

Copyright The Philosophy of Science Association 2020.  
Preprint (not copyedited or formatted).  
Please use DOI when citing or quoting. DOI: <https://doi.org/10.1086/710072>.

# Does von Neumann Entropy Correspond to Thermodynamic Entropy?

Eugene Chua

*University of California San Diego*

*Department of Philosophy*

*eychua@ucsd.edu*

## **Abstract**

Conventional wisdom holds that the von Neumann entropy corresponds to thermodynamic entropy, but Hemmo and Shenker (2006) have recently argued against this view by attacking von Neumann's (1955) argument. I argue that Hemmo and Shenker's arguments fail due to several misunderstandings: about statistical-mechanical and thermodynamic domains of applicability, about the nature of mixed states, and about the role of approximations in physics. As a result, their arguments fail in all cases: in the single-particle case, the finite particles case, and the infinite particles case.

## **Acknowledgements**

I would like to thank Craig Callender, Eddy Keming Chen, Erik Curiel, Tim Maudlin, John Norton, Sai Ying Ng, Adrian K. Yee, and participants at the 2019

summer school *The Nature of Entropy I* for their feedback, discussion and comments - they have contributed to this paper in one way or another. I would also like to thank two anonymous reviewers for pushing me to improve on various aspects of this paper.

# Does Von Neumann Entropy Correspond to Thermodynamic Entropy?

## **Abstract**

Conventional wisdom holds that the von Neumann entropy corresponds to thermodynamic entropy, but Hemmo and Shenker [2006] have recently argued against this view by attacking von Neumann's [1955] argument. I argue that Hemmo and Shenker's arguments fail due to several misunderstandings: about statistical mechanical and thermodynamic domains of applicability, about the nature of mixed states, and about the role of approximations in physics. As a result, their arguments fail in all cases: in the single-particle case, the finite particles case, and the infinite particles case.

# 1 Introduction

According to conventional wisdom in physics, von Neumann entropy corresponds to phenomenological thermodynamic entropy. The origin of this claim is von Neumann's (1955) argument that his proposed entropy corresponds to the thermodynamic entropy, which appears to be the only explicit argument for the equivalence of the two entropies. However, Hemmo and Shenker (**H&S**) (2006) – and earlier, Shenker (1999) – have argued that this correspondence fails, contrary to von Neumann. If so, this leaves conventional wisdom without explicit justification.

Correspondence can be understood, at the very least, as a numerical consistency check: in this context, this means that the von Neumann entropy has to be included in calculating thermodynamic entropy to ensure consistent accounting in contexts where both thermodynamic and von Neumann entropy are physically relevant. Successful correspondence provides strong evidence of equivalence. While it does not guarantee equivalence, it seems to be at least a necessary condition for equivalence. If thermodynamic entropy and von Neumann entropy correspond, then we have reason to think that von Neumann entropy is rightfully thermodynamic in nature, since proper accounting of thermodynamic entropy would demand von Neumann entropy. By contrast, a failure of correspondence seems to entail that the von Neumann entropy is *not* thermodynamical in nature, since it is irrelevant to thermodynamic calculations in contexts where both entropies are physically significant (e.g. when a system has both quantum degrees of freedom *and* is sufficiently large to warrant thermodynamical considerations).

Although Henderson (2003), in my view, has successfully criticized Shenker's earlier

argument, little has been done in the philosophical literature to evaluate **H&S**'s more recent arguments.<sup>1</sup> This lacuna is striking because, as I mentioned, von Neumann's argument appears to be the only explicit argument for correspondence for the two entropies.

My goal in this paper is to fill this lacuna by providing a novel set of criticisms to **H&S**. Here's the plan: I introduce key terms (§2) and then present von Neumann's thought-experiment which aims to establish the correspondence between thermodynamic entropy and von Neumann entropy; along the way, a novel counterpart to the usual argument for correspondence is discussed (§3). I then present and criticize **H&S**'s arguments for the single-particle case in the context of thermodynamics (§4.1) and in the context of statistical mechanics (§4.2), the N-particles case (§4.3), and the infinite-particles case (§4.4). I conclude that their argument fails in all cases – in turn, we have good reasons to reject their claim that the von Neumann entropy fails to correspond to thermodynamic entropy, and hence the claim that von Neumann entropy is not thermodynamic in nature.

---

<sup>1</sup>It is only slightly better in the physics literature: Deville and Deville (2013) appears to be the only paper to critique **H&S**. On the philosophical side, one (very recent) exception is Prunkl (ms), though she restricts discussion to the single-particle case and appears to conflate information entropy with thermodynamic entropy. See §4.1/§4.2 for why this is not obviously right.

## 2 Key Terms

Let me first define the notions of thermodynamic entropy and von Neumann entropy. Following H&S, I define the *change in thermodynamic entropy*  $S_{TD}$  between two thermodynamic states in an *isothermal quasi-static* process,<sup>2</sup> as:

$$\Delta S_{TD} = \frac{1}{T} \int P dV \quad (1)$$

We will restrict our discussion to ideal gases in equilibrium (i.e. systems where pressure  $P$ , volume  $V$ , and temperature  $T$  remain constant).

Next, the *von Neumann entropy*  $S_{VN}$ , for any pure or mixed quantum system, is defined as:

$$S_{VN} = -kTr(\rho \log \rho) \quad (2)$$

where  $k$  is the Boltzmann constant and  $Tr(\cdot)$  is the trace function. Generally, the density matrix  $\rho$  is such that:

$$\rho = \sum_{n=1}^i p_n |\psi_n\rangle \langle \psi_n| \quad (3)$$

where  $\psi_1, \psi_2, \dots, \psi_n$  correspond to the number of pure states in a statistical mixture represented by  $\rho$ , with  $p_1, p_2, \dots, p_n$  being their associated classical probabilities (which must sum to unity). In the case where there is only one pure state possible for a system (e.g. when we are absolutely certain about its quantum state), then  $n = 1$ , with

---

<sup>2</sup>There is no change in temperature in an isothermal quasi-static process, which is why  $T$  is taken to be constant. As a matter of historical note, von Neumann uses an isothermal set-up in his argument, with a box containing a quantum ideal gas coupled to a (much larger) heat sink ensuring constant temperature over time (von Neumann 1955, 361/371).

probability 1, so the appropriate density matrix is  $\rho = |\psi\rangle\langle\psi|$ . For such a system in a pure state (i.e. represented by a single state vector in Hilbert space),  $S_{\text{VN}} = 0$ . For mixed states (i.e. states which cannot be represented by a single state vector in Hilbert space, hence *mixture* of pure states or a *mixed state*),  $\text{Tr}(\rho \log \rho) < 1$  and  $S_{\text{VN}} > 0$  in general. A mixed state is often said to represent our *ignorance* about a system – this will suffice as a first approximation (more on how to interpret this ignorance in §4.2).

*Prima facie*,  $S_{\text{VN}}$  and  $S_{\text{TD}}$  appear to share nothing in common, apart from the word ‘entropy’. However, von Neumann claims that there are important correlations between the two, which suggests a correspondence between  $S_{\text{TD}}$  and  $S_{\text{VN}}$ .

### 3 Von Neumann’s Thought-Experiment

For the sake of parity, I adopt **H&S**’s presentation of von Neumann’s thought-experiment,<sup>3</sup> which aims to show that changes in thermodynamic entropy can only be made consistent with the laws of thermodynamics if we considered the von Neumann entropy as contributing to the calculation of the thermodynamic entropy. *Fig. 1.* depicts the stages of the thought-experiment.

We begin, in stage one, with a box with a partition in the middle. On one side of the partition there is a gas at volume  $V$ , constant temperature  $T$ , and constant pressure  $P$ . Each gas particle starts off having the pure state spin-up along the x-direction  $|\psi_x^\uparrow\rangle$ , which is equivalent to a superposition of spin-up and spin-down pure states along the

---

<sup>3</sup>It is not clear to me that von Neumann’s original 1932/1955 argument is exactly the same as the argument **H&S** reproduces. However, for the sake of argument, I will refer to **H&S**’s version as von Neumann’s argument in this paper.

z-direction, labelled  $|\psi_z^\uparrow\rangle$  and  $|\psi_z^\downarrow\rangle$  respectively. According to standard quantum mechanics, the state of each particle is thus  $\frac{1}{\sqrt{2}}(|\psi_z^\uparrow\rangle + |\psi_z^\downarrow\rangle)$ .

In this context, particles with quantum behavior may be taken to be *ideal gases*, i.e. sets of particles each of which do not interact with other particles and take up infinitesimal space. Following von Neumann’s assumptions (von Neumann 1955, 361),<sup>4</sup> each gas particle is understood as a quantum particle with a spin degree of freedom contained inside a large impenetrable box, and each gas particle is put inside an even larger container isolated from the environment (i.e. the box we began with). This ensures that each spin degree of freedom is incapable of interacting with other particles. These boxes’ sizes also ensure that the positions of these boxes (and hence of the particles) can be approximately classical. Since the container is much larger than each gas particle, this ensures that the gas particles take up negligible space relative to the massive container. Accepting these assumptions, we may then take these quantum particles to behave like an ideal gas.<sup>5</sup> Following **H&S**, we further assume that the position degrees of freedom of the gas particles have no interaction with the spin degrees of freedom at this point, and “due to the large mass of the boxes, the position degrees of freedom of the gas may be taken to be classical and represented by a quantum mechanical mixture”. (Hemmo and Shenker 2006, 155)

Moving on, stage two involves a spin measurement along the z-axis on all the particles in the container, with a result being an equally weighted statistical mixture of

---

<sup>4</sup>These assumptions are borrowed from Einstein (1914). For more, see Peres (Peres 2002, 271).

<sup>5</sup>I shall follow everyone in this debate in assuming that the above set-up is physically possible.



particles with either  $|\psi_z^\uparrow\rangle$  or  $|\psi_z^\downarrow\rangle$  states. As a result, the spin state of each particle is then represented instead by a density matrix  $\rho_{spin}$ , such that:

$$\rho_{spin} = \frac{1}{2}(|\psi_z^\uparrow\rangle\langle\psi_z^\uparrow| + |\psi_z^\downarrow\rangle\langle\psi_z^\downarrow|) \quad (4)$$

More precisely, there should be terms for the measurement device too, when truly considering the entire system.  $\rho_{spin}$  describes only the subsystem (i.e. the quantum ideal gas) *sans* measurement device, i.e. a state with the measurement device traced out – this is in line with von Neumann’s focus on the entropy changes due to changes in the subsystem (von Neumann 1955, 358–379). I follow Henderson (2003) and **H&S** in talking about the system’s state as though I have already traced the measurement device out whenever measurement is involved.

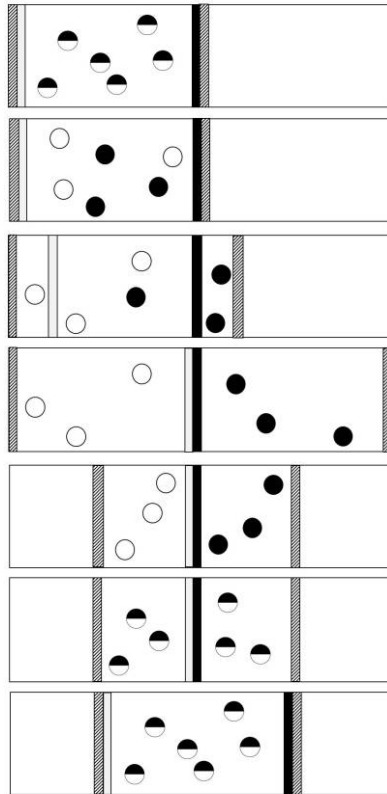
Stages three and four are where the particles are (reversibly) separated according to their spin states by a semi-permeable wall into two sides of the box, each with volume  $V$ .<sup>6</sup> As a result of this separation, we in effect double the mixture’s volume. The gas expands to fill up volume  $V$  on each side.

Stage five involves an isothermal and quasi-static compression of the mixture so that we return to a total volume  $V$  (effectively halving the volume on each side of the box), while pressure on both sides becomes equal. Importantly, due to this compression,  $S_{TD}$  decreases due to the decrease in volume.

---

<sup>6</sup>This semi-permeable wall can be assumed to be a black box which reversibly separates particles to different sides based on their different orthogonal/disjoint states; see (von Neumann 1955, 367–370) for discussion. I follow everyone in the debate in accepting this assumption.

Stage six brings all the particles into the pure spin state  $|\psi_x^\uparrow\rangle$  quasi-statically and without work done, while stage seven removes the semi-permeable wall, such that the system returns to its original state.



*Fig. 1: from top to bottom, stage one to stage seven, as described by H&S.*

Now consider how  $S_{VN}$  and  $S_{TD}$  change across the various stages. Stage seven ends with the body of gas having the same thermodynamic state (same  $V$ , same  $P$ , and constant  $T$ ) as stage one. Furthermore, all the thermodynamic transformations performed were reversible, and removing the wall alone does no additional work. Thus, the system at stage one must have the same thermodynamic entropy as stage seven, i.e.  $\Delta S_{TD} = 0$ , since  $S_{TD}$  depends only on the initial and final state of the system.

$\Delta S_{VN} = 0$  from stage one to seven too, since the system is in the same state in both the first and seventh stages.

Since stage six does not involve thermodynamic transformations, there is no change in  $S_{TD}$ . Likewise, the transformation of  $\rho_{spin}$  to  $|\psi_x^\uparrow\rangle$  here does not change  $S_{VN}$  as the transformation can be performed unitarily. This is possible as a result of our separation of the gases to different sides of the box according to their spin-eigenstates - given this, we can perform unitary operations on each side of the box (or perform the more general measurement procedure recommended by (von Neumann 1955, 365-367)), to transform them into the same state as stage one. Both unitary transformations and von Neumann's procedure do not increase  $S_{VN}$ , and so there is no change in  $S_{VN}$  at stage six as a result.

There are no changes in  $S_{TD}$  or  $S_{VN}$  in stages three and four. While there is an increase in the gas's volume, as noted above, from  $V$  to  $2V$ , and hence an accompanying increase in  $S_{TD}$  by  $n.R.\log 2$ ,<sup>7</sup> there is also a compensating change in the thermodynamic entropy of mixing<sup>8,9</sup> by  $-n.R.\log 2$  which exactly compensates this

---

<sup>7</sup>Here,  $n$  refers to the number of moles of gas in the system, and  $R$  is the gas constant.

<sup>8</sup>Henderson (2003) explains the mixing entropy, describing the mixing of different gases, crisply: "After separation, each separated gas occupies the original volume  $V$  alone. To return to the mixture, each gas is compressed to a volume  $c_i V$  (where  $c$  is the concentration of the  $i^{th}$  gas). The compression requires work  $W = -n.k.T \sum_i c_i \log c_i$  to be invested, and the entropy of the gas is reduced by  $\Delta S = -n.k. \sum_i c_i \log c_i$ . An increase in entropy of the same amount must then be associated with the mixing step of removing the partitions. This is the 'mixing entropy'." (Henderson 2003, 292) Separation simply results in a decrease in entropy of the same amount.

<sup>9</sup>Tim Maudlin raised the following objection to the applicability of the entropy of mix-

increase in  $S_{TD}$ .<sup>10</sup> Since the particles are in orthogonal spin states at this stage, there are no quantum effects (e.g. ‘collapse’ effects) from simply filtering the gases with the semi-permeable walls, and hence  $S_{VN}$  does not change either.<sup>11</sup>

However, importantly, there is a decrease in  $S_{TD}$  in stage five, of  $-n.R.\log 2$  due to the isothermal compression and decrease in volume. Yet, nowhere else is there any further change in  $S_{TD}$ . We have to account for why the overall change in  $S_{TD}$  from the first to the seventh stages is 0.

As von Neumann argues, only one possibility remains. While  $S_{TD}$  remains constant in stage two, notice that there was an increase in  $S_{VN}$ , of  $-N.k. - \log 2 = \frac{N.R}{N_A}.\log 2 = n.R.\log 2$ ,<sup>12</sup> as a result of the spin measurement. This is equivalent to the change of  $S_{TD}$  in stage five. The state of each particle changes from a ing in this context when a version of this paper was presented at a summer school. Mixing should have a thermodynamic effect only when differences between the gases are already assumed to be thermodynamically relevant: for example, mixing differently colored gases should not have a thermodynamic effect unless the difference in color is thermodynamically relevant. It is, however, not clear whether the difference in spin is a thermodynamically relevant one, and might amount to begging the question. This is a good point, but one that I am setting aside for now, since everyone in the debate accepts the assumption that separating the gases here decreases the entropy of mixing. As we shall see later, a more fundamental issue arises with using the entropy of mixing in the ‘single particle’ case.

<sup>10</sup>see (H&S 2006, 157, fn. 4).

<sup>11</sup>This is argued for in (von Neumann 1955, 370–376).

<sup>12</sup> $N$  is the total number of particles: since each particle is assumed to be non-interacting and independent from others under the ideal gas assumption, their entropies are additive.  $N_A$  is Avogadro’s number.

pure state  $\frac{1}{\sqrt{2}}(|\psi_z^\uparrow\rangle + |\psi_z^\downarrow\rangle)$  to a mixed state represented by  $\rho_{spin}$ , and hence  $S_{VN}$  for the gas increases on the whole. In order to ensure that entropic changes are consistent, von Neumann thinks that we should accept  $S_{VN}$ 's contribution to  $S_{TD}$  in this context, where both quantum effects and thermodynamical considerations are at play. Without accepting  $S_{VN}$  in our entropic accounting, we end up with a violation of thermodynamics since we have a reversible thermodynamic cycle with non-zero change in  $S_{TD}$ , *contra* the Second Law. In other words, we should accept that  $S_{VN}$  corresponds to  $S_{TD}$ .

Furthermore, the correspondence of  $S_{VN}$  and  $S_{TD}$  in this context can be defended from another perspective, apart from considerations about consistency from the thermodynamic perspective: consistent accounting from the perspective of *quantum mechanics* also demands correspondence. This is simply a change in perspective with regards to the thought-experiment, but, to my knowledge, this argument has not been explicitly made in the literature, thus underselling the case for correspondence in von Neumann's thought experiment.

Instead of arguing for correspondence by considering thermodynamic consistency, i.e. ensuring that  $\Delta S_{TD} = 0$  throughout the cycle, we can also consider consistency from the quantum mechanical perspective. We started and ended with the same spin state, and so it should be the case that  $\Delta S_{VN} = 0$  throughout the cycle. Yet, there is an inconsistency: if we only consider the increase of  $S_{VN}$  in stage two as a result of measurement, we should end in stage seven with an *increase* in  $S_{VN}$ , *not*  $\Delta S_{VN} = 0$ . As described, there is nowhere else in the thought-experiment where  $S_{VN}$  changes. However, there *is* a decrease in  $S_{TD}$  in stage five due to the *thermodynamic process* of isothermal compression, exactly balancing out the increase in  $S_{VN}$ . Hence, we can ensure consistency, i.e. that  $\Delta S_{VN} = 0$ , only by taking  $S_{VN}$  to correspond to  $S_{TD}$ . In other words, just as the

thermodynamic accounting of  $S_{\text{TD}}$  is consistent only if we consider  $S_{\text{VN}}$ , the quantum entropic accounting of  $S_{\text{VN}}$  is *also* consistent only if we consider  $S_{\text{TD}}$ . Consistency from a quantum mechanical perspective also demands correspondence between  $S_{\text{VN}}$  and  $S_{\text{TD}}$ .

Though the debate has largely focused only on how the thought-experiment demonstrates one direction of correspondence, of  $S_{\text{VN}}$  to  $S_{\text{TD}}$  as a result of thermodynamical considerations, the correspondence demonstrated by this thought-experiment in fact goes *both ways*. Of course, since von Neumann was focused on demonstrating the *thermodynamic* nature of  $S_{\text{VN}}$  (specifically the irreversibility of measurement), rather than the *quantum* nature of  $S_{\text{TD}}$ , it was natural that he chose to approach it the way he did.

## 4 Hemmo and Shenker's Arguments

H&S disagree with von Neumann's argument, and criticize it by considering three cases: the single-particle case, the finite but large  $N$  particles case, and the infinite particles case.

### 4.1 Single Particle Case - Thermodynamics

H&S first consider von Neumann's argument in the single particle case (see Fig. 2). They claim that the argument does not go through here, since  $S_{\text{TD}}$  actually remains constant, contrary to our thought-experiment's description. In other words, using thermodynamical considerations, they find that  $S_{\text{VN}}$  should not be included in our accounting for  $S_{\text{TD}}$ .

Here's their argument. Consider the stages where there are entropic changes. In

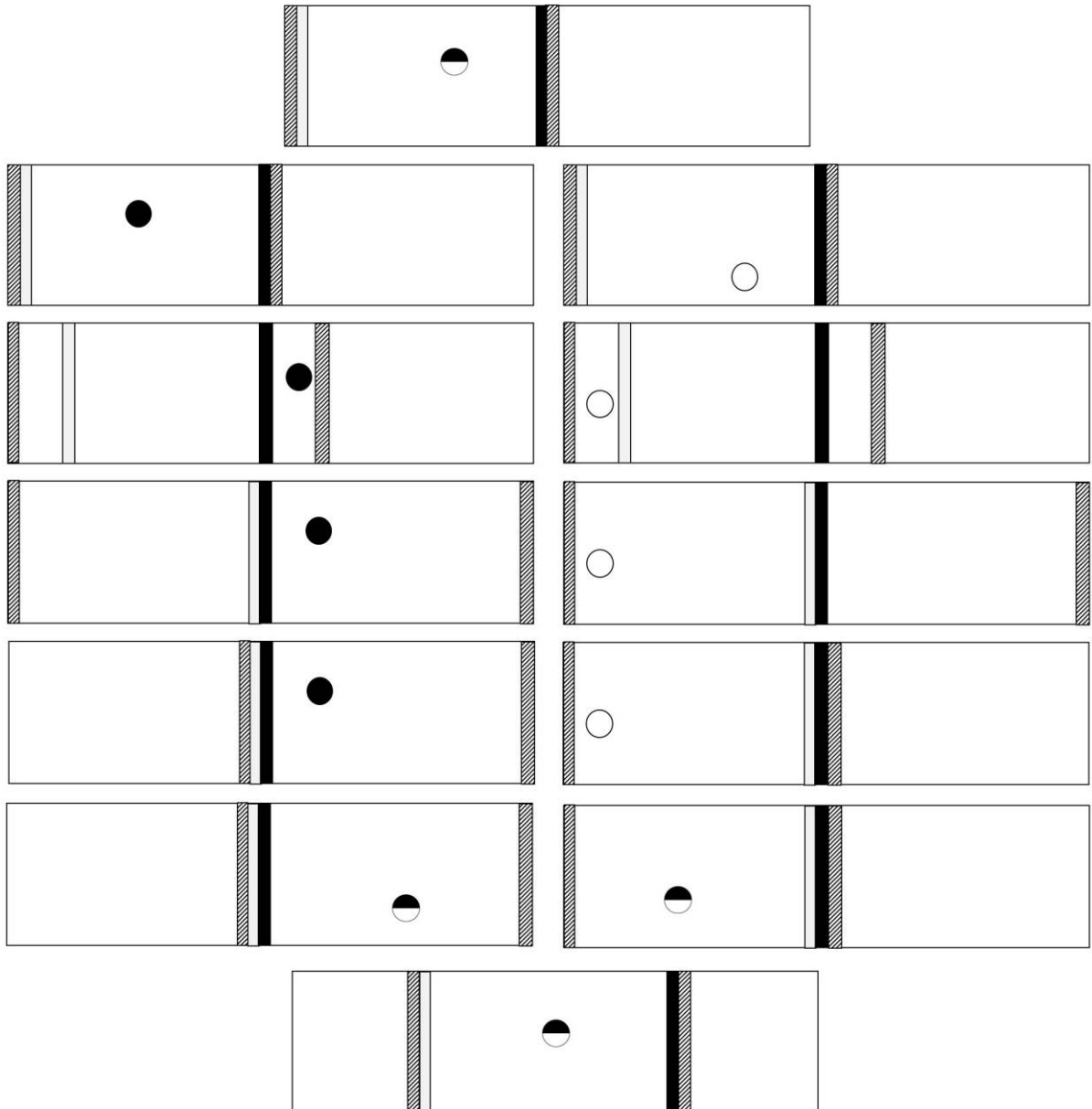
stage two when the spin measurement was performed,  $S_{VN}$  increases as before, since it tracks the change of the particle's spin state from pure to mixed.

Contrariwise,  $S_{TD}$  does not change in stage five (isothermal quasi-static compression) *nor anywhere else* (this will be important later). After stage two, the single particle is in either the  $|\psi_z^\uparrow\rangle$  state or the  $|\psi_z^\downarrow\rangle$  state. After stages three and four, with the expansion and separation via semi-permeable wall, there is a particle only in *one* side of the box, and not the other. We make an  $S_{TD}$ -conserving location measurement<sup>13</sup> to figure out which side of the box is empty and which side the particle is at, so as to compress the box against the empty side. The compression is then performed as per before. However, this compression does *not* decrease  $S_{TD}$ :<sup>14</sup> to restore the volume of the 'gas' to  $V$  no work needs to be done, since we are compressing against vacuum. Since there is a change in  $S_{VN}$  in this cycle, but no change in  $S_{TD}$ , the apparent answer, in order to do our entropic accounting, is to *ignore*, not incorporate,  $S_{VN}$  into  $S_{TD}$ . Hence  $S_{VN}$  does not correspond to  $S_{TD}$ .

---

<sup>13</sup>Prunkl (ms) claims that the location measurement leads to a violation of the Second Law. If true, this makes **H&S's** argument even more problematic. Here, for the sake of argument, I assume that the location measurement is unproblematic.

<sup>14</sup>As an anonymous reviewer rightfully notes, the location measurement is important for ensuring  $\Delta S_{TD} = 0$  here. Without the location measurement, we might end up compressing in the wrong direction against the side with the gas, rather than the empty vacuum - this will have thermodynamic effects since we are doing work on the gas. However, the **H&S** set-up emphasizes the location measurement, and I will play along for the sake of argument.



*Fig. 2: from top to bottom, stage one to stage seven for the single particle case as described by **H&S**.*

Their analysis is problematic. Though their ultimate point in this analysis – that  $S_{TD}$



fails to correspond to  $S_{VN}$  – still holds, it does not hold in the way they claim. In fact, the way it fails suggests to us that we should *disregard* the single particle case.

For the single particle case, they claim that “... [ $S_{TD}$ ] is null throughout the experiment.” (H&S 2006, 162) This then allows them to claim that thermodynamic accounting for  $S_{TD}$  is consistent *only if* we did not consider  $S_{VN}$ . This then supports their claim that  $S_{VN}$  does not correspond to  $S_{TD}$  since adding  $S_{VN}$  into the thermodynamic accounting actually renders the otherwise consistent calculations inconsistent.

They are right to say that the stage five compression (after location measurement) has no thermodynamic effect because we are compressing against vacuum: no work needs to be done, and so  $\Delta S_{TD} = 0$  for stage five. *However*, I claim that  $\Delta S_{TD} \neq 0$  for the single particle case *overall*, because  $\Delta S_{TD} \neq 0$  in stages three and four in this context.

As far as I can tell, H&S did not analyze stage three and four, i.e. the isothermal expansion and separation, in terms of the single particle case at all. Rather, they seem to have assumed that  $\Delta S_{TD} = 0$  in these stages as with the original case of the macroscopic gas.<sup>15</sup> However, this assumes that there is both a change in entropy of  $n.R.\log 2$  due to isothermal expansion *and* a change in the entropy of mixing of  $-n.R.\log 2$  due to separation, as they say so themselves for the original case: “The increase of thermodynamic entropy due to the volume increase  $\Delta S = \frac{1}{T} \int P dV$  is exactly compensated by the decrease of thermodynamic mixing entropy  $\Delta S = \sum w_k \ln w_k$  (where  $w_k$  is the relative frequency of molecules of type  $k$ ) due to the separation.” (H&S 2006, 157, fn. 4, emphasis mine)

In the single particle case, it makes sense that isothermal expansion should still

---

<sup>15</sup>Prunkl (ms) appears to do the same.

increase  $S_{TD}$ , since the single particle ‘gas’ is expanding against a piston and doing work. However, it *does not make physical sense* to speak of the entropy of mixing here at all, since *there is no separation of gases* in the single particle case. The entropy of mixing is explicitly defined for systems where *different* gases are separated from/mixed with one another via semi-permeable walls, but a single particle cannot be separated from/mixed with itself. The quote above makes this conceptual point explicit: by **H&S**’s own lights, the relative frequency of a single particle is simply unity (and null for particles of other types), so the entropy of mixing is  $1 \ln 1 = 0$ . *There is no thermodynamic entropy of mixing in the single particle case.*

Discounting the entropy of mixing, however, we find that  $\Delta S_{TD} = n.R.\log 2 \neq 0$  for stages three and four, and hence for the entire process, contrary to **H&S**’s claim. Interestingly, correspondence *does* fail to obtain between  $S_{TD}$  and  $S_{VN}$ , since  $\Delta S_{TD} + \Delta S_{VN} = 2n.R.\log 2 \neq 0$ , despite the process being reversible *ex hypothesi*: incorporating  $S_{VN}$  into thermodynamic accounting violates the Second Law.

However, on this new analysis, we gain some clarity as to why the single particle case is problematic. While it is true that incorporating  $S_{VN}$  into the thermodynamic accounting violates the Second Law,  $S_{TD}$  accounting *by itself* also violates the Second Law (contrary to **H&S**). Even without considering  $S_{VN}$ ,  $\Delta S_{TD} \neq 0$  despite the process being reversible. Thermodynamic accounting is inconsistent here *no matter what we do*, which suggests that the reversible process they described for the single particle case is thermodynamically unsound: if so, any argument **H&S** make in this context may be disregarded.

The upshot: I agree with **H&S** that correspondence fails for the single particle case, but not *why* it fails. It is *not* because the process they described is already

thermodynamically consistent without taking  $S_{VN}$  into account. Rather, it is because the process is already thermodynamically *inconsistent* anyway.

In recent work, John Norton argued that thermodynamically reversible processes for single-particle systems are impossible in principle, which might explain *why* the process described by **H&S** is thermodynamically unsound: it was not justified to assume the process was reversible for a single particle system. For Norton, a reversible process is “loosely speaking, one whose driving forces are so delicately balanced around equilibrium that only a very slight disturbance to them can lead the process to reverse direction. Because such processes are arbitrarily close to a perfect balance of driving forces, they proceed arbitrarily slowly while their states remain arbitrarily close to equilibrium states.” (Norton 2017, 135) Norton notes that these thermodynamic equilibrium states are balanced not because there are no fluctuations, but because these fluctuations are negligible for macroscopic systems. However, fluctuations relative to single-particle systems are large, and generally prevent these systems from being in equilibrium states at any point of the process, rendering reversible processes impossible in the single particle case. (Norton 2017, 135) If reversible processes are impossible for single particle systems in general, then it should come as no surprise that the particular single particle reversible process used by **H&S** is likewise thermodynamically unsound, as my analysis above suggests. If so, **H&S**'s claim that correspondence fails in this process is simply besides the point, since this process is not thermodynamic at all.

Since any reversible process cannot be realized for single particle systems in general, the issue seems not to be with any particular process *per se*, but with the single particle case *simpliciter*. To my knowledge, no one prior to **H&S** discussed von Neumann's experiment in terms of a single particle; von Neumann (1955), Peres (1990, 2002),

Shenker (1999) and Henderson (2003) all explicitly or implicitly assume a large (or infinite) number of particles. This is for good reason. As **H&S** acknowledge, and as we have seen: “The case of a single particle is known to be problematic as far as arguments in thermodynamics are concerned”. (**H&S** 2006, 158) Matter in phenomenological thermodynamics is assumed to be continuous.<sup>16</sup> A ‘gas’ composed of one particle can be many things, but it is surely not continuous in any commonly accepted sense. In other words, it is just not clear whether the domain of thermodynamics should apply to the single-particle case at all.

As Myrvold (2011) notes, Maxwell also made a similar claim with regards to phenomenological thermodynamics *in general*; it does not and should not hold in the single particle case. On his view, the laws of phenomenological thermodynamics, notably the Second Law, must be continually violated on small scales:

If we restrict our attention to any one molecule of the system, we shall find its motion changing at every encounter in a most irregular manner.

If we go on to consider a finite number of molecules, even if the system to which they belong contains an infinite number, the average properties of this group, though subject to smaller variations than those of a single molecule, are still every now and then deviating very considerably from the theoretical mean of the whole system, because the molecules which form the group do not submit their procedure as individuals to the laws which prescribe the behaviour of the average or mean molecule.

---

<sup>16</sup>See Compagner (1989) for a discussion of the so-called ‘continuum limit’ as a counterpart to the thermodynamic limit in phenomenological thermodynamics.

Hence the second law of thermodynamics is continually being violated, and that to a considerable extent, in any sufficiently small group of molecules belonging to a real body. As the number of molecules in the group is increased, the deviations from the mean of the whole become smaller and less frequent [...] (Maxwell 1878, 280)

The Second Law, and hence phenomenological thermodynamics, should not be expected to hold true universally in small scale cases, and especially *not* in the single-particle case. Von Neumann and everyone else in the debate should have recognized this point. Why, then, should it matter that the thought-experiment succeeds or fails in this case? Phenomenological thermodynamics does not apply to single-particle cases. There is thus no profit in trying to establish correspondence between  $S_{VN}$  and  $S_{TD}$  in this case. Indeed, if we took seriously Maxwell's claim that the Second Law fails at small scales, a failure of thermodynamic entropic accounting might even be *expected*; it does not rule out the possible thermodynamic nature of  $S_{VN}$  even though the sum of  $S_{VN}$  and  $S_{TD}$  might be inconsistent with the Second Law. In short, it is not clear *why* the single-particle case is relevant to the discussion at hand.

**H&S's** reasoning is untenable, because they fail to respect the context of phenomenological thermodynamics by bringing it into a context where it is not expected to hold. Instead, it seems more appropriate that the single-particle case is precisely beyond the purview of classical thermodynamics, requiring an analogue that only corresponds to classical thermodynamics at the appropriate scales and limits. We may then take  $S_{VN}$  to be the analogue of  $S_{TD}$  in this case, only approximating  $S_{TD}$  as the system in question approaches the context suitable for traditional thermodynamic

analysis. If so, we may see von Neumann as merely demonstrating that  $S_{VN}$  corresponds, not at *all* domains but *in the domain where thermodynamics is taken to hold*, to  $S_{TD}$ .

## 4.2 Single Particle Case Redux - Statistical Mechanics and Information

Given the foregoing discussion, **H&S** might insist that  $S_{VN}$  fails to correspond to  $S_{TD}$  *even when* take into account a more relevant domain for single particles – statistical mechanics.

After directly arguing that  $S_{VN}$  does not correspond to  $S_{TD}$  (**H&S** 2006, 162–165), they further argue that  $S_{VN}$  does not correspond to information entropy (more on this below) in the single-particle case. *Prima facie*, this should seem *irrelevant* to von Neumann’s argument, which was to establish the correspondence of the *thermodynamic*  $S_{TD}$  and *quantum*  $S_{VN}$ : why should *information* entropy’s failure to correspond with  $S_{VN}$  be a worry at all?

Here’s one plausible worry, on a charitable reading. *If* information entropy corresponds to  $S_{TD}$ , *and* **H&S** shows that  $S_{VN}$  fails to correspond to information entropy, then we might conclude, indirectly, that  $S_{VN}$  does not correspond to  $S_{TD}$  after all.<sup>17</sup> This argument assumes that information entropy *does* correspond to  $S_{TD}$ , an assumption **H&S** seem to hold as well: this is in line with the so-called ‘subjectivist’

---

<sup>17</sup>Caveat: I am not committed to the information entropy’s relationship to thermodynamics. One may, like Earman and Norton (1998, 1999), be skeptical that information entropy is related to  $S_{TD}$  at all, in which case **H&S**’s argument here is simply irrelevant.

view of statistical mechanics.<sup>18</sup> Furthermore, my above argument against the misapplication of phenomenological thermodynamics does not seem to apply here, since this argument is being made in the context of statistical mechanics and its particle picture, with no commitment to phenomenological thermodynamics.

However, **H&S** do not do much to motivate the linkage between information entropy and  $S_{TD}$ ; indeed, in their words, “a linkage between the Shannon information and thermodynamic entropy has not been established” (**H&S** 2006, 164). Without this link, the failure of correspondence between the information entropy and  $S_{VN}$  appears, at best, irrelevant to the correspondence between  $S_{TD}$  and  $S_{VN}$ . Nevertheless, I will take a charitable view here and assume that there *is* a correspondence between information entropy and  $S_{TD}$ , for the sake of assessing their argument. Here’s a plausible (if arguable) sketch: if one were a subjectivist like Jaynes (1957), one might take the Gibbs entropy in statistical mechanics to be a special case of the information entropy. After all, both have the form:

$$- \sum_i p_i \ln p_i \quad (5)$$

with  $i$  being the number of possible states with associated probabilities of occurring  $p_i$ , with the Gibbs entropy being multiplied by an additional Boltzmann’s constant  $k$ .<sup>19</sup> We know that statistical mechanics corresponds to phenomenological thermodynamics at the thermodynamic limit so we can think of the Gibbs entropy, and hence information entropy, as corresponding to  $S_{TD}$ . I take this to be in line with what **H&S** have in mind: “to the extent that the Shannon information underwrites the thermodynamic entropy, it

---

<sup>18</sup>Notably, see Jaynes (1957).

<sup>19</sup>Using the so-called Planck units, where  $k = 1$ , Gibbs entropy and information entropy are then formally equivalent.

does so via statistical mechanics” (2006, 165). *Assuming* that the above picture is plausible, a failure of correspondence between  $S_{VN}$  and the information entropy provides evidence against the correspondence between  $S_{VN}$  and  $S_{TD}$ .

Their argument comes into two parts. Ignoring  $S_{TD}$  for the time being (which does not change throughout the cycle for the single-particle case – see §4.1), they claim that we can consider the stage five location measurement to be a decrease in *information entropy* of  $\ln 2$ , as a result of learning information about which one of two parts of the box the particle is in. On first glance, this seems to resolve the arithmetic inconsistency in entropic accounting:  $\ln 2$  is exactly the increase in  $S_{VN}$  as a result of the spin state changing from a pure state  $|\psi_x^\uparrow\rangle$  to the mixed state  $\rho_{spin}$  in stage two. In other words, for both the information and von Neumann entropy’s accounting to be correct (i.e. net change of zero across the cycle), we must consider  $S_{VN}$  as corresponding to information entropy. Now, since information entropy also corresponds, *ex hypothesi*, to  $S_{TD}$ , we have an indirect argument for the correspondence of  $S_{VN}$  to  $S_{TD}$ .

However, **H&S** claim that this argument fails for *collapse* interpretations, i.e. interpretations of quantum mechanics on which a superposed quantum state ontologically collapses into a pure state upon measurement (either precisely or approximately).<sup>20</sup> They allow that, on *no-collapse* interpretations, e.g. Bohmian or many-worlds interpretations, the location measurement in stage five does not decrease  $S_{VN}$ , since the state of the system never changes in light of measurements, and so the above argument goes through.

Let us see what they could mean by this claim by following the state of the particle through the cycle. At stage two, everyone agrees that the state of the system is  $\rho_{spin}$

---

<sup>20</sup>On GRW-type approaches, though, collapse occurs with or without measurement, but measurement increases the likelihood of collapse, roughly speaking.



following the z-spin measurement;  $S_{VN}$  increases by  $\ln 2$ . At this point, the particle's position degrees of freedom remain independent from its spin degrees of freedom, as per our ideal gas assumption, though we might assume the particle starts out on the left half of the box, with the mixture of position states  $\rho_{pos}(L)$  with 'L' representing the left side. (Consider *Fig. 1* but with only one particle). Following the semipermeable wall's filtering at stages three and four, the location of the particle becomes classically correlated with the spin. Let's say that the semipermeable wall sends  $|\psi_z^\uparrow\rangle$  particles to the left, represented by  $\rho_{pos}(L)$ , and  $|\psi_z^\downarrow\rangle$  particles to the right, represented by  $\rho_{pos}(R)$ . As such, the (mixed) state of the particle is now:

$$\rho_{particle} = \frac{1}{2} \left( |\psi_z^\uparrow\rangle \langle \psi_z^\uparrow| \otimes \rho_{pos}(L) + |\psi_z^\downarrow\rangle \langle \psi_z^\downarrow| \otimes \rho_{pos}(R) \right) \quad (6)$$

For no-collapse interpretations, **H&S** agree that the state of the particle stays the same as above after the location measurement in stage five. We perform the compression in stage five and remove the partition at the end of stage six, thereby removing the classical correlations between position and spin. No further change in either information entropy or  $S_{VN}$  occurs, and hence the correspondence goes through (**H&S** 2006, 164) - the spin state remains mixed until unitarily transformed into a pure state and completing the cycle.

For collapse interpretations, they claim that the location measurement *decreases*  $S_{VN}$  by  $\ln 2$ , because, on collapse interpretations, the state of the particle upon the measurement, depending on which side the particle is found, becomes:

$$\rho_{particle} = \begin{cases} |\psi_z^\uparrow\rangle \langle \psi_z^\uparrow| \otimes \rho_{pos}(L) \\ |\psi_z^\downarrow\rangle \langle \psi_z^\downarrow| \otimes \rho_{pos}(R) \end{cases} \quad (7)$$

The spin state of the system here effectively goes from being a mixed state to a pure state as a result of this measurement:  $S_{VN}$  decreases by  $\ln 2$ . Summing up the entropy changes, there was a decrease of  $\ln 2$  in information entropy, and a net change of zero for  $S_{VN}$  as a result of the increase in stage two and the decrease in stage five. Overall, then, the change is *not* zero but  $-\ln 2$ ; our accounting has gone awry, and there is a failure of correspondence between  $S_{VN}$  and information entropy. If this is right,  $S_{VN}$  does not correspond to  $S_{TD}$ .

However, I think that **H&S** are wrong to claim that  $S_{VN}$  decreases following the location measurement for collapse interpretations. As Prunkl (Prunkl ms, 11–12) notes, there is an inconsistency here. Everyone, *including H&S*, agrees that the spin state of the particle is mixed – *not* pure – after stage two’s spin measurement, *even on* collapse interpretations (**H&S** 2006, 160). In that case, why does the particle’s spin become *pure* after the location measurement?

I think this results from a confusion over the nature of mixed states. In particular, they seem to have adopted what Hughes (Hughes 1992, §5.4, §5.8) call the “ignorance interpretation” of mixed states, confusing what I call *classical* and *quantum* ignorance. They seem to be assuming that mixed states simply represents *classical* ignorance, i.e. the lack of knowledge about a *particular system*: a system represented by a mixed state *really is* in a pure state, but we know not which. This is why the location measurement is supposed to reveal to us the pure state of this system (by revealing which side it is on

and hence the correlated spin state) and hence ‘wash away’ our classical ignorance of the real state of the system - post-measurement, we know exactly which pure state *this* system is in, unlike pre-measurement; hence  $S_{VN}$  decreases.

However, as Hughes (Hughes 1992, 144/150) argues, this interpretation of mixed states – as representing classical ignorance about which pure state a *particular* system is in – cannot be the right interpretation of all mixed states. To begin, a mixed state can be decomposed in non-unique ways in general. Here’s a simple example: a mixed state representing a mixture of  $|\psi_z^\uparrow\rangle$  and  $|\psi_z^\downarrow\rangle$  can *also* represent a mixture of  $|\psi_x^\uparrow\rangle$  and  $|\psi_x^\downarrow\rangle$  and so on. If we insist that a mixed state represent our classical ignorance about the real state of a particular system, then we end up having to say that a system’s state is really *both* either  $|\psi_z^\uparrow\rangle$  or  $|\psi_z^\downarrow\rangle$ , *and* either  $|\psi_x^\uparrow\rangle$  or  $|\psi_x^\downarrow\rangle$ . Of course, this is impossible given quantum mechanics. The defender of the classical ignorance interpretation might insist that we simply pick one pair of possible pure states but not both at once. In general, however, there’s no way to do that non-arbitrarily given some density matrix. Furthermore, this problem only worsens when we consider that there are usually more than just two ways to decompose a density matrix - a principled choice based on the mixed state alone is not feasible. The mixed state cannot be a representation of classical ignorance.

Instead, to paraphrase Hughes (Hughes 1992, 144–145), mixed states should be (minimally) interpreted as such: if we prepared in *the same way* an ensemble of systems, each described with the same mixed state, i.e. a mixture of pure states with certain weights, then the relative frequency of any given measurement outcome from the ensemble is exactly what we would get if the ensemble comprised of various ‘sub-ensembles’ each in one of the pure states in the mixture, with the relative frequency

of each sub-ensemble in the ensemble given by the corresponding weights.

In other words, the sort of *quantum* ignorance relevant in the right interpretation of mixed states is not whether we are ignorant about the *real* state of *this* particular system, but whether we are ignorant about the *measured* states of an *ensemble* of *identically prepared* systems like this one. If this is right, quantum ignorance cannot be ‘washed away’ upon measurement of a single system unlike the sort of ignorance **H&S** were implicitly assuming, and it seems like this quantum ignorance is precisely what remains after the location measurement.

This was roughly Henderson’s (2003) criticism against Shenker (1999), which is why it is puzzling that **H&S** (2006) commit the same mistake:

This preparation produces the pure states  $[|\psi_z^\downarrow\rangle]$  and  $[|\psi_z^\uparrow\rangle]$  with equal probabilities. In a particular trial, the observer may take note of the measurement result, and he therefore discovers that he has say a  $[|\psi_z^\uparrow\rangle]$ . If he applies a projective measurement in the  $\{[|\psi_z^\uparrow\rangle], [|\psi_z^\downarrow\rangle]\}$  basis, he could predict that he will measure  $[|\psi_z^\uparrow\rangle]$ . However, this does not mean that, if someone handed him another state prepared in the same way, he could again predict that the outcome of his measurement would be  $[|\psi_z^\uparrow\rangle]$ . In this sense the observer does not know the state of the system which is being prepared, and it is because of this ignorance that the state is mixed. Looking at the measurement result does not remove the fact that there is a probability distribution over the possible outcomes. (Henderson 2003, 294)

This applies to the location measurement in stage five too: measuring the location of the particle in *this* case does not change the state of the particle from a mixed one to a pure

one even on collapse interpretations. Firstly, it seems quite irrelevant whether we adopt a collapse or no-collapse interpretation, because the collapse mechanism applies to superposed pure states, not statistical mixtures. If anything, collapse had already happened in the stage two measurement procedure, yet everyone including **H&S** (**H&S** 2006, 160) accepts that the system is in a *mixed* state after stage two even for collapse interpretations. More importantly, there remains a probability distribution over the states of the particle as a result of stage two's spin measurement, even *after* the location measurement. Given an ensemble of particles prepared from stages one to five in the same way, we are *still* not be able to predict with certainty whether an ensemble of particles would all be measured on the left or right sides of the box (and hence all spin-up or spin-down) as a result of the mixed state resulting from stage two, only that half of the ensembles will be on the left and the other half will be on the right. Quantum ignorance remains – the system remains in a mixed state even after the location measurement, as:

$$\rho_{particle} = \frac{1}{2} \left( |\psi_z^\uparrow\rangle \langle \psi_z^\uparrow| \otimes \rho_{pos}(L) + |\psi_z^\downarrow\rangle \langle \psi_z^\downarrow| \otimes \rho_{pos}(R) \right) \quad (8)$$

This is exactly the state of the system in no-collapse interpretations, i.e. quantum ignorance does *not* discern between collapse and no-collapse interpretations. What *has* gone away is the classical ignorance that **H&S** (mistakenly) assumed was relevant for mixed states, ignorance about *this particular system's* state. By measuring the system's location, we come to learn of the correlations between location measurement and the particle's spin. This ignorance does not change the mixed state to a pure state: instead, this loss of classical ignorance – *gain in information* – is represented as a *decrease* in

information entropy just as before, and this information is what we *use* to perform the compression in stage five.

As a result, there is no additional decrease in  $S_{VN}$  in stage five for collapse interpretations; the entropy accounting lines up after all, as with no-collapse interpretations: the decrease in information entropy *does* correspond to the increase in  $S_{VN}$ , and so information entropy does correspond to  $S_{VN}$  after all. **H&S**'s argument does not establish the failure of correspondence between  $S_{VN}$  and  $S_{TD}$  via the failure of  $S_{VN}$  and information entropy to correspond.

To sum up, their arguments in the single-particle case are either ill-motivated and irrelevant to von Neumann and our discussion of correspondence when considered in terms of phenomenological thermodynamics, or outright fails when considered in the more relevant domain of (informational approaches to) statistical mechanics. Either way, their argument does not support the failure of correspondence between  $S_{VN}$  and  $S_{TD}$ .<sup>21</sup>

---

<sup>21</sup>Let me briefly note that their argument in the two particles case fails for similar reasons. On the one hand, from the perspective of phenomenological thermodynamics, their argument is irrelevant: following Maxwell and others, two particles do not a thermodynamic system make. On the other hand, in the domain of statistical mechanics, the analysis in terms of information entropy is irrelevant from non-informational views of statistical mechanics. From an informational perspective, however, their argument rests again on the supposed difference between collapse and no-collapse interpretations of mixed states. Since this difference is non-existent, their argument likewise fails apart in that case.

### 4.3 Finitely Many Particles

H&S's argument in the case of finitely many particles rests on the assumption of *equidistribution*, i.e. that the particles will be equally distributed across the left and right sides of the box after separation by the semi-permeable wall.

Assuming equidistribution, the increase in  $S_{VN}$  given the spin measurement in stage two is  $N \ln 2$  (H&S 2006, 169). Furthermore, the decrease in thermodynamic entropy in the fourth stage is  $N \ln 2$  as well. The entropic accounting therefore seems to work out.

However, H&S press further on the 'rough' nature of equidistribution when  $N$  is large but finite: they claim that the change in  $S_{VN}$  will only be  $N \ln 2$  when  $N$  is infinite, since equidistribution only truly holds when  $N \rightarrow \infty$ . In other cases,  $S_{VN}$  will strictly only *approximate*  $S_{TD}$ , and hence  $S_{VN}$  and  $S_{TD}$  combined will never be *exactly* zero; hence, "Von Neumann's argument goes through as an approximation" (H&S 2006, 169). However, they claim that this state of affairs suggest, instead, that von Neumann's argument strictly *fails*: "[...] since Von Neumann's argument is meant to establish a conceptual identity between the quantum mechanical entropy and thermodynamic entropy, we think that such an implication is mistaken [...] no matter how large  $N$  may be, as long as it is finite, the net change of entropy throughout the experiment will not be exactly zero." (H&S 2006, 169)

As I have already discussed in §4.1, it is not clear to me that von Neumann's goal really was to establish *strict identity* (what they call "conceptual identity"), i.e. correspondence between  $S_{VN}$  and  $S_{TD}$  in *all* domains. Rather, it seems to be the establishing of correspondence *only* in domains where  $S_{TD}$  is taken to hold. If so, their argument here simply misses the point.

Furthermore, as is well-known, the particle analogue of thermodynamics, statistical mechanics, become equivalent to phenomenological thermodynamics only when  $N = \infty$ , viz. when  $N$  arrives at the thermodynamic limit. As such, to complain that  $S_{VN}$  does not match up to  $S_{TD}$  outside of this domain is to demand the unreasonable, since it is not clear that even statistical mechanics, the bona fide particle analogue of thermodynamics, can satisfy this demand. Since  $S_{VN}$  approximates  $S_{TD}$  the same way statistical mechanical entropies approximate  $S_{TD}$  (and becomes equivalent at  $N = \infty$ ), and physicists generally accept that statistical mechanics corresponds to thermodynamics nevertheless, why should this problem of approximation be particularly problematic for  $S_{VN}$ ? I think **H&S** take too seriously the notion of conceptual identity involved in von Neumann's thought-experiment to be *strict* equality, though I suspect a better way to understand von Neumann's strategy is to understand  $S_{VN}$  as an approximation to  $S_{TD}$  that is more fundamental than  $S_{TD}$  in small  $N$  cases, but becomes part of the  $S_{TD}$  calculus in domains where  $S_{TD}$  applies.

To have a case against  $S_{VN}$  as a quantum analogue of  $S_{TD}$  in the case of finitely many particles, **H&S** must explain what exactly the problem is with approximations in *this* case, if it has worked out so well for the case of statistical mechanics and thermodynamics. If not, they might just be "taking thermodynamics too seriously".<sup>22</sup>

One might say something stronger: unless they can justify why we cannot use approximations at all in science, they do not have a case at all. As they note themselves,  $S_{TD}$  is itself only *on average approximately*  $-N \ln 2$  (**H&S** 2006, 169), only being equal to  $-N \ln 2$  when  $N = \infty$ . So, in fact, the approximate quantity of  $S_{VN}$ ,  $\sim N \ln 2$ , exactly matches the approximate quantity of  $S_{TD}$ ,  $\sim -N \ln 2$ , in the case of finitely many

---

<sup>22</sup>See Callender (2001).



particles. Unless there is something wrong with approximations in physics *in general*, this, then, is in fact a case of  $S_{VN}$  corresponding to  $S_{TD}$ , contrary to their argument.

#### 4.4 Infinitely Many Particles

**H&S** consider von Neumann's argument in the infinite particles case in two different ways: one as  $N \rightarrow \infty$  and one as  $N = \infty$ . As they rightly point out, the two cases are very different for calculations of physical quantities.

Consider stage two and stage five in this context. **H&S** emphasize that a spin measurement is "a physical operation which takes place in time" (**H&S** 2006, 170), which constrains what is physically possible.

For the case where  $N \rightarrow \infty$ , stage two is to be understood as a succession of physical measurements where "we measure individual quantities of each of the particles separately and only then count the relative frequencies" (**H&S** 2006, 170), before coming up with a density matrix describing this state. In this case, as with the case described in §4.3,  $S_{VN}$  approaches  $N \ln 2$  as  $N \rightarrow \infty$ . Their complaint here consist of two premises: one, that, as with §4.3,  $S_{VN}$  never reaches  $N \ln 2$  unless  $N = \infty$ . Two, that since measurements are physical measurements, we can never perform an infinite series of these measurements, and so we can never measure infinite particles. *A fortiori* the measurable  $S_{VN}$  can never arrive at  $N \ln 2$ , and so the entropic accounting is again supposed to be inconsistent if we consider both  $S_{VN}$  and  $S_{TD}$ .

However, it is clear that their argument is moot given a clear understanding of the sort of thermodynamics we are interested in (see §4.3). While it is true that  $S_{VN}$  will never reach  $N \ln 2$ , recall that  $S_{TD}$  (or, more likely, one of its statistical mechanical

analogues, given the domain of finitely many particles merely *approaching*  $\infty$  rather than  $N = \infty$ ) will likewise never reach  $N \ln 2$ . In other words, it does not matter that we can never perform an infinite series of these measurements, and hence never come to know of  $S_{VN}$  at the thermodynamic limit, since we can likewise never have a thermodynamic entropy equivalent to  $N \ln 2$  unless we are at the thermodynamic limit. The two entropies, then, in fact *correspond* in this case.

What of the second case? Here, **H&S** concede that “arithmetically Von Neumann’s argument goes through at the infinite limit” (**H&S** 2006, 172), which makes sense because, as I have insisted so far, von Neumann’s strategy was never to demonstrate the *strict identity* of  $S_{VN}$  and  $S_{TD}$ , i.e. the correspondence of  $S_{VN}$  and  $S_{TD}$  in *all* domains. Instead, it was to show that  $S_{VN}$  corresponds to  $S_{TD}$  *only in the domain where phenomenological thermodynamics hold*, in all other cases merely *approximating*  $S_{TD}$  in large  $N$  cases or replacing it altogether (in e.g. single-particle cases). I maintain that **H&S**’s main mistake was to confuse the domain where phenomenological thermodynamics hold, with domains where they do not hold.

**H&S** complain that “[...] real physical systems are finite. This means that Von Neumann’s argument does not establish a conceptual identity between the Von Neumann entropy and thermodynamic entropy of physical systems. Identities of physical properties mean that the two quantities refer to the same magnitude in the world.” (**H&S** 2006, 172) In line with what I have said in §4.1, it seems that there was no physically meaningful theoretical term in phenomenological thermodynamics that *could* refer to some quantity in the single-particle case, which was why von Neumann needed to come up with a new measure of entropy to begin with. Furthermore, extending a concept to a new domain does not require strict identity, as we have seen and understood for a long

time in the case of statistical mechanics and phenomenological thermodynamics.

As Peres (2002) summarizes: “There should be no doubt that von Neumann’s entropy. . . is equivalent to the entropy of classical thermodynamics. (This statement must be understood with the same vague meaning as when we say that the quantum notions of energy, momentum, angular momentum, etc., are equivalent to the classical notions bearing the same names).” (Peres 2002, 174) ‘Equivalence’ here should not be understood in terms of strict (or conceptual) identity i.e. correspondence at all domains. Rather, we should understand equivalence loosely as correspondence in the suitable domains of application, and successful extension of old concepts in these domains to new domains. As Peres noted above, ‘equivalence’ should be understood in the context of discovery, where one is trying to develop new concepts which are analogous to old ones in different domains. For von Neumann, we have a theory (phenomenological thermodynamics) that is well-understood, but also another theory (quantum mechanics) that we want to understand in light of the former theory. Finding correspondence provides us with ways to *extend* concepts from the original theory to the new theory: for example, with  $S_{VN}$  we may now define ‘something like’  $S_{TD}$  whereas before there was no way to talk about these cases. The same goes for statistical mechanics: by finding a correspondence between e.g. temperature to mean kinetic energy in the thermodynamic limit, we can *extend* the notion of ‘something like’ temperature beyond its original domain into systems with small numbers of particles, whereas before there was, again, no way to talk about these cases.

I see nothing wrong in these cases in the context of discovery. We should give up a strong and untenable notion of conceptual identity in this context. If so, **H&S**’s objection loses much bite.

They further claim that “the fact that the behavior of the two quantities coincides approximately for a very large number of particles is not enough, because in any ensemble of finite gases there are systems in which the identity will not be true. This means that in a real experiment the Von Neumann entropy is not identical with the thermodynamic entropy.” (H&S 2006, 172) This again reveals a confusion between phenomenological and statistical thermodynamics. If they want to talk about particles *at all*, it seems they must adopt some form of statistical mechanical picture with microscopic variables, given phenomenological thermodynamics’ emphasis on purely macroscopic variables like volume or temperature. Yet, if so, they must recognize that thermodynamic entropy  $S_{TD}$  is in general not strictly identical to statistical mechanical entropy, e.g. the Gibbs entropy or information entropy (briefly discussed in §4.2) *either*. Their complaint about approximate coincidences not being enough for (the relevant sort of) equivalence thus weakens significantly, especially since they must assume some such equivalence (which cannot be strict identity) to even talk about particles within the context of phenomenological thermodynamics to begin with. Furthermore, statistical mechanics is evidently empirically successful in explaining and predicting traditionally thermodynamic phenomena despite this ‘non-equivalence’ – it is not clear why this ‘non-equivalence’ should matter if, for all practical purposes, statistical mechanics is the conceptual successor of thermodynamics. Of course, if they could come up with a principled reason why approximations should not be allowed *period, while* accounting for statistical mechanics’ empirical success in accounting for thermodynamic behavior, then this could change. As of now, I see no such argument forthcoming.

## 5 Conclusion, and Some Open Questions

Given the above, I hope to have shown that **H&S**'s argument against the correspondence of  $S_{\text{VN}}$  and  $S_{\text{TD}}$  – to my knowledge the only one in the philosophical literature – fails to hold in all three cases considered (§4.1 – §4.4), as a result of their misunderstanding about domains where phenomenological thermodynamics should hold and domains where it should not. This is compounded with misunderstandings about the role of approximations and the relevant interpretation of density matrices and ignorance in quantum mechanics. I conclude that their argument fails on the whole; the correspondence holds for now.

Of course, even if **H&S**'s claims were debunked, this does not yet amount to a positive argument for the equivalence between von Neumann entropy and thermodynamic entropy. Even assuming correspondence, correspondence does not entail equivalence. However, the former does provide good *prima facie* reasons to believe the latter, especially given the novel take on correspondence I provided in the end of §3: we can accept the correspondence based on thermodynamic considerations about the Second Law and  $S_{\text{TD}}$  accounting, but *also* based on quantum mechanical considerations about  $S_{\text{VN}}$  accounting. The correspondence supports a 'two-way street' – equivalence – between  $S_{\text{TD}}$  and  $S_{\text{VN}}$ .

While I hope to have conclusively refuted **H&S**'s argument, this is but the beginning of further inquiry into questions arising from this supposed correspondence. Amidst the tangle of entropies, there remains much more housekeeping to be done for philosophers of physics.

## References

- Compagner, A. (1989). Thermodynamics as the continuum limit of statistical mechanics. *American Journal of Physics* 57, 106 – 117. doi: 10.1119/1.16103.
- Deville, A. and Y. Deville (2013). Clarifying the link between von Neumann and thermodynamic entropies. *Eur. Phys. J. H* 38, 57–81. DOI: 10.1140/epjh/e2012-30032-0.
- Earman, J. and J. Norton (1998). Exorcist XIV: The wrath of Maxwell’s Demon. Part I. From Maxwell to Szilard. *Studies in History and Philosophy of Modern Physics* 29, 435–471.
- Earman, J. and J. Norton (1999). Exorcist XIV: The wrath of Maxwell’s Demon. Part II. From Szilard to Landauer and beyond. *Studies in History and Philosophy of Modern Physics* 30, 1–40.
- Einstein, A. (1914/1997). Contributions to Quantum Theory. In *The Collected Papers of Albert Einstein, Volume 6 (English). The Berlin Years: Writings, 1914-1917. (English translation supplement) Translated by Alfred Engel*, pp. 20–26. Princeton University Press: Princeton. Available at: <https://press.princeton.edu/titles/6161.html>.
- Hemmo, M. and O. Shenker (2006). Von Neumann’s Entropy Does Not Correspond to Thermodynamic Entropy. *Philosophy of Science* 73(2), 153–174. <http://www.jstor.org/stable/10.1086/510816>.
- Henderson, L. (2003). The Von Neumann Entropy: A Reply to Shenker. *British Journal for the Philosophy of Science* 54(2), 291–296. <http://www.jstor.org/stable/3541968>.

- Hughes, R. I. G. (1992). *The Structure and Interpretation of Quantum Mechanics*. Cambridge, Mass: Harvard Univ. Press. <https://doi.org/10.2307/2186092>.
- Jaynes, E. T. (1957). Information Theory and Statistical Mechanics. *The Physical Review* 106(4), 620–630.
- Maxwell, J. C. (1878). Tait’s “Thermodynamics”. *Nature* 17, 257–259, 278–280.
- Myrvold, W. (2011). Statistical mechanics and thermodynamics: A Maxwellian view. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 42(2), 237–243.
- Norton, J. (2017). Thermodynamically reversible processes in statistical physics. *American Journal of Physics* 85(135), 135–145.
- Peres, A. (1990). Thermodynamic Constraints on Quantum Axioms. In Zurek, W. H. (Ed.), *Complexity, Entropy, and the Physics of Information, The Proceedings Of The Workshop On Complexity, Entropy and the Physics of Information Held May-June, 1989 in Santa Fe, New Mexico, Santa Fe Institute Studies in the Sciences of Complexity, vol. VIII*, pp. 345–356. Westview Press.
- Peres, A. (2002). *Quantum Theory: Concepts and Methods (2<sup>nd</sup> Ed.)*. Dordrecht: Kluwer Academic Publishers.
- Prunkl, C. (manuscript). On the Equivalence of  $-\text{Tr } \rho \ln \rho$  and Thermodynamic Entropy. Available at <https://arxiv.org/pdf/1810.09585.pdf> (last accessed 30 Aug 2019).
- Shenker, O. (1999). Is  $-ktr(\rho \log \rho)$  The Entropy in Quantum Mechanics? *British Journal for the Philosophy of Science* 50, 33–48.

von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics*. Princeton University Press: Princeton.