

Article

Entropy of the Canonical Occupancy (Macro) State in the Quantum Measurement Theory

Arnaldo Spalvieri 

Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy; arnaldo.spalvieri@polimi.it

Abstract: The paper analyzes the probability distribution of the occupancy numbers and the entropy of a system at the equilibrium composed by an arbitrary number of non-interacting bosons. The probability distribution is obtained through two approaches: one involves tracing out the environment from a bosonic eigenstate of the combined environment and system of interest (the empirical approach), while the other involves tracing out the environment from the mixed state of the combined environment and system of interest (the Bayesian approach). In the thermodynamic limit, the two coincide and are equal to the multinomial distribution. Furthermore, the paper proposes to identify the physical entropy of the bosonic system with the Shannon entropy of the occupancy numbers, fixing certain contradictions that arise in the classical analysis of thermodynamic entropy. Finally, by leveraging an information-theoretic inequality between the entropy of the multinomial distribution and the entropy of the multivariate hypergeometric distribution, Bayesianism of information theory and empiricism of statistical mechanics are integrated into a common “infomechanical” framework.

Keywords: occupancy numbers; multivariate hypergeometric distribution; multinomial distribution; canonical typicality; Gibbs correction factor; Sackur-Tetrode entropy formula



Citation: Spalvieri, A. Entropy of the Canonical Occupancy (Macro) State in the Quantum Measurement Theory. *Entropy* **2024**, *26*, 107. <https://doi.org/10.3390/e26020107>

Academic Editors: Carlo Cafaro and Roberto Zivieri

Received: 13 December 2023

Revised: 10 January 2024

Accepted: 22 January 2024

Published: 24 January 2024



Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The concept of physical entropy is controversial since the times of Boltzmann and Gibbs. The subsequent excerpt from an interview with Shannon, whose crucial contribution to the comprehension of entropy is now widely acknowledged not only in physics [1] but also in diverse fields such as medicine, seismology, and finance (refer to [2–4] for recent applications of Shannon entropy in these domains), is available in [5]: “My greatest concern was what to call it. I thought of calling it ‘information,’ but the word was overly used, so I decided to call it ‘uncertainty.’ When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, “You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one really knows what entropy really is, so in a debate you will always have the advantage.””

Over the years, the number of different interpretations and definitions of entropy has grown, a recent collection of heterogeneous “entropies” being reported in [6].

In this fragmented and ambiguous context, one of the most debated points is the relationship between information and physical entropy. The controversy about the role of information in physics has its roots in the famous thought experiments of Maxwell and Szilard, and is still an object of discussion today. From the one side, Landauer claims in [7] that

Information is a physical entity.

From the other side, Maroney writes in his thesis [8] that

The Szilard Engine is unsuccessful as a paradigm of the information-entropy link, and Maroney and Timpson reiterate the same concept in [9]:

rejecting the claims that information is physical provides a better basis for understanding the fertile relationship between information theory and physics.

Independently of the different opinions and arguments, it is a matter of fact that entropy is the key quantity that allows us to derive all the thermodynamic macro-properties of many-particle systems from the statistical properties of the constituent microscopic particles.

One standard approach to entropy in the classical (non-quantum) regime is based on the Gibbs entropy of microstates, which, for systems at equilibrium with the environment, is the Shannon entropy of the probability distribution resulting from entropy maximization under the constraints imposed on the system by the environment that surrounds it. The principle of maximization of entropy of microstates, often referred to as MaxEnt, dates back to the seminal paper of Jaynes [10] and inspired Shannon's information-theoretic view of entropy. Regrettably, the probability distribution of microstates does not account for indistinguishability of particles. At the same time, indistinguishability of particles is essential; therefore, it is imperative to incorporate it into the framework. This is typically accomplished by subtracting from the entropy of microstates the Gibbsian correction term $\log(N!)$, where N is the number of the system's particles.

The Gibbsian correction term enters the scene almost like a sudden intervention, akin to a *deus ex machina*, but the subtraction of $\log(N!)$ generates a glaring contradiction: when the entropy of microstates is smaller than $\log(N!)$, the difference between the entropy of microstates and $\log(N!)$ becomes negative, as it actually happens to the Sackur–Tetrode entropy formula at a low temperature-to-density ratio. Since the entropy of any random variable is guaranteed to always be non-negative (this statement is, according to [11], a formulation of the third law of thermodynamics), entropy of microstates minus $\log(N!)$ cannot be entropy. This pathology is not surprising since, after the subtraction of $\log(N!)$, the random variable for which the "entropy" is calculated is no longer specified. Consequently, "entropy" transforms into a mere formula representing an unspecified physical system. In this controversial situation, more than 100 years after its introduction, the Gibbsian $\log(N!)$ is still an object of research and debate today; see, e.g., [11–15], see also [16] for the case of small thermodynamic systems.

This paper recovers coherency of the picture by considering, in place of the classical setting, the more general quantum setting, where entropy is the von Neumann entropy of the mixed state of the system. The mixed state of a system is commonly represented by a density operator living in the Hilbert space $\mathcal{H}^{\otimes N}$ obtained as the tensor product N times of the single-particle Hilbert space \mathcal{H} . Exactly as it happens with the entropy of microstates, the von Neumann entropy of this density operator is appropriate for systems of distinguishable particles, but it does not capture the indistinguishability of particles. We observe that, to capture the indistinguishability of particles of a bosonic system, i.e., a system whose state remains unchanged when particles' indexes are permuted, an eigenbasis that conveniently represents its quantum state is that of the vectors of the occupancy numbers, giving rise to the so-called second quantization formalism. The consideration of indistinguishability of particles and, with it, of the occupancy numbers, leads us to define the bosonic density operator, i.e., the density operator in a system's bosonic subspace. Due to the inherent capability of the bosonic density operator to capture the indistinguishability of particles, its von Neumann entropy is the entropy of the system. This allows us to surpass the standard approach to entropy based on distinguishable particles and resolves the puzzle of the Gibbsian $\log(N!)$, which, as shown in this paper, now enters the scene through the main entrance alongside other quantum correction terms ensuring non-negativity of entropy.

Our exploration of the quantum approach is situated within the framework of contemporary quantum thermodynamics, whose origins can be attributed to [17,18]. The two cited papers, which have deeply influenced the successive literature, e.g., [19–23], derive the mixed state of a system at equilibrium with the environment by tracing out the environment from the universe, which, in the thermal case, is the union of the system and the heat bath that thermalizes the system. The analysis of [17,18] shows that, as the number of particles

of the universe tends to infinity, the mixed state of the system is the same, at least in the weak sense, both if the universe is in a pure state or in a mixed state. Purity of the state of the universe makes the introduction of a statistical ensemble of “universes” unnecessary, thus overcoming the subjectivism that is inherent in the Bayesian approach.

The outline of the paper is as follows. Section 2 introduces the bosonic Hilbert subspace and its eigenbasis. In Section 3, we derive the bosonic density operator for the system under consideration by tracing out the environment from the universe, assuming that the universe is in a bosonic eigenstate. With this assumption, we show that the probability distribution that weights the projectors of the bosonic Hilbert subspace of the system is the multivariate hypergeometric distribution. As the number of bosons of the universe tends to infinity, the multivariate hypergeometric distribution converges to the multinomial distribution, which we identify as the canonical distribution of the occupancy numbers. Section 4 shows that, if, as in the Bayesian approach, the universe is assumed to be in a mixed bosonic state, then the distribution of the occupancy numbers of the system is multinomial provided that the occupancy numbers of the universe are multinomially distributed too. Section 5 discusses the application of the mentioned probability distributions to the entropy of physical systems and places our work within the framework of quantum information theory, unveiling an engaging connection between Bayesianism and empiricism in physics. To illustrate the intrinsic capacity of the canonical bosonic density operator to capture the indistinguishability of particles in systems at equilibrium, Section 6 shows that its von Neumann entropy fits the entropy of the ideal gas in a container. Section 7 sketches the future application of our approach to the Szilard engine. Finally, in Section 8 we draw the conclusions.

2. The Bosonic Hilbert Subspace

Let $\{c \in \mathbb{C}\}$, $\mathbb{C} = \{1, 2, \dots, |\mathbb{C}|\}$, be the set of quantum numbers (the colors) allowed to a boson, let $\{\bar{c} = (c_1, c_2, \dots, c_N)\}$ (the overline denotes vectors and their size is not explicitly mentioned for brevity when it can be inferred from the context) be the set of microstates of a system made by N non-interacting bosons (the balls), and let $\{|\bar{c}\rangle\}$ be the complete set of eigenstates that span the Hilbert space $\mathcal{H}^{\otimes N}$ of the system. Let $\delta(\cdot)$ be the indicator function and let

$$n_c(c_1, c_2, \dots, c_N) \stackrel{\text{def}}{=} \sum_{i=1}^N \delta(c - c_i), \quad c = 1, 2, \dots, |\mathbb{C}|,$$

be the number of bosons whose quantum number is c . In the following, the dependency on (c_1, c_2, \dots, c_N) of $n_c(c_1, c_2, \dots, c_N)$ will be omitted when possible and the vector of the occupancy numbers $\bar{n} = (n_1, n_2, \dots, n_{|\mathbb{C}|})$ will be called occupancy macrostate. The quantum state of the system of indistinguishable bosons is

$$|\bar{n}\rangle = \sum_{\bar{c} \in \mathbb{C}^N(\bar{n})} \frac{|\bar{c}\rangle}{\sqrt{W(\bar{n})}}; \tag{1}$$

see, e.g., [24] and chapter 7 of [25]. In the above equation, $\mathbb{C}^N(\bar{n})$ is the set of the $W(\bar{n})$ equiprobable microstates whose occupancy macrostate is \bar{n} and $W(\bar{n})$ is the multinomial coefficient,

$$W(\bar{n}) = \begin{cases} \frac{N!}{\prod_{c=1}^{|\mathbb{C}|} n_c!}, & n_c \geq 0, \forall c \in \mathbb{C}, \\ 0, & \text{elsewhere,} \end{cases}$$

i.e., the number of distinct permutations of the entries of any one of the elements of $\mathbb{C}^N(\bar{n})$.

The set $\{|\bar{n}\rangle, \bar{n} \in \mathbb{N}\}$, with

$$|\mathbb{N}| = \binom{N + |\mathbb{C}| - 1}{|\mathbb{C}| - 1}$$

(see [26]) is a complete set of bosonic eigenstates for the bosonic Hilbert subspace of the Hilbert space of the system.

3. Empirical Approach

We want to obtain the system’s bosonic density operator $\hat{\nu}_{\bar{u}}$, where \bar{u} is the occupancy macrostate of the universe, tracing out the environment made by $U - N$ bosons from the projector $|\bar{u}\rangle \langle \bar{u}|$ of the universe made by U bosons. The assumption that the universe is in an eigenstate appears to be plausible, especially for large values of U and within the context of the Eigenstate Thermalization Hypothesis (ETH) [27]. The first step is to write $|\bar{u}\rangle$ in the form of bosonic purification:

$$\begin{aligned} |\bar{u}\rangle &= \sum_{\bar{n} \in \mathbb{N}} \sum_{\bar{c}_1^N \in \mathbb{C}^N(\bar{n}), \bar{c}_{N+1}^U \in \mathbb{C}^{U-N}(\bar{u}-\bar{n})} \frac{|\bar{c}_1^N, \bar{c}_{N+1}^U\rangle}{\sqrt{W(\bar{u})}} \\ &= \sum_{\bar{n} \in \mathbb{N}} \sqrt{\frac{W(\bar{n})W(\bar{u}-\bar{n})}{W(\bar{u})}} |\bar{n}, \bar{u}-\bar{n}\rangle, \end{aligned}$$

where \bar{c}_i^j represents the windowing of vector \bar{c} between i and j and it is understood that the set $\mathbb{C}^N(\bar{n})$ ($\mathbb{C}^{U-N}(\bar{u}-\bar{n})$) is empty when one or more entries of \bar{n} ($\bar{u}-\bar{n}$) is negative. Observing that

$$\langle \bar{u}-\bar{n} | \bar{u}-\bar{n}' \rangle |\bar{n}\rangle \langle \bar{n}'| = \begin{cases} |\bar{n}\rangle \langle \bar{n}'|, & \bar{u}-\bar{n} = \bar{u}-\bar{n}', \\ 0, & \bar{u}-\bar{n} \neq \bar{u}-\bar{n}', \end{cases}$$

and that the condition $\bar{u}-\bar{n} = \bar{u}-\bar{n}'$ forces $\bar{n} = \bar{n}'$, we conclude that

$$\begin{aligned} \hat{\nu}_{\bar{u}} &= \text{Tr}_{\mathcal{H}^{\otimes(U-N)}} |\bar{u}\rangle \langle \bar{u}| \\ &= \sum_{\bar{n} \in \mathbb{N}} \frac{W(\bar{u}-\bar{n})W(\bar{n})}{W(\bar{u})} |\bar{n}\rangle \langle \bar{n}|, \end{aligned} \tag{2}$$

where $\text{Tr}_{\mathcal{H}^{\otimes(U-N)}}$ is the operator that traces out the Hilbert space of the environment. If the set $\mathbb{C}^N(\bar{n})$ ($\mathbb{C}^{U-N}(\bar{u}-\bar{n})$) is empty, then one or more entries of \bar{n} ($\bar{u}-\bar{n}$) is negative; in which case, the multinomial coefficient $W(\bar{n})$ ($W(\bar{u}-\bar{n})$) is zero by definition. The fraction appearing in (2) is the multivariate hypergeometric distribution, which is the distribution of the occupancy numbers of colors in drawing without replacement N balls out of an urn containing U balls with color occupancy numbers \bar{u} . In many textbooks and papers, the multivariate hypergeometric distribution is expressed by binomial coefficients, as in the equation

$$P_{\bar{u}}(\bar{n}) = \frac{W(\bar{u}-\bar{n})W(\bar{n})}{W(\bar{u})} = \frac{\prod_{c=1}^{|\mathbb{C}|} \binom{u_c}{n_c}}{\binom{U}{N}}, \tag{3}$$

where $P(\cdot)$ is the probability of the random variable inside the round brackets and \bar{u} in the subscript means that the probability distribution $\{P_{\bar{u}}(\bar{n})\}$ depends on the vector of known parameters \bar{u} .

In the thermodynamic limit, the multivariate hypergeometric distribution converges to the multinomial distribution, i.e., the distribution of the occupancy numbers of colors

in drawing with replacement N times a ball out of an urn containing colored balls with relative frequency of color c equal to $P(c)$:

$$\begin{aligned} \lim_{U \rightarrow \infty} P_{\bar{u}}(\bar{n}) &= W(\bar{n}) \lim_{U \rightarrow \infty} \frac{W(\bar{u} - \bar{n})}{W(\bar{u})} \\ &= W(\bar{n}) \lim_{U \rightarrow \infty} \prod_{c=1}^{|\mathcal{C}|} (U^{-1} u_c)^{n_c} \end{aligned} \tag{4}$$

$$= W(\bar{n}) \prod_{c=1}^{|\mathcal{C}|} (P(c))^{n_c}, \tag{5}$$

$$\lim_{U \rightarrow \infty} \hat{v}_{\bar{u}} = \hat{v} = \sum_{\bar{n} \in \mathbb{N}} W(\bar{n}) \prod_{c=1}^{|\mathcal{C}|} (P(c))^{n_c} |\bar{n}\rangle \langle \bar{n}|,$$

where (4) is Stirling’s formula, while, by regarding \bar{u} as the result of a PVM operated on the universe, Equation (5) is recognized to be the Law of Large Numbers (LLN) for the empirical one-particle distribution. Concentration inequalities that bound the probability of deviations of the empirical probability from its expectation, that is the probability of occurrence of non-typical bosonic eigenstates of the universe, can be found in [28] for hypergeometrically distributed random variables, in [29] for multinomially distributed random vectors. Paper [30] demonstrates that the multinomial distribution is the maximum entropy distribution of the occupancy numbers with constrained one-particle distribution. See also [31,32] for the multinomial distribution in statistical mechanics.

By equiprobability of the disjoint microstates belonging to the same occupancy macrostate, we see that, in (3),

$$\frac{W(\bar{u} - \bar{n}(c_1, c_2, \dots, c_N))}{W(\bar{u})} = P_{\bar{u}}(c_1, c_2, \dots, c_N).$$

Clearly

$$\begin{aligned} \lim_{U \rightarrow \infty} P_{\bar{u}}(c_1, c_2, \dots, c_N) &= \prod_{c=1}^{|\mathcal{C}|} (P(c))^{n_c(c_1, c_2, \dots, c_N)} \\ &= \prod_{i=1}^N P(c_i), \end{aligned}$$

which explicitly shows that, in the thermodynamic limit, microstates outcoming from the mentioned PVM become independent and identically distributed random variables. What happens is that the LLN cancels the dependencies inside the joint distribution of microstates induced by the constraints imposed on the colors of the N balls by the occupancy numbers \bar{u} of the universe in drawing without replacement. Furthermore, the probability distribution $\{\prod_{i=1}^N P(c_i)\}$ also becomes independent of the specific result \bar{u} of the PVM operated on the universe, in the sense that the empirical one-particle distribution $\{U^{-1} u_c\}$ for $U \rightarrow \infty$ is the same for almost every bosonic eigenstate of the universe, i.e., for the typical bosonic eigenstates of the universe. The one-particle distribution and, with it, the product $\{\prod_{i=1}^N P(c_i)\}$, will depend on the physical constraints imposed on the universe. When the constraint is the temperature, it is widely recognized that $\{\prod_{i=1}^N P(c_i)\}$ is the Boltzmann distribution, which can be found by Jaynes’ constrained maximization of entropy [10].

Independency and identical distribution of the eigenstates of the individual bosons of the system lead to the following density operator in $\mathcal{H}^{\otimes N}$

$$\hat{\rho} = \sum_{\vec{c} \in \mathcal{C}^N} \prod_{i=1}^N P(c_i) |\vec{c}\rangle \langle \vec{c}|,$$

which we claim to be the canonical density operator in $\mathcal{H}^{\otimes N}$. Papers [17,18] claim that the system’s density operator in $\mathcal{H}^{\otimes N}$ converges to the canonical density operator as $U \rightarrow \infty$ for almost every pure state of the universe, i.e., for the typical pure states of the universe. For this reason convergence to the canonical state is called *canonical typicality* in [18]. Typicality, which can be intended in various senses, is often invoked in statistical mechanics. For instance, paper [33] uses the properties of the information-theoretic typical set to characterize weak convergence of microstates to equiprobability in the context of classical statistical mechanics. We point out that the claim of [17,18] is compatible with our claim of convergence to the bosonic canonical state if we do not pretend that typical bosonic eigenstates are typical states, as illustrated in Figure 1.

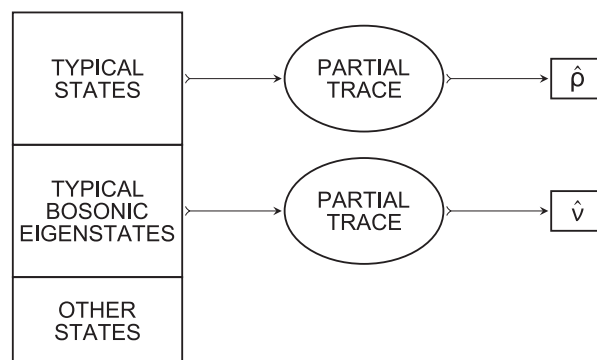


Figure 1. The box on the left represents the pure states of the universe. Although in the thermodynamic limit almost every pure state of the universe is typical, still there is space for infinitely many typical bosonic eigenstates and for infinitely many other pure states that are neither typical states nor typical bosonic eigenstates, among which are all the eigenstates of $\mathcal{H}^{\otimes U}$.

4. Bayesian Approach

In the Bayesian approach, the vector of known parameters \vec{u} becomes a random vector whose probability distribution, the Bayesian prior, characterizes the bosonic density operator of the universe. In the Bayesian setting, the hypergeometric multivariate distribution of the density operator $\hat{v}_{\vec{u}}$ of (2) is the Bayesian likelihood, i.e., a conditional distribution where the random condition is the random vector of parameters \vec{u} . The probability distribution of the density operator of the system, the Bayesian marginal, is obtained by tracing out the environment from the bosonic density operator of the universe. The Bayesian approach is self-consistent when the prior is multinomial because the marginal turns out to be multinomial too, regardless of the number of particles of the universe:

$$\begin{aligned} \hat{v} &= \sum_{\vec{u} \in \mathcal{U}} W(\vec{u}) \prod_{c=1}^{|\mathcal{C}|} (P(c))^{u_c} \text{Tr}_{\mathcal{H}^{\otimes (U-N)}} |\vec{u}\rangle \langle \vec{u}| \\ &= \sum_{\vec{u} \in \mathcal{U}} W(\vec{u}) \prod_{c=1}^{|\mathcal{C}|} (P(c))^{u_c} \hat{v}_{\vec{u}} \\ &= \sum_{\vec{n} \in \mathbb{N}} W(\vec{n}) \prod_{c=1}^{|\mathcal{C}|} (P(c))^{n_c} |\vec{n}\rangle \langle \vec{n}|, \end{aligned} \tag{6}$$

where the third equality is obtained by substituting (2) in the second line and manipulating. The comparison between the Bayesian (6) and the empirical (2) clarifies the relationship between the two approaches. The Bayesian approach captures the pre-measurement randomness of the mixed state of the universe by a mathematical procedure based on constrained entropy maximization whose result is the multinomial prior of the universe. Instead, the empirical approach renounces to mathematics and imports the result \bar{u} of the measurement of the occupancy numbers of the universe in the multivariate hypergeometric distribution of the occupancy numbers of a system that is separated from the universe after the measurement.

5. Quantum Entropy and Quantum Information

Since $\hat{\nu}_{\bar{u}}$ ($\hat{\nu}$) is an eigendecomposition, its von Neumann entropy is equal to the Shannon entropy of the multivariate hypergeometric (multinomial) distribution. Specifically, when the density operator is an eigendecomposition, the associated density matrix is diagonal, leading to

$$S(\hat{\nu}) \stackrel{\text{def}}{=} -\text{Tr}(\hat{\nu} \log(\hat{\nu})) = - \sum_{\bar{n} \in \mathbb{N}} P(\bar{n}) \log(P(\bar{n})),$$

where $S(\cdot)$ is the von Neumann entropy of the density operator inside the round brackets and the rightmost term is the Shannon entropy. For this reason, in the following, we make no distinction between the von Neumann entropy and the Shannon entropy, calling it simply “entropy”. It is worth emphasizing that, here, we completely skip the notion of phase space, leading to the exact probability distribution of the quantum occupancy numbers, and, as a consequence, to the exact entropy. Conversely, the standard phase space approach inherently leads to approximations of entropy. These approximations require improvement at low temperature/density ratios, see e.g., [34], while still retaining some approximations. See also [11] for the role of the multinomial coefficient in the solution of the Gibbs paradox in the phase-space approach.

The entropy of the entanglement between the environment and the system of interest is the entropy of the multivariate hypergeometric distribution, which, owing to the distribution’s symmetry, is equal to the entropy of both the system and the environment after their separation. It is

$$\begin{aligned} S(\hat{\nu}_u) &= \log(W(\bar{u})) - E\{\log(W(\bar{u} - \bar{n}))\} - E\{\log(W(\bar{n}))\} \\ &= -E\{\log(P_{\bar{u}}(\bar{c}))\} - \log(N!) + \sum_{c=1}^{|\mathcal{C}|} E\{\log(n_c!)\}, \end{aligned} \quad (7)$$

where $E\{\cdot\}$ is the classical expectation computed over the probability distribution of the random variable inside the expectation, the base of the logarithm is Euler’s number, and the Boltzmann constant in front of the logarithm is omitted for brevity. The term $-E\{\log(P_{\bar{u}}(\bar{c}))\}$ is the Gibbs entropy of microstates, i.e., the entropy of the system of distinguishable particles, while the term $-E\{\log(W(\bar{n}))\}$ is due to indistinguishability of particles, which, in the average case, prevents access to $E\{\log(W(\bar{n}))\}$ units of information. Since $(W(\bar{n}))^{-1}$ is the conditional probability of microstates given the macrostate, its expectation is the conditional Shannon entropy of microstates given the macrostate; so, $S(\hat{\nu}_u)$ is the mutual information between microstates and macrostates. The term $\log(N!)$ in (7) was introduced by Gibbs to force compatibility between the non-quantized phase-space (differential) entropy of microstates and the physical entropy of systems of indistinguishable particles. We observe that, while the probability that two or more classical particles have the same position and momentum in the phase space is zero because position and momentum are dense variables, the probability that two or more quantum particles occupy the same quantum state is not zero. This non-zero probability is captured by the sum of expectations in (7). As the entropy of microstates becomes lower and lower, this sum moves closer and

closer to $\log(N!)$, until becoming equal to $\log(N!)$ when all the particles occupy the ground state. This also prevents the system's entropy from becoming negative when the entropy of microstates becomes vanishingly small.

In the canonical case, the entropy of the bosonic density operator is the entropy of the multinomial distribution:

$$S(\hat{\nu}) = -NE\{\log(P(c))\} - E\{\log(W(\bar{n}))\}. \tag{8}$$

See [35,36] for the calculation of the above entropy and see [32] for approximations of the entropy of the multinomial distribution in the context of statistical mechanics. Equation (8) is Equation (11) of [32], where the authors call the entropy of the distribution of the occupancy numbers *entropy fluctuations*. Apart from certain exceptions, the authors of [32] consider these entropy fluctuations negligible compared to the entropy of the system, failing to recognize that the entropy of the occupancy numbers is the entropy of a system of indistinguishable particles.

Let us consider the thermal state. The one-particle Boltzmann distribution for a system at the thermal equilibrium at temperature T Kelvin degrees is

$$P(c) = Z^{-1}e^{-\epsilon_c/k_B T},$$

where ϵ_c is the energy eigenvalue of eigenstate $|c\rangle$, $k_B = 1.38 \cdot 10^{-23}$ J/K is the Boltzmann constant and Z is the one-particle partition function

$$Z = \sum_{c=1}^{|\mathcal{C}|} e^{-\epsilon_c/k_B T},$$

where $|\mathcal{C}|$ is large enough to make the impact of the truncation of the sum negligible. In the non-quantum regime, where the approximation

$$E\{\log(W(\bar{n}))\} \approx \log(N!)$$

holds, the entropy $S(\hat{\nu})$ of (8) fits the classical relationship between thermodynamic entropy and temperature:

$$\begin{aligned} \frac{\partial S(\hat{\nu})}{N\partial E\{\epsilon_c\}} &\approx -\frac{\partial}{\partial E\{\epsilon_c\}} E\{\log(P(c))\} \\ &= \frac{\partial}{\partial E\{\epsilon_c\}} E\{\log(Ze^{\epsilon_c/k_B T})\} \\ &= \frac{1}{k_B T} \end{aligned} \tag{9}$$

where it easy to see that

$$\frac{\partial \log(Z)}{\partial E\{\epsilon_c\}} = -E\{\epsilon_c\} \frac{\partial}{\partial E\{\epsilon_c\}} \frac{1}{k_B T}.$$

In the impossibility, for the reasons already discussed, to take the quantum entropy of microstates $S(\hat{\rho})$ and/or $S(\hat{\rho}) - \log(N!)$ as candidates to the role of quantum thermodynamic entropy, the right hand side of (8), despite the deviation of its partial derivative w.r.t. the expected energy from (9) in the quantum regime, is a solid candidate for the quantum thermodynamic entropy. Actually, the equality between the partial derivative of the classical thermodynamic entropy w.r.t. the expected energy and (9) is a genuinely classical equality that, due to the quantum effects that become important at a low temperature, can be no more true when the quantum thermodynamic entropy is considered. Also, the expectation of the logarithm of the multinomial coefficient can make (8) non-extensive when the probability that two or more particles occupy the same quantum state is non-negligible;

see [37] for non-extensivity of (8) in the case of a monoatomic gas at thermal equilibrium in a container. Again, extensivity of thermodynamic entropy is a genuinely classical concept, so we should not be surprised that there is a lack of extensivity when quantum effects are non-negligible.

A digression is now necessary in order to delve into the relationship between our entropy of macrostates expressed by Equations (7) and (8), and the observational entropy investigated in [6] and in other papers of the same author published from 2019 to the present day, including the recent [38]. Even if the authors of [6] attribute to their observational entropy a meaning that is clearly different from the meaning of our entropy of macrostates, to a certain extent, the two risk being confused, at least formally, because the observational entropy includes a term that is exactly the same as that for our entropy of macrostates. Precisely to avert this risk, we hereafter show that, in our setting, the observational entropy is equal to the Gibbs entropy of microstates. Therefore, its meaning cannot be that of the entropy of macrostates expressed by Equations (7) and (8). Hereafter, the vector (i_1, i_2, \dots, i_n) appearing in the definition of observational entropy of multiple coarse graining given in Equation (19) of [6] is identified with our $\bar{n} = (n_1, n_2, \dots, n_{|C|})$. Thus, the observational entropy defined in Equation (19) of [6] is

$$-\sum_{\bar{n}} p_{\bar{n}} \log\left(\frac{p_{\bar{n}}}{V_{\bar{n}}}\right). \tag{10}$$

It is formally admissible and it seems aligned with the spirit of [6] to take for the projector $\hat{P}_{n_1} \hat{P}_{n_2} \dots \hat{P}_{n_{|C|}}$ of [6]

$$\hat{P}_{n_1} \hat{P}_{n_2} \dots \hat{P}_{n_{|C|}} = \hat{P}_{\bar{n}} = \sum_{\bar{c} \in \mathbb{C}^N(\bar{n})} |\bar{c}\rangle \langle \bar{c}|,$$

leading to

$$\hat{P}_{\bar{n}} \hat{P}_{\bar{n}} = \hat{P}_{\bar{n}}$$

and, by the definition of $V_{\bar{n}}$ given in [6], to

$$V_{\bar{n}} = \text{Tr}(\hat{P}_{\bar{n}}) = W(\bar{n}),$$

that, in full accordance with the spirit of [6], is the dimension of the subspace of $\mathcal{H}^{\otimes N}$ occupied by the macrostate \bar{n} . Using the definition of the probability $p_{\bar{n}}$ of macrostate \bar{n} given in [6], with

$$\hat{\rho} = \sum_{\bar{c} \in \mathbb{C}^N} P(\bar{c}) |\bar{c}\rangle \langle \bar{c}|,$$

we get

$$p_{\bar{n}} = \text{Tr}(\hat{P}_{\bar{n}} \hat{\rho} \hat{P}_{\bar{n}}) = W(\bar{n}) P(\bar{c}), \quad \forall \bar{c} \in \mathbb{C}^N(\bar{n}),$$

which is equal to the probability of macrostate \bar{n} proposed in this paper and is actually an alternative way to define it, though. Note the difference; we derived it from the partial trace operated on the eigenstate of the universe and/or from the MaxEnt principle. Plugging $p_{\bar{n}}$ and $V_{\bar{n}}$ in (10), from our Equations (7) and/or (8), we conclude that the observational entropy (10) is equal to the Gibbs entropy of microstates. This is enough to avert the risk of confusion between our entropy of macrostates and the observational entropy, because the only case where the entropy of macrostates is equal to the observational entropy occurs when $V_{\bar{n}} = 1$ for all the vectors \bar{n} , i.e., when particles are distinguishable and the entropy of macrostates is the Gibbs entropy of microstates.

The following inequalities sandwich the Boltzmann entropy $\log(W(E\{\bar{n}\}))$ between the two terms in the right-hand side of (8):

$$\begin{aligned} -NE\{\log(P(c))\} &\geq \log(W(E\{\bar{n}\})) \\ &\geq E\{\log(W(\bar{n}))\}, \end{aligned} \tag{11}$$

where, with some abuse of notation, the factorials of the real numbers in the denominator of $W(E\{\bar{n}\})$ are intended as $x! = \Gamma(x + 1)$, where $\Gamma(\cdot)$ is the Gamma function. The first inequality is (11.22) of [26], the second inequality is obtained by applying the Jensen inequality

$$E\{f(n_c)\} \geq f(E\{n_c\}), \forall c \in \mathbb{C},$$

to the convex (upward) function $f(n_c) = \log(n_c!)$. In statistical mechanics, it is standard to derive from Stirling’s formula an approximation between the two terms of (11).

Note that, if we pretend that entropy is a variable of state, then the probability distribution of the occupancy numbers must depend only on the state of the system. However, the multivariate hypergeometric distribution depends also on the state of the universe. The dependency becomes progressively weaker as the number of particles in the universe approaches infinity. However, it remains evident that this makes the empirical approach incompatible with the notion of entropy as a state variable. In conclusion, entropy can only be considered a variable of state if we embrace the Bayesian approach.

We hereafter introduce the “empirical” information, sketching a new engaging connection between the Bayesian and the empirical approaches to information in quantum measurement theory. Let us regard the PVM operated on the universe as a POVM operated on the system, with the environment in the role of ancilla. The difference between the entropy of the multinomial Bayesian marginal and the expectation over the multinomial Bayesian prior $\{P(\bar{u})\}$ of the entropy of the multivariate hypergeometric Bayesian likelihood of the system is equal to the Holevo upper bound χ above the accessible quantum information that any POVM can achieve [39]:

$$\chi = S(\hat{v}) - \sum_{\bar{u} \in \mathbb{U}} P(\bar{u})S(\hat{v}_{\bar{u}}) \geq 0.$$

Regardless of the value of U , in place of the above Bayesian information, one could consistently consider the following empirical information, which does not need the definition of a prior:

$$S(\hat{v}_{\text{emp}}) - S(\hat{v}_{\bar{u}}) \geq 0, \tag{12}$$

where $S(\hat{v}_{\text{emp}})$ is the Shannon entropy of the empirical multinomial distribution $\{P_{\text{emp}}(\bar{n})\}$ of the occupancy numbers of the system,

$$P_{\text{emp}}(\bar{n}) = W(\bar{n}) \prod_{c=1}^{|\mathbb{C}|} (U^{-1}u_c)^{n_c}.$$

In (12), the one-particle probability distribution $\{U^{-1}u_c\}$ is the same for the multivariate hypergeometric distribution and for the empirical multinomial distribution. Therefore, the inequality is guaranteed by the maximum entropy property of the multinomial distribution demonstrated in [30]. The difference (12) is the “empirical information” brought by the PVM operated on the universe about the system. As $U \rightarrow \infty$, both $S(\hat{v}_{\bar{u}})$ and $S(\hat{v}_{\text{emp}})$ tend to $S(\hat{v})$ and both the Bayesian information and the empirical information tend to zero. If, after the POVM, a PVM is operated on the system, the total information brought by the two measurements is $S(\hat{v})$ in the Bayesian approach and $S(\hat{v}_{\text{emp}})$ in the empirical approach. As $U \rightarrow \infty$, the total empirical information becomes equal to the total Bayesian information.

6. Entropy of the Ideal Gas in a Container

In the case of an ideal monoatomic gas in a cubic container of side L , one particle of the gas is modelled as a quantum “particle in a box” with three degrees of freedom, whose energy eigenvalues with aperiodic boundary conditions are

$$\epsilon_c = (c_x^2 + c_y^2 + c_z^2) \frac{h^2}{8mL^2},$$

where c consists of the three quantum numbers (c_x, c_y, c_z) , m is the mass of the particle and $h = 6.626 \cdot 10^{-34}$ J · s is the Planck constant.

When the temperature-to-density ratio is high, it becomes possible to employ two approximations. In the first one, the partition function is approximated to an integral, see Equation (19.54) of [40], leading to

$$-NE\{\log(P(c))\} \approx \frac{3N}{2} \left(1 + \log\left(\frac{2\pi mk_B TL^2}{h^2}\right) \right).$$

In the second one, with the idea that the probability that two or more particles occupy the same state is negligible, the denominator of the multinomial coefficient is ignored and, for large number of particles, the logarithm of the numerator $\log(N!)$ is approximated to $N \log(N) - N$, leading to

$$-E\{\log(W(\bar{n}))\} \approx -N \log(N) + N.$$

Plugging the two approximations in (8), one gets the textbook Sackur–Tetrode formula:

$$S(\hat{v}) \approx N \left(\log\left(\frac{L^3}{N} \left(\frac{2\pi mk_B T}{h^2}\right)^{\frac{3}{2}}\right) + \frac{5}{2} \right).$$

Note that, as already mentioned in the introduction, the exact entropy of the multinomial distribution (8) is guaranteed to be non-negative, while the Sackur–Tetrode formula becomes negative at low temperature-to-density ratios, where the two mentioned approximations do not hold; see [37,41] for details.

7. Future Work

We hereafter sketch the application of our proposed approach to the state of Szilard engine after the insertion of the piston, leaving the complete analysis of the Szilard cycle to future work. In the case of N particles, after the insertion of the piston, the number b of particles that are found in one of the two sub-containers, say the one of volume V' , is a binomial random variable $(N, V'/V)$, where V is the total volume. The probability distribution of the occupancy numbers is

$$P(\bar{n}', \bar{n}'', b) = P(\bar{n}', \bar{n}'' | b)P(b) = P(\bar{n}' | b)P(\bar{n}'' | b)P(b),$$

where $\{P(\bar{n}' | b)\}$ ($\{P(\bar{n}'' | b)\}$) is the probability distribution of the occupancy numbers of a gas with b ($N - b$) particles in the sub-container of volume V' ($V - V'$). For $N = 1$ and $V' = V/2$, the entropy after the insertion of the wall is

$$\log(2) - \sum_{c=1}^{|\mathbb{C}|} P(c) \log(P(c)),$$

where the famous $\log(2)$ of Landauer [42] comes from the binary equiprobable random variable b and $\{P(c)\}$ is the probability distribution of one particle in a box of volume V' . We have evaluated the partition function of the Boltzmann distribution with the parametrization of [43], that is the mass of the particle $m = 9.11 \cdot 10^{-31}$ kg, temperature $T = 300$ K, and one-dimensional box of size $L = 20 \cdot 10^{-9}$ m. We obtain that the entropy of the single particle with one degree of freedom before the insertion of the piston is 1.988 in k_B units. After the insertion of the piston, the size of the one-dimensional box is equal to $10 \cdot 10^{-9}$ m. The entropy of one particle in this box is, in k_B units, equal to 1.243. Considering also the $\log(2)$ that comes from the binary equiprobable random variable, the entropy fall due to the insertion of the piston results equal to $1.988 - 0.693 - 1.243 = 0.052$, in excellent agreement with the entropy fall shown in Figure 3 of [43], where the result is derived by the phase space approach.

8. Conclusions

Entropy is a macroscopic property of a physical system and, at the same time, it is a mathematical property of randomness. As such, it must be a property of the randomness of system's macrostates. However, despite the necessity of introducing the somewhat ad hoc term $-\log(N!)$ as a kind of *deus ex machina*, from Boltzmann to the present day, the "entropy" of a system at equilibrium is commonly intended as the entropy of microstates, be it the Gibbs entropy or, to some extent, the Boltzmann entropy $\log(W(E\{\bar{n}\}))$. This author has come to the conclusion that the misunderstanding arises from the following two conceptual errors.

The first error consists in regarding $\log(W(\bar{n}))$ as the "entropy of a macrostate". Even an author as exceptionally deep as Jaynes writes the following in [44]: *To emphasize this, note that a "thermodynamic state" denoted by $X = \{X_1, \dots, X_n\}$ defines a large class $C(X)$ of microstates compatible with X . Boltzmann, Planck, and Einstein showed that we may interpret the entropy of a macrostate as $S(X) = k\log(W(C))$, where $W(C)$ is the phase volume occupied by all the microstates in the chosen reference class C .* This misconception is pervasive in the entire statistical mechanics. The influential authors of [45] attribute to Einstein the idea that an individual macrostate can have entropy: *In fact, already Einstein (1914, Equation (4a)) argued that the entropy of a macro state should be proportional to the log of the "number of elementary quantum states" compatible with that macro state...* But entropy is a property of a statistical ensemble; therefore, one individual macrostate cannot have entropy. Moreover, if by entropy of the macrostate we mean the entropy of the microstates that constitute it, then we must acknowledge that this entropy has no physical significance because the *elementary quantum state compatible with that macrostate* is physically meaningless when particles are indistinguishable. This misconception is so widespread that it can be found also in standard textbooks. We hereafter quote a passage from the introduction to chapter 16 of [40]: *specification of a macrostate constitutes incomplete information.* In the absence of a formal definition of information, this statement risks becoming misleading. Actually, we have shown that the entropy of the occupancy numbers is the complete information about the system because the entropy of microstates belonging to the same macrostate, due to indistinguishability of particles, is not informative, as it is subtracted in (8) from the entropy of microstates.

The second error is the lack of consideration of the absolute randomness of macrostates. Navigating through the literature of statistical mechanics, you may come across passages like the one of [31] that is hereafter reported: *A crucial observation in statistical mechanics is that the distribution of all macrostate variables gets sharply peaked and narrow as system size N increases. ... In the limit $N \rightarrow \infty$, the probability of measuring a macrostate becomes a Dirac delta. ...* Clearly, the quoted statement is wrong because the absolute width of the probability distribution of the occupancy numbers, and, more generally, of any macroscopic observable, such as the system's energy, becomes broader and broader as the number of particles grows, while, by the LLN, its relative width, i.e., the width compared to the mean value, sharpens. Certainly, it is apparent that

$$\lim_{N \rightarrow \infty} N^{-1}S(\hat{v}) = 0,$$

which shows that the relative randomness, i.e., the randomness per particle, becomes vanishingly small as $N \rightarrow \infty$. Meanwhile, the absolute randomness, represented by $S(\hat{v})$, increases as N becomes larger and larger.

In the end, the lack of formal specification of the physical role of microstates, together with the lack of consideration of the absolute randomness of macrostates, lead to the questionable belief that all the properties of a system of a large numbers of particles, including "entropy," depend on the macrostate $E\{\bar{n}\}$, because it contains "the overwhelming majority" of microstates. From this standpoint, the subtraction of $\log(N!)$ is seen as a technical maneuver that becomes necessary to circumvent the challenge posed by the indistinguishability of particles. Note that the idea itself that one macrostate can contain the overwhelming

majority of microstates is inherently questionable. This is because the ratio $|\mathbb{C}^N(\bar{n})|/|\mathbb{C}^N|$ tends to zero as $N \rightarrow \infty$, regardless of the value of \bar{n} .

In conclusion, this paper made clear that, since the quantum state of a system of indistinguishable particles (bosons) is completely specified by the occupancy numbers of the quantum states allowed to system's particles, the entropy the physical system is the Shannon entropy of the random occupancy numbers \bar{n} , which is obtained by subtracting the expectation of $\log(W(\bar{n}))$ from the entropy of microstates. Recognizing that $E\{\log(W(\bar{n}))\}$ is the conditional Shannon entropy of microstates given the macrostate, we equivalently express the above concept by saying that the entropy of the physical system is equal to the mutual information between microstates and macrostates.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflict of interest.

References

1. The bit juggler. *Nat. Phys.* **2023**, *19*, 919. [[CrossRef](#)]
2. Sklenarova, B.; Chladek, J.; Macek, M.; Brazdil, M.; Chrastina, J.; Jurkova, T.; Dolezalova, I. Entropy in scalp EEG can be used as a preimplantation marker for VNS efficacy. *Sci. Rep.* **2023**, *13*, 18849. [[CrossRef](#)]
3. Rey-Devesa, P.; Prudencio, J.; Benítez, C.; Bretón, M.; Plasencia, I.; León, Z.; Ibáñez, J.M. Tracking volcanic explosions using Shannon entropy at Volcán de Colima. *Sci. Rep.* **2023**, *13*, 9807. [[CrossRef](#)]
4. Drzazga-Szczesniak, E.A.; Szczepanik, P.; Kaczmarek, A.Z.; Szczesniak, D. Entropy of financial time series due to the shock of war. *Entropy* **2023**, *25*, 823. [[CrossRef](#)]
5. Myron, T.; McIrvine, E.C. Energy and information. *Sci. Am.* **1971**, *225*, 179–190.
6. Dominik, S.; Deutsch, J.M.; Aguirre, A. Quantum coarse-grained entropy and thermalization in closed systems. *Phys. Rev. A* **2019**, *99*, 012103.
7. Rolf, L. Information is a physical entity. *Phys. A Statistical Mech. Appl.* **1999**, *263*, 63–67.
8. Maroney, O.J.E. Information and entropy in quantum theory. *arXiv* **2004**, arXiv:quant-ph/0411172.
9. Maroney, O.J.; Timpson, C.G. Timpson. How is there a physics of information? On characterizing physical evolution as information processing. *Phys. Perspect. Comput. Perspect. Phys.* **2018**, 103–126.
10. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *4*, 620–630. [[CrossRef](#)]
11. Swendsen, R.H. Probability, entropy, and Gibbs' paradox (es). *Entropy* **2018**, *20*, 450. [[CrossRef](#)] [[PubMed](#)]
12. Swendsen, R.H.; Wang, J.-S. Gibbs volume entropy is incorrect. *Phys. Rev. E* **2015**, *92*, 020103. [[CrossRef](#)] [[PubMed](#)]
13. Barsky, E. Gibbs Paradox as a Derivative of Composition Entropy. *Refract. Ind. Ceram.* **2022**, *63*, 365–369. [[CrossRef](#)]
14. Sasa, S.I.; Hiura, K.; Nakagawa, N.; Yoshida, A. Quasi-static Decomposition and the Gibbs Factorial in Small Thermodynamic Systems. *J. Stat. Phys.* **2022**, *189*, 31. [[CrossRef](#)]
15. Tasaki, H. The best answer to the puzzle of Gibbs about $N!!$: A note on the paper by Sasa, Hiura, Nakagawa, and Yoshida. *arXiv* **2022**, arXiv:2206.05513.
16. Yūto, M.; Ueda, M. Gibbs paradox revisited from the fluctuation theorem with absolute irreversibility. *Phys. Rev. Lett.* **2017**, *118*, 060601.
17. Sandu, P.; Short, A.J.; Winter, A. Entanglement and the foundations of statistical mechanics. *Nat. Phys.* **2006**, *2*, 754–758.
18. Goldstein, S.; Lebowitz, J.L.; Tumulka, R.; Zanghì, N. Canonical typicality. *Phys. Rev. Lett.* **2006**, *96*, 050403. [[CrossRef](#)]
19. Christian, B.; Gemmer, J. Dynamical typicality of quantum expectation values. *Phys. Rev. Lett.* **2009**, *102*, 110403.
20. Spakovsky, V.; Gemmer, M.R.J. Some trends in quantum thermodynamics. *Entropy* **2014**, *16*, 3434–3470. [[CrossRef](#)]
21. Goold, J.; Huber, M.; Riera, A.; Del Rio, L.; Skrzypczyk, P. The role of quantum information in thermodynamics—A topical review. *J. Phys. A Math. Theor.* **2016**, *49*, 143001. [[CrossRef](#)]
22. Parrondo, J.M.; Horowitz, J.M.; Sagawa, T. Thermodynamics of information. *Nat. Phys.* **2015**, *11*, 131–139. [[CrossRef](#)]
23. Sebastian, D.; Campbell, S. *Quantum Thermodynamics: An Introduction to the Thermodynamics of Quantum Information*; Morgan and Claypool Publishers: San Rafael, CA, USA, 2019.
24. Beggi, A.; Siloi, I.; Benedetti, C.; Piccinini, E.; Razzoli, L.; Bordone, P.; Paris, M.G. Back and forth from Fock space to Hilbert space: A guide for commuters. *Eur. J. Phys.* **2018**, *39*, 065401. [[CrossRef](#)]
25. Kardar, M. *Statistical Physics of Particles*; Cambridge University Press: Cambridge, UK, 2007.
26. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*, 2nd ed.; Wiley Series in Telecommunications and Signal Processing; Wiley: New York, NY, USA, 2006.
27. Deutsch, J.M. Eigenstate thermalization hypothesis. *Rep. Prog. Phys.* **2018**, *81*, 082001. [[CrossRef](#)] [[PubMed](#)]

28. Evan, G.; Wellner, J.A. Exponential bounds for the hypergeometric distribution. *Bernoulli Off. J. The Bernoulli Soc. Math. Stat. Probab.* **2017**, *23*, 1911.
29. Agrawal, R. Finite-sample concentration of the multinomial in relative entropy. *IEEE Trans. Inf. Theory* **2022**, *66*, 6297–6302. [[CrossRef](#)]
30. Hal, S.; Cover, T.M. Maximum entropy and the lottery. *J. Am. Stat. Assoc.* **1989**, *84*, 980–985.
31. Rudolf, H.; Thurner, S.; Gell-Mann, M. How multiplicity determines entropy and the derivation of the maximum entropy principle for complex systems. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 6905–6910.
32. Županović, P.; Domagoj, K. Relation between Boltzmann and Gibbs entropy and example with multinomial distribution. *J. Phys. Commun.* **2018**, *2*, 045002. [[CrossRef](#)]
33. Spalvieri, A. The Shannon-McMillan theorem proves convergence to equiprobability of Boltzmann’s microstates. *Entropy* **2021**, *23*, 899. [[CrossRef](#)]
34. Qiu, T.; Fei, Z.; Pan, R.; Quan, H.T. Quantum corrections to the entropy and its application in the study of quantum Carnot engines. *Phys. Rev. E* **2020**, *101*, 032113. [[CrossRef](#)] [[PubMed](#)]
35. Kaji, Y. Bounds on the entropy of multinomial distribution. In Proceedings of the 2015 IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015.
36. Cheraghchi, M. Expressions for the entropy of binomial-type distributions. In Proceedings of the 2018 IEEE International Symposium on Information Theory (ISIT), Vail, CO, USA, 17–22 June 2018.
37. Spalvieri, A. The Exact Entropy Formula of the Ideal Gas and its Information-Theoretic Interpretation. *arXiv* **2023**, arXiv:2305.04652.
38. Safranek, D.; Juzar, T. Quantifying information extraction using generalized quantum measurements. *Phys. Rev. A* **2023**, *108*, 032413. [[CrossRef](#)]
39. Nielsen, M.A.; Chuang, I.L. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2010.
40. Sekerka, R.F. *Thermal Physics: Thermodynamics and Statistical Mechanics for Scientists and Engineers*; Elsevier: Amsterdam, The Netherlands, 2015.
41. Spalvieri, A. Infomechanics of Independent and Identically Distributed Particles. *arXiv* **2022**, arXiv:2204.06952.
42. Landauer, R. Irreversibility and heat generation in the computing process. *IBM J. Res. Dev.* **1961**, *5*, 183–191. [[CrossRef](#)]
43. Alhun, A.; Sisman, A.; Kosloff, R. Landauer’s Principle in a quantum Szilard engine without Maxwell’s Demon. *Entropy* **2020**, *22*, 294.
44. Jaynes, E.T. The Gibbs paradox. In *Maximum Entropy and Bayesian Methods*; Springer: Dordrecht, The Netherlands, 1992; pp. 1–21.
45. Goldstein, S.; Lebowitz, J.L.; Tumulka, R.; Zanghì, N. Gibbs and Boltzmann entropy in classical and quantum mechanics. In *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*; World Scientific Publishing: Singapore, 2020 ; pp. 519–581.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.