisher information it must actually be a function. Define then

$$\Psi_N = \sqrt{\mu_N}, \quad \Gamma_N = |\Psi_N\rangle\langle\Psi_N| \quad (3.1)$$

The first object above is a bosonic wave-function, namely  $\Psi_N \in L^2_{\text{sym}}(\mathbb{R}^{dN}, \mathbb{C})$  satisfies

$$\Psi_N(x_1, \dots, x_N) = \Psi_N(x_{\sigma(1)}, \dots, x_{\sigma(N)})$$

in analogy with (2.1). The second object in (3.1) is the  $L^2$ -orthogonal projector on the complex linear span of  $\Psi_N$ . It is a bosonic state, i.e. a positive trace-class operator with trace 1, acting on  $L^2_{\text{sym}}(\mathbb{R}^{dN}, \mathbb{C})$ . Note that we do not use the usual quantization of classical mechanics: we simply use that a classical state of position *or* velocity variables can be directly embedded in a quantum formalism.

We shall write Fisher informations as quantum kinetic energies of  $\Psi_N$  or  $\Gamma_N$

$$\mathcal{I}_s[\mu_N] = \langle \Psi_N | H_N | \Psi_N \rangle_{L^2} = \text{Tr}(H_N \Gamma_N) \quad (3.2)$$

with

$$H_N = \sum_{j=1}^N h_{x_j}, \quad h = (-\Delta)^s$$

---

<sup>4</sup>This is the derivative of the entropy along the fractional heat flow.

and  $h_{x_j}$  acting on the variable  $x_j$ . We do not emphasize the dependence on  $s$ , for our proofs shall be based solely on the following

**Properties 3.1 (Quantum kinetic energies).**

*The kinetic energy*

$$L^2(\mathbb{R}^d, \mathbb{C}) \ni u \mapsto \langle u | h | u \rangle \in \mathbb{R}^+ \cup \{+\infty\}$$

with  $h$  as above is

- (1) *Positivity preserving*

$$\langle u | h | u \rangle \geq \langle |u| | h | |u| \rangle$$

- (2) *Convex as a function of  $|u|^2$ :*

$$L^1(\mathbb{R}^d, \mathbb{R}^+) \ni \rho \mapsto \langle \sqrt{\rho} | h | \sqrt{\rho} \rangle \in \mathbb{R}^+ \cup \{+\infty\}$$

*is convex.*

- (3) *With locally compact resolvent. For a locally bounded  $V : \mathbb{R}^d \mapsto \mathbb{R}$  with*

$$V(x) \xrightarrow{|x| \rightarrow \infty} +\infty$$

*the operator  $h+V$  has compact resolvent, i.e.  $(h+V+c)^{-1}$  is compact as an operator on  $L^2$ , where  $c$  is a constant sufficiently large for the inverse to make sense.*

*Remarks.* The first property for  $s = 1$  is just the straightforward (at least for smooth functions  $u = |u|e^{i\varphi}$ ) identity

$$|\nabla u|^2 = |\nabla |u||^2 + |u|^2 |\nabla \varphi|^2,$$

a particular case of the diamagnetic inequality [15, Theorem 7.12]. For  $s > 1$  it follows immediately from the last definition in (2.5), the triangle inequality, and (2.6). Note that “positivity preserving” usually means something stronger (but also true in the case at hand), namely that the heat flow associated with  $h$  preserves positivity of functions. The property we require usually goes hand-in-hand with this “true” positivity-preserving property.

The second property can be found in [15, Theorems 7.8 and 7.13]. The third property follows from the Sobolev compact embedding in  $L^2$ .  $\square$

We also recall the notion of reduced density matrix, extending that of marginal (2.7). We define the  $n$ -th reduced density matrix  $\Gamma_N^{(n)}$  by a partial trace

$$\Gamma_N^{(n)} := \text{Tr}_{n+1 \rightarrow N} \Gamma_N. \quad (3.3)$$

This is the operator on  $L^2_{\text{sym}}(\mathbb{R}^{dn})$  defined by the relation

$$\text{Tr} \left( A_n \Gamma_N^{(n)} \right) = \text{Tr} \left( A_n \otimes \mathbf{1}^{\otimes (N-n)} \Gamma_N \right)$$

for any bounded operator  $A_n$  on  $L^2_{\text{sym}}(\mathbb{R}^{dn})$ . Note that  $\Gamma_N$ , as a trace-class (in particular, Hilbert-Schmidt) operator on  $L^2$  has an integral kernel

$$\Gamma_N(x_1, \dots, x_N; y_1, \dots, y_N) = \overline{\Psi_N(y_1, \dots, y_N)} \Psi_N(x_1, \dots, x_N).$$



The integral kernel of  $\Gamma_N^{(n)}$  is then given, formally, as

$$\Gamma_N^{(n)}(x_1, \dots, x_n; y_1, \dots, y_n) = \int_{\mathbb{R}^{d(N-n)}} \Gamma_N(x_1, \dots, x_n, z_{n+1}, \dots, z_N; y_1, \dots, y_n, z_{n+1}, \dots, z_N) dz_{n+1} \dots dz_N.$$

One use of these objects is that, if  $N$  is a multiple of  $n$ ,

$$\mathcal{I}_s[\mu_N] = \text{Tr}(H_N \Gamma_N) = \frac{N}{n} \text{Tr}\left(H_n \Gamma_N^{(n)}\right) \quad (3.4)$$

taking partial traces and using that  $H_N$  is a sum of terms acting on one variable at a time.

**3.2. Superadditivity.** We now prove Theorem 2.2. The following considerations are very much in the spirit of the Hoffmann-Ostenhof<sup>2</sup> inequality [10, 13].

From (3.2) and (3.3) we have, similarly to in (3.4),

$$\mathcal{I}_s[\mu_N] = \text{Tr}\left(H_n \Gamma_N^{(n)}\right) + \text{Tr}\left(H_{N-n} \Gamma_N^{(N-n)}\right) \quad (3.5)$$

where  $\mu_N$  and  $\Gamma_N$  are related by (3.1). Now,  $\Gamma_N^{(n)}$  being a positive trace-class operator with unit trace, the spectral theorem implies the existence of an orthonormal basis  $(u_j)$  of  $L^2(\mathbb{R}^{dn})$  such that

$$\Gamma_N^{(n)} = \sum_j \lambda_j |u_j\rangle\langle u_j|$$

where the positive numbers  $\lambda_j$  add to 1. Thus

$$\text{Tr}\left(H_n \Gamma_N^{(n)}\right) = \sum_j \lambda_j \langle u_j | H_n | u_j \rangle$$

and, using Items 1 and 2 in Properties 3.1 we get

$$\text{Tr}\left(H_n \Gamma_N^{(n)}\right) \geq \sum_j \lambda_j \langle |u_j| | H_n | |u_j| \rangle \geq \langle \sqrt{\rho_n} | H_n | \sqrt{\rho_n} \rangle$$

with

$$\rho_n := \sum_j \lambda_j |u_j|^2.$$

The proof is concluded by observing that

$$\rho_n = \mu_N^{(n)} \quad (3.6)$$

with  $\mu_N^{(n)}$  the  $n$ -th marginal of  $\mu_N$  and arguing similarly for the second term of (3.5).

To see the truth of (3.6), identify any bounded function  $V_n$  of  $n$  variables in  $\mathbb{R}^d$  with the corresponding multiplication operator on  $L^2(\mathbb{R}^{dn})$ . Then

$$\begin{aligned} \int_{\mathbb{R}^{dn}} V_n \rho_n &= \text{Tr} \left( V_n \Gamma_N^{(n)} \right) \\ &= \text{Tr} \left( V_n \otimes \mathbf{1}^{\otimes(N-n)} \Gamma_N \right) \\ &= \left\langle \Psi_N | V_n \otimes \mathbf{1}^{\otimes(N-n)} | \Psi_N \right\rangle \\ &= \int_{\mathbb{R}^{dN}} V_n(x_1, \dots, x_n) |\Psi_N(x_1, \dots, x_N)|^2 dx_1, \dots, dx_N \\ &= \int_{\mathbb{R}^{dn}} V_n \mu_N^{(n)}. \end{aligned}$$

This completes the proof.

**3.3. Affinity.** Our proof of Theorem 2.4 is based on the manifest affinity of (3.2) as a function of  $\Gamma_N$ , and the quantum de Finetti theorem, a generalization of the classical de Finetti-Hewitt-Savage theorem, see [20, 19] for review. Readers familiar with the classical theorem could note that the version of the quantum theorem I use has a proof (see [11] and [14, Appendix A]) essentially identical to that of Hewitt-Savage [9].

First observe that, as per Item 2 in Properties 3.1, Definition 2.3 and (2.11) we immediately have that

$$\mathcal{I}_s[\mu] \leq \int_{\mathcal{P}(\mathbb{R}^d)} \mathcal{I}_s[\rho] dP(\rho).$$

We aim at a corresponding lower bound.

Define, for any  $N \in \mathbb{N}$ , a bosonic quantum wave-function and a bosonic quantum state as

$$\Psi_N = \sqrt{\mu^{(N)}}, \quad \Gamma_N = |\Psi_N\rangle\langle\Psi_N|$$

where  $\mu^{(N)}$  is the  $N$ -th marginal of  $\mu \in \mathcal{P}_{\text{sym}}(\mathbb{R}^{dN})$ . Let then, for  $k \leq N$ ,  $\Gamma_N^{(k)}$  be the reduced density matrix (3.3) of  $\Gamma_N$ . At fixed  $k$ , the sequence  $(\Gamma_N^{(k)})_N$  is by definition bounded in the trace-class. Modulo a (not-relabeled) subsequence we thus have

$$\Gamma_N^{(k)} \xrightarrow{\star} \gamma^{(k)} \tag{3.7}$$

in the trace-class, as  $N \rightarrow \infty$ . The latter being [24, 23] the dual of the compact operators (equipped with the operator norm, and where the duality bracket between two operators  $A, B$  is given by  $\text{Tr}(AB)$ ), this means

$$\text{Tr} \left( K_k \Gamma_N^{(k)} \right) \rightarrow \text{Tr} \left( K_k \gamma^{(k)} \right) \tag{3.8}$$

for any compact operator  $K_k$  on  $L_{\text{sym}}^2(\mathbb{R}^{dk})$ . By a diagonal extraction argument we can assume that, for all  $k \geq 0$  the sequences  $(\Gamma_N^{(k)})_N$  converges weakly- $\star$ , along a common subsequence in  $N$ .

Let  $V$  be the potential such that (2.12) holds and

$$h^V = h + V.$$

Then

$$\begin{aligned} \mathrm{Tr} \left( \sum_{j=1}^k h_{x_j}^V \Gamma_N^{(k)} \right) &= \frac{k}{N} \left\langle \sqrt{\mu^{(N)}} \left| \sum_{j=1}^N h_{x_j} + V(x_j) \right| \sqrt{\mu^{(N)}} \right\rangle_{L^2} \\ &\leq k \left( \mathcal{I}_s[\mu] + \int_{\mathbb{R}^d} \mu^{(1)} V \right) < \infty \end{aligned}$$

by assumption. Hence, for all  $k \geq 0$  and any constant  $c > 0$

$$\mathrm{Tr} \left( \sum_{j=1}^k (h + V + c)_{x_j} \Gamma_N^{(k)} \right) \leq C_{k,c}$$

independently of  $N$ . Using cyclicity of the trace, the positive operator  $L_k^{1/2} \Gamma_N^{(k)} L_k^{1/2}$  with  $L_k = \sum_{j=1}^k (h + V + c)_{x_j}$  is bounded in the trace-class. As above this implies that (modulo a further extraction of a subsequence) it converges weakly- $\star$  (in the sense of (3.8))

$$L_k^{1/2} \Gamma_N^{(k)} L_k^{1/2} \xrightarrow{\star} L_k^{1/2} \gamma^{(k)} L_k^{1/2} \quad (3.9)$$

where the limit is identified by testing the convergence with smooth finite-rank operators and recalling (3.7). As per Item 3 in Properties 3.1,  $h + V + c$  has compact inverse provided  $c$  is chosen large enough. Consequently, so does  $L_k$ . Then, by (3.9),

$$\mathrm{Tr} \Gamma_N^{(k)} = \mathrm{Tr} \left( L_k^{-1} L_k^{1/2} \Gamma_N^{(k)} L_k^{1/2} \right) \rightarrow \mathrm{Tr} \left( L_k^{-1} L_k^{1/2} \gamma^{(k)} L_k^{1/2} \right) = \mathrm{Tr} \left( \gamma^{(k)} \right).$$

This proves convergence of the trace-class norm

$$1 = \mathrm{Tr} \left( \Gamma_N^{(k)} \right) \rightarrow \mathrm{Tr} \gamma^{(k)}$$

and hence that actually

$$\Gamma_N^{(k)} \rightarrow \gamma^{(k)} \quad (3.10)$$

strongly in the trace-class norm (see [4] or [24, Addendum H]).

We now use the strong<sup>5</sup> quantum de Finetti theorem (see [11, 14] and [19, 20] for review) to obtain the existence and uniqueness of a probability measure  $Q$  over the unit sphere of  $L^2(\mathbb{R}^d)$  such that

$$\gamma^{(k)} = \int |u^{\otimes k}\rangle \langle u^{\otimes k}| dQ(u). \quad (3.11)$$

By (3.4) (with  $n = 1$ )

$$\frac{1}{N} \mathcal{I}_s[\mu^{(N)}] = \mathrm{Tr} \left( h \Gamma_N^{(1)} \right)$$

and a lower semi-continuity argument gives

$$\liminf_{N \rightarrow \infty} \mathrm{Tr} \left( h \Gamma_N^{(1)} \right) \geq \mathrm{Tr} \left( h \gamma^{(1)} \right)$$

---

<sup>5</sup>”Strong” refers to the fact that we use strong trace-class convergence of reduced density matrices. This is the quantum analogue of the Hewitt-Savage theorem where one uses tightness of the marginals. There is also a ”weak” quantum de Finetti theorem, relying only on weak- $\star$  convergence of reduced density matrices. We do not use it here.

so that, combining the two observations and (3.11),

$$\mathcal{I}_s[\mu] \geq \int \langle u|h|u \rangle dQ(u).$$

Recalling Item 1 of Properties 3.1 this yields

$$\mathcal{I}_s[\mu] \geq \int \langle |u||h||u| \rangle dQ(u) = \int \langle \sqrt{|u|^2} |h| \sqrt{|u|^2} \rangle dQ(u)$$

and the proof will be complete once we have proven the next lemma:

**Lemma 3.2 (Identification of de Finetti measures).**

Let  $P \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$  be the classical de Finetti measure defined in Theorem 2.4 and  $Q \in \mathcal{P}(L^2(\mathbb{R}^d))$  the quantum de Finetti measure defined in (3.11). We have, for any bounded continuous function  $\Phi$  over  $\mathcal{P}(\mathbb{R}^d)$

$$\int \Phi(\rho) dP(\rho) = \int \Phi(|u|^2) dQ(u).$$

*Proof.* The monomial functions  $\Phi$  of the form (see [6, Section 1.7])

$$\Phi(\rho) = M_{k, \varphi_k}(\rho) = \int_{\mathbb{R}^{dk}} \varphi_k \rho^{\otimes k}$$

with  $k \in \mathbb{N}$  and  $\varphi_k$  bounded continuous over  $\mathbb{R}^{dk}$  generate a subalgebra of the continuous functions on the space of probability measures. This subalgebra is dense by the Stone-Weierstrass theorem and it thus suffices to test the claim against all the above monomials.

We identify the function  $\varphi_k$  with the multiplication operator thereby to write

$$\begin{aligned} \int M_{k, \varphi_k}(\rho) dP(\rho) &= \int_{\mathbb{R}^{dk}} \varphi_k \mu^{(k)} \\ &= \int_{\mathbb{R}^{dN}} \varphi_k(x_1, \dots, x_k) \mu^{(N)}(x_1, \dots, x_N) dx_1 \dots dx_N \\ &= \text{Tr} \left( \varphi_k \otimes \mathbf{1}^{\otimes(N-k)} \Gamma_N \right) \\ &= \text{Tr} \left( \varphi_k \Gamma_N^{(k)} \right) \rightarrow \text{Tr} \left( \varphi_k \gamma^{(k)} \right) \end{aligned}$$

using Equations (2.11)-(2.7)-(3.1)-(3.3) and, in the last step, the fact that multiplication by a bounded function is a bounded operator to pass to the limit using (3.10). Since the left-hand side actually does not depend on  $N$  we deduce

$$\int M_{k, \varphi_k}(\rho) dP(\rho) = \text{Tr} \left( \varphi_k \gamma^{(k)} \right).$$

As per (3.11) this implies the desired

$$\int M_{k, \varphi_k}(\rho) dP(\rho) = \int M_{k, \varphi_k}(|u|^2) dQ(u).$$

This being true for all  $k$  and  $\varphi_k$ , the proofs of the lemma and the theorem (approximating  $\rho \mapsto \langle \sqrt{\rho} |h| \sqrt{\rho} \rangle$  by a sequence of continuous functions) are both complete.  $\square$

## REFERENCES

- [1] J. BOURGAIN, H. BRÉZIS, AND P. MIRONESCU, *Another look at Sobolev spaces*, in *Optimal control and Partial Differential equations*, IOS Press, 2001, pp. 439–455.
- [2] ———, *Limiting embedding theorems for  $W^{s,p}$  when  $s \uparrow 1$  and applications*, *J. Anal. Math.*, 87 (2002), pp. 77–101.
- [3] E. CARLEN, *Superadditivity of Fisher’s information and logarithmic Sobolev inequalities*, *J. Funct. Anal.*, 101 (1991), pp. 194–211.
- [4] G. DELL’ANTONIO, *On the limits of sequences of normal states*, *Comm. Pure Appl. Math.*, 20 (1967), p. 413.
- [5] N. FOURNIER, M. HAURAY, AND S. MISCHLER, *Propagation of chaos for the 2d viscous vortex model*, *J. Eur. Math. Soc.*, 16 (2014), pp. 1423–1466.
- [6] F. GOLSE, *On the Dynamics of Large Particle Systems in the Mean Field Limit*, ArXiv e-prints 1301.5494, (2013). Lecture notes for a course at the NDNS+ Applied Dynamical Systems Summer School “Macroscopic and large scale phenomena”, Universiteit Twente, Enschede (The Netherlands).
- [7] M. HAURAY, *Limite de champ moyen et propagation du chaos pour des systèmes de particules, limites gyro-cintique et quasi-neutre pour les plasmas*. Habilitation thesis, 2014.
- [8] M. HAURAY AND S. MISCHLER, *On Kac’s chaos and related problems*, *J. Func. Anal.*, 266 (2014), pp. 6055–6157.
- [9] E. HEWITT AND L. J. SAVAGE, *Symmetric measures on Cartesian products*, *Trans. Amer. Math. Soc.*, 80 (1955), pp. 470–501.
- [10] M. HOFFMANN-OSTENHOF AND T. HOFFMANN-OSTENHOF, *Schrödinger inequalities and asymptotic behavior of the electron density of atoms and molecules*, *Phys. Rev. A*, 16 (1977), pp. 1782–1785.
- [11] R. L. HUDSON AND G. R. MOODY, *Locally normal symmetric states and an analogue of de Finetti’s theorem*, *Z. Wahrscheinlichkeitstheor. und Verw. Gebiete*, 33 (1975/76), pp. 343–351.
- [12] M. K.-H. KIESSLING, *The Hartree limit of Born’s ensemble for the ground state of a bosonic atom or ion*, *J. Math. Phys.*, 53 (2012), p. 095223.
- [13] M. LEWIN, *Mean-Field limit of Bose systems: rigorous results*, arXiv:1510.04407. Proceedings of the International Congress of Mathematical Physics, 2015
- [14] M. LEWIN, P. NAM, AND N. ROUGERIE, *Derivation of Hartree’s theory for generic mean-field Bose systems*, *Adv. Math.*, 254 (2014), pp. 570–621.
- [15] E. H. LIEB AND M. LOSS, *Analysis*, vol. 14 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2nd ed., 2001.
- [16] W. MASJA AND J. NAGEL, *Über äquivalente normierung der anisotropen Funktionalräume  $H^\mu(\mathbb{R}^n)$* , *Beiträge zur Analysis*, 12 (1978), pp. 7–17.
- [17] V. MAZ’YA AND T. SHAPOSHNIKOVA, *On the Bourgain, Brezis, and Mironescu theorem concerning limiting embeddings of fractional Sobolev spaces*, *J. Func. Anal.*, 195 (2002), pp. 230–238.
- [18] D. ROBINSON AND D. RUEELLE, *Mean entropy of states in classical statistical mechanics*, *Commun. Math. Phys.*, 5 (1967), pp. 288–300.
- [19] N. ROUGERIE, *De Finetti theorems, mean-field limits and Bose-Einstein condensation*. arXiv:1506.05263, 2014. LMU lecture notes.
- [20] ———, *Théorèmes de De Finetti, limites de champ moyen et condensation de Bose-Einstein*, Les cours Peccot, Spartacus IDH, Paris, 2016. Cours Peccot, Collège de France : février-mars 2014.
- [21] S. SALEM, *Propagation of chaos for fractional Keller Segel equations in diffusion dominated and fair competition cases*, *Journal de Mathématiques Pures et Appliquées*, (2019).
- [22] ———, *Propagation of chaos for the boltzmann equation with moderately soft potentials*. arXiv:1910.01883, 2019. in preparation.
- [23] R. SCHATTEN, *Norm Ideals of Completely Continuous Operators*, vol. 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete, Folge*, 1960.
- [24] B. SIMON, *Trace ideals and their applications*, vol. 35 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, 1979.
- [25] G. TOSCANI, *The fractional Fisher information and the central limit theorem for stable laws*, *Ric. Mat.* 65, 65 (2016), pp. 71–91.
- [26] ———, *The information-theoretic meaning of Gagliardo-Nirenberg type inequalities*, (2018).

[27] ———, *Score functions, generalized relative Fisher information and applications*, (2018).

UNIVERSITÉ GRENOBLE-ALPES & CNRS, LPMMC (UMR 5493), B.P. 166, F-38042 GRENOBLE,  
FRANCE

*E-mail address:* `nicolas.rougerie@grenoble.cnrs.fr`