Enhanced Photonic Maxwell's Demon with Correlated Baths

Guilherme L. Zanin[1,2,†], Michael Antesberger[1], Maxime J. Jacquet[1,3], Paulo H. Souto Ribeiro[2], Lee A. Rozema[1], and Philip Walther[1,4]

[1]University of Vienna, Faculty of Physics, Vienna Center for Quantum Science and Technology (VCQ), Vienna, Austria

[2]Departamento de Física, Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina 88040-900, Brazil

[3]Laboratoire Kastler Brossel, Sorbonne Université, CNRS, ENS-Université PSL, Collège de France, Paris 75005, France.

[4]Christian Doppler Laboratory for Photonic Quantum Computer, Faculty of Physics, University of Vienna, 1090 Vienna, Austria

[†]Correspondence to: guilherme.zanin@univie.ac.at

Maxwell's Demon is at the heart of the inter-relation between quantum information processing and thermodynamics. In this thought experiment, a demon generates a temperature gradient between two thermal baths initially at equilibrium by gaining information at the single-particle level and applying classical feed-forward operations, allowing for the extraction of work. Here we implement a photonic version of Maxwell's Demon with active feed-forward in a fibre-based system using ultrafast optical switches. We experimentally show that, if correlations exist between the two thermal baths, the Demon can generate a temperature difference over an order of magnitude larger than without correlations, and so extract more work. Our work demonstrates the great potential of photonic experiments – which provide a unique degree of control on the system – to access new regimes in quantum thermodynamics.

1 Introduction

Thermodynamics was conceived as a phenomenological theory for the equilibrium properties of macroscopic systems ranging from gas (ensembles of 'many' small systems) to black holes. For these systems, theories like that of Maxwell's for heat [1] provide a complete description of quantities such as temperature or work, and the manipulation of thermal states – as well as their emergence from fluctuations at the smallest scale – can be described in terms of (quantum) information processing [2–4]. Reciprocally, thermodynamic considerations ought to limit quantum information processing itself [5–7]. This inter-relation is perhaps best manifested in the protocol known as 'Maxwell's demon' (MD) [1, 8–15]. As illustrated in Fig. 1 **(a)**, the MD monitors the motion of gas particles inside a partitioned box. Based on its previous knowledge of the particle's velocity [16, 17], the MD can sort fast particles from slow particles by opening an aperture in the partition such that all fast particles eventually end up on one side of the partition and all slow particles on the other. So the MD brings the system out of equilibrium, and one can then exploit the engineered temperature differential to extract work. In modern experiments [18–29], we say that the MD generates work by measuring individual particles to gain information [16, 17], and depending on these measurement results, it applies different operations to the bath (i.e. opening the door, or leaving it closed). These measurement-dependent operations are called feed-forward operations. Here, work and energy are bound to the available information, and are ultimately physically limited by it [17, 30].

The 'power' of the MD, the temperature imbalance it can generate per measurement, can be greatly enhanced by operating with correlated particles: Consider a situation where, for every particle moving in a certain direction in the left bath, there is a twin particle moving in the same direction in the right bath (see Fig. 1 **(c)**). If the MD observes a particle in the left bath it also learns something about an (unobserved) particle in the right bath. Because of this additional information, the MD can generate a larger temperature difference per measurement [31] by simply modifying its feed-forward protocol to make use of the new information.

While correlations are difficult to engineer in

most systems, doing so is very easy in photonics. Here, we present the first experimental realisation of a photonic MD operating with correlations and active feed-forward. The photonic MD was first proposed and a proof-of-principle experiment was demonstrated in [32], and the idea to enhance the MD's power was proposed in [31].

The working principle of an uncorrelated photonic MD is depicted in Fig. 1 (**e**). In a photonic MD, the baths are realised with optical modes, each containing a thermal state of photons whose average photon number directly sets the temperature of the bath. These two optical modes are then sent to individual detectors, D_A and D_B. When the two optical modes have an equal photon number (the baths have an equal temperature), the detectors register the same counts and no work can be extracted. When a photon is subtracted from this state at a beamsplitter and detected in the reflection port with detector Dem_A or Dem_B, bunching will occur, meaning that the photon number in the transmitted beam will be temporarily doubled (as depicted in red in Fig. 1 (**e**)). This photon extraction provides information about the average photon number in the beam, based on which a classical feed-forward operation may then be applied to create an imbalance in the photon number between the two beams. This energy imbalance can in turn be used to extract work. The gain of information and the generation of a temperature difference through classical feed-forward is precisely the functioning of an MD — a photonic MD is thus a device that can generate a photon number imbalance between two thermal states that initially have the same photon number. The power of a photonic MD is simply quantified by the photon number difference that can be generated to extract work.

The working principle of the photonic MD was demonstrated only recently in a proof-of-principle experiment [32]. There, the feed-forward at the core of the extraction of work was not applied, instead the operation of the MD was shown by post-processing the data. Detectors D_A and D_B were replaced by photodiodes connected to a capacitor. In the case where the photon number at D_A and D_B was balanced, the photo-currents generated by the photodiodes cancelled out, creating a zero net charge on the capacitor. If, however, there were more photons in one beam, the capacitor's polarity could have been flipped (classical feed-forward operation) such that the photo-currents would no longer cancel out and the capacitor would always be charged positively, allowing the extraction of work. In [32], instead of flipping the capacitor's polarity, the operation of the photonic MD was quantified by recording the charge across the capacitor in real time. It was found that a positive (negative) charge of the capacitor would correspond to a click at Dem_A (Dem_B). Hence, no work was extracted.

In our photonic MD, we implement active feed-forward to physically swap the paths of the beams dependent on detection events at Dem_A and Dem_B using an ultra-fast optical switch (methods adapted from [33]). While our active feed-forward would allow for the extraction of work, we do not extract work here. We instead monitor the photon number imbalance directly using single-photon detectors, as this characterises the maximum work that could be extracted. This detection scheme is also more compatible with the quantum states of light that we use to generate correlations to increase the power of the MD.

Correlations in quantum thermodynamics have been used in numerous experiments in quantum thermodynamics [34–44], but not in terms of information processing by the MD. Here, we compare the statistics obtained with uncorrelated thermal states (produced classically) and split thermal states to those obtained with correlated and anticorrelated thermal states produced by spontaneous parametric down-conversion and a path-entangled (so-called N00N [45]) state, respectively. We show that the choice of correlations is essential in enabling the MD to extract work. Our experiments evidence that particular correlations allow for an order of magnitude increase of the power of the MD, while others prevent the MD from extracting any work at all. Furthermore, we derive a figure of merit for the power of the MD in the regime of low photon-numbers and demonstrate that information processing with correlations and anticorrelations vastly outperforms operation with the other states.

Although we generate some of these correlations using quantum states of light, the enhancement of the power of the MD does not rely on

the quantum nature of the correlations. It remains an open question if quantum correlations could further enhance the power of the MD. Nevertheless, correlations are ubiquitous in quantum information processing, and so our photonic implementation settles Maxwell's demon therein.

2 The power of Maxwell's Demon

In our experiment, the MD consists of a logic that monitors photon arrival events at two detectors and conditionally controls the path light takes. Specifically, the experiment (Fig. 2 **(d)**) is based on two separate spatial modes for light propagation, corresponding to two paths (labelled In$_A$ and In$_B$ in Fig. 1 **(e)**) in the setup, in which we send states with various statistics. The MD consists of two single-photon detectors (Dem$_A$ and Dem$_B$), two variable-reflectivity beamsplitters (implemented with tunable directional couplers, labelled TDC), the ultra-fast optical switches (UFOS), and the logic (labelled TTM) used to control the UFOS. The detectors monitor photon arrival events in the reflected port of the TDCs, and send signals to the coincidence logic which then sets the state of the UFOS to route the light.

The feed-forward protocol functions as follows: The thermal states (generated by either of the sources Fig. 2 **(a)**, **(b)** or **(b)**+**(c)**) enter the setup (Fig. 2 **(d)**) in modes In$_A$ and In$_B$. First, the temperature of the thermal baths are set equal by fixing the UFOS in the 'bar state' (to transmit mode A to D_A and B to D_B), and equalising the photon rates N_A and N_B measured by single-photon detectors in modes D_A and D_B. Second, we make the MD's feedforward active. Microscopic knowledge about incoming photons in modes In$_A$ and In$_B$ is acquired via the detector clicks in the reflected arms Dem$_A$ and Dem$_B$. Based on this information, the MD operates the UFOS (or not) in order to route the beam with a higher photon number in mode A or B to D_A, thus creating a photon-number imbalance at the output. The exact details of this routing depend on the states considered and we discuss them below.

To characterise the temperature gradient generated by the MD, we directly measure the photon rates N_A and N_B in the two beams exiting the experiment with single-photon detectors in modes D_A and D_B when the feed-forward is ac-

tive. We then measure $\Delta N = N_A - N_B$, the photon number difference between these detectors. Although we do not extract any work here, work could in principle be extracted from this photon-number difference for example by charging a capacitor as proposed in [32], or by moving a particle via the differential in radiation pressure thus engineered. We consider $\Delta N / N$ — the change in photon number per incident photon — as our figure of merit for the 'power' of the MD. Here, N is the number of photons incident in mode In$_A$ which we will always set equal to the number of photons in mode In$_B$. We estimate N by making the feed-forward inactive, measuring the photon number with the detectors in modes D_A and D_B, and then dividing by the beamsplitter reflectivity. We begin by presenting our derivations of this figure of merit for the different input states in the regime of low photon-numbers, which we define as one- and two-photon detection events.

Uncorrelated thermal states. — We first consider two uncorrelated thermal states with the same photon number, as in Fig. 1 **(e)**. The density matrix of a single-mode thermal state of temperature T ($\beta = 1/(k_b T)$) is $\rho_{\text{Th}} = \sum_{n=0}^{\infty} P_n |n\rangle \langle n|$, where $P(n) = \bar{n}^n \left((1+\bar{n})^{n+1}\right)^{-1}$ and $\bar{n} = (e^{\beta \hbar \omega} - 1)^{-1}$ is the average photon number in this state ($\hbar$ and k_B are the Planck and Boltzmann constant, respectively and ω is the angular frequency of the photon). This state is sent into mode In$_A$ towards the beamsplitter, with vacuum incident into the other unlabelled mode. At the beamsplitter, some of the light is reflected towards detector Dem$_A$ while the rest is transmitted into mode A. The probability to detect n photons at detector Dem$_A$ and m photons in mode A is $P(m,n) = \frac{(n+m)!}{n!m!} \frac{\bar{n}^{n+m}}{(1+\bar{n})^{n+m+1}} R^{2n}(1-R^2)^m$, where R is the reflection amplitude of the beamsplitter [46]. The subtraction of one photon from a thermal state leads to a two-fold increase in the average photon number left in mode A, $\bar{n}' = 2\bar{n}$ [47] (hence the two red photons in Fig. 1 **(e)**). Moreover, if the initial thermal state is a single-mode thermal state, the remaining state is a multimode thermal state [48]. The same is true for the lower B modes. Importantly, neglecting this detection event results in a thermal state with a lower average photon number $\bar{n}' = (1-|R|^2)\bar{n}$ in mode A or B [46]; hence, the MD's measurement does not inject any energy into either mode. As
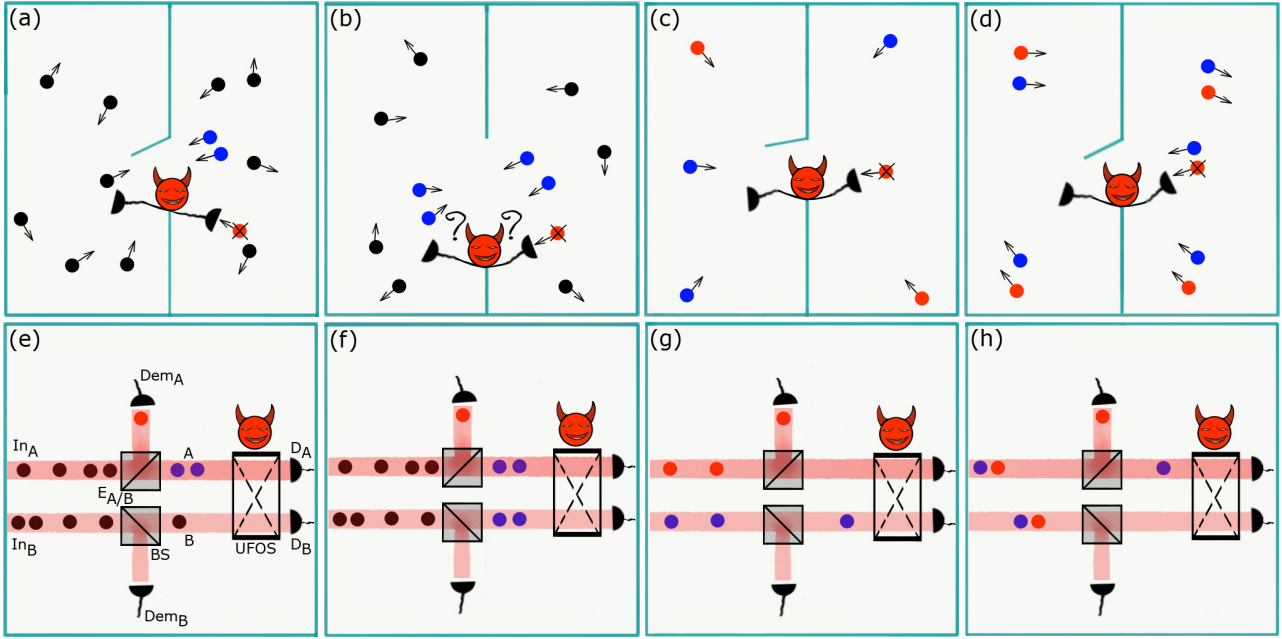
Figure 1: **Principle of operation of the photonic Maxwell's Demon (MD):** the MD controls the path of particles/photons in the system. The colours red and blue represent correlated photons; i.e. when the demon measures a red particle, it learns the location of the blue particle. Top row, particle picture of the photonic MD; bottom row, actual implementation of the photonic MD. **(a)** and **(e)**, uncorrelated thermal states; **(b)** and **(f)**, split thermal states; **(c)** and **(g)** correlated states; **(d)** and **(h)**, anti-correlated states. **(e)** The mode labels shown here also apply to panels **(f)**-**(g)**. $\mathrm{In}_{A/B}$ and $\mathrm{E}_{A/B}$ label the input modes of the MD's beamsplitter. The various thermal states are input into $\mathrm{In}_{A/B}$, while vacuum is always incident in modes $\mathrm{E}_{A/B}$. We use A/B and $\mathrm{Dem}_{A/B}$ to label the output modes of the beamsplitters. Modes $\mathrm{Dem}_{A/B}$ are sent to the MD's detectors, while A/B are sent to UFOS so the MD can swap them if it wishes. Finally, $\mathrm{D}_{A/B}$ label the output modes of the UFOS, which are directly sent to the detectors to measure the photon number imbalance. In the photonic implementation without correlations (**(a)** and **(e)**), a detection at the demon's detector Dem_A heralds more photons in mode A (indicated by the two blue particles), corresponding to an increased temperature. With classical correlations arising from split thermal baths (**(b)** and **(f)**), detection at Dem_A increases the photon number in both mode A and mode B by the same amount, rendering the demon powerless. The MD with correlated (**(c)** and **(g)**) and anti-correlated (**(d)** and **(h)**) baths learns about particles in both baths. Detecting the red photon at Dem_A tells the MD that there is now one more photon in mode B for correlated baths. While detecting the red photon at Dem_A for for anti-correlated baths tells the demon that there are no photons in mode B.

a result, if the MD detects a photon at detector Dem_A the remaining photon number in mode A is higher. Since the MD will attempt to increase the photon number in mode D_A, it will set the UFOS in the 'bar state' so that light incident in mode A (B) is transmitted to mode D_A (D_B). However, detection of a photon at detector Dem_B means that there are more photons in mode B, so the MD will swap modes A and B by setting the UFOS to the 'cross state'. When no photons are detected, or a photon is detected in each mode, the MD sets the UFOS to the 'bar state'. Thus a photon number imbalance can be generated based on the MD's information (click/no-click), which depends on the beamsplitter's reflection amplitude R. In Appendix C, we

derive the power of the MD when using uncorrelated thermal states for low photon-numbers, obtaining

$$\frac{\Delta N}{N} = \frac{2\overline{n}}{(1-\overline{n})^2} R^2 (1-R^2). \qquad (1)$$

Note that here $\overline{n}$ is the expectation value of the photon number per mode (i.e., the number of photons per coherence time, which is $\tau_c = 5.42\mu$s for our uncorrelated thermal photons and $\tau_c \approx 2$ ps for the photons generated by spontaneous parametric down-conversion). In contrast, N is the total number of photons sent into mode $\mathrm{In_A}$ (and equally into mode $\mathrm{In_B}$) during the experiment. Hence, $\overline{n} = \frac{\tau_C}{T} N$, where T is the total measurement time. From Eq. 1, we see that the power of the MD increases with $\overline{n}$ (in the limit
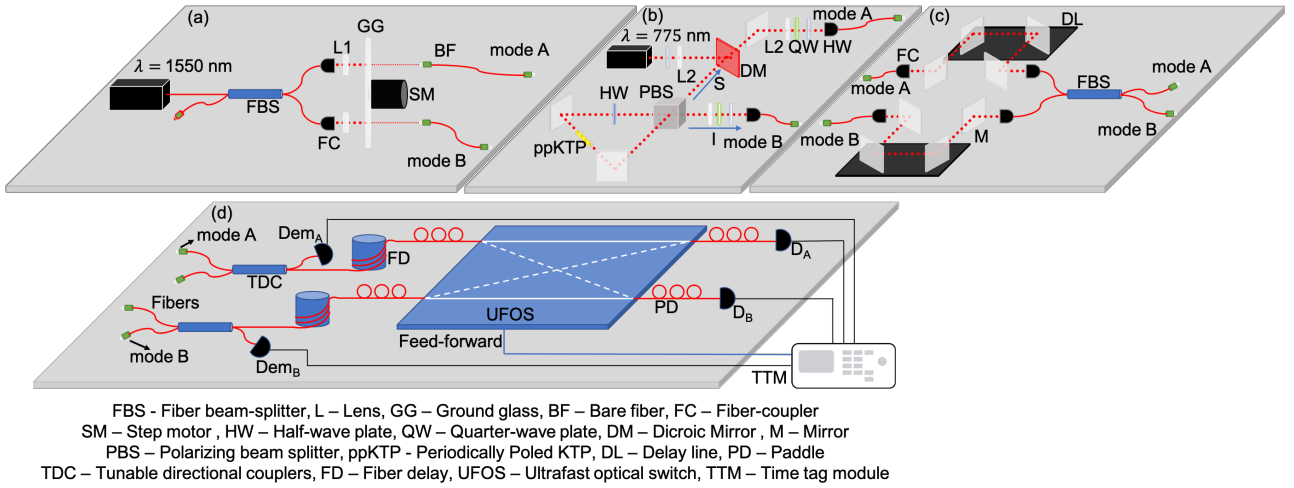
FBS - Fiber beam-splitter, L – Lens, GG – Ground glass, BF – Bare fiber, FC – Fiber-coupler
SM – Step motor , HW – Half-wave plate, QW – Quarter-wave plate, DM – Dicroic Mirror , M – Mirror
PBS – Polarizing beam splitter, ppKTP - Periodically Poled KTP, DL – Delay line, PD – Paddle
TDC – Tunable directional couplers, FD – Fiber delay, UFOS – Ultrafast optical switch, TTM – Time tag module

Figure 2: **Active photonic Maxwell's Demon.** The two photon sources: **(a)**, uncorrelated thermal states and **(b)**, correlated thermal states (Sagnac loop for SPDC). Splitting one mode of one of the sources shown in **(a)** on a beamsplitter (not shown) yields the split thermal state. Passing photons from **(b)** through the Hong-Ou-Mandel interferometer **(c)** yields the two-photon N00N state. The Maxwell's Demon in **(d)** monitors the statistics of the incoming photons with the detectors in modes Dem_A and Dem_B and conditionally controls the path the photons take through the ultra-fast optical switch (UFOS) to the detectors in modes D_A and D_B.

of $\overline{n} \ll 1$), indicating that this effect is driven by the thermal nature of the beams. Since Eq. 1 also depends on the reflection amplitude R, we will tune this parameter to optimise the power of the MD.

Split thermal states. — Now, if instead of two statistically uncorrelated thermal states we send two classically correlated thermal states (created by splitting a single thermal state at a balanced beamsplitter) into the setup, the state in modes In_A and In_B is $\rho_{\mathrm{Split}} = \sum_{n,m=0}^{\infty} P(n,m) |n,m\rangle \langle n,m|_{\mathrm{In}_A,\mathrm{In}_B}$, with $P(n,m)$ as defined above (for $R = 1/\sqrt{2}$). Since the thermal states in each mode originated from the same source, a detection at Dem_A (or equivalently Dem_B) affects the state in both modes A and B: this amounts to effectively increasing the photon number in both arms simultaneously (Fig. 1 **(f)**), as in [49].

Therefore, when operating with split thermal states, the MD cannot possibly know in which arm the bunching effect will be stronger — its power in this case is zero: $\Delta N/N = 0$, for all values of the reflectivity of the MD's beamsplitter (see Appendix C). This result holds in general, not just in the regime of low photon-numbers [31]. In terms of the partitioned box example, we can imagine a hole in the partition through which particles can freely escape to either side such that, when the MD measures one particle, the

temperature is increased on both sides of the partition (Fig. 1 **(b)**). This is not a perfect analogy, as in our photonic implementation, particles are not exchanged after the MD's measurement. Rather, the thermal state is distributed among two optical modes and photon subtraction in either of these modes increases the average photon number equally in both modes, as in [49] for example.

Correlated thermal states. — Unlike the above correlations, which effectively make the MD powerless, the correlations generated by type-II spontaneous parametric down conversion (SPDC) increase its power [31]. Ideally, for this state whenever there is a photon in mode In_A, a twin photon is found in mode In_B (correlated particles are denoted as the red and blue particles in Fig. 1 **(c)** and **(g)**). In particular, the ideal state is $|\psi\rangle = \frac{1}{\cosh s} \sum_{n=0}^{\infty} (\tanh s)^n |n,n\rangle_{\mathrm{In}_A,\mathrm{In}_B}$, where s specifies the brightness of the source. Although photons are always emitted in pairs, tracing over photons in mode In_A (In_B) leaves photons in mode In_B (In_A) in a thermal state [46], which means that the power of the MD could be enhanced as with uncorrelated thermal states. However, we will now see that the correlations generated by SPDC have a much larger effect, and in fact counteract the thermal bunching.

The key resource for this state are the correlations which provide the MD with more informa-

tion than it can obtain from the thermal bunching. For example, if exactly one photon is detected at Dem_A, the MD has removed one photon from mode A, and it now knows that there will be one more photon in mode B (and vice-versa). This is illustrated Fig. 1 **(g)**, where the photons come in pairs. For every red photon in mode Dem_A there is a correlated blue photon in mode B. Hence, when the MD detects a photon in Dem_A there are now zero photons in mode A, and exactly one photon in mode B. Therefore, the MD will move the remaining photon from B to mode D_A by putting the UFOS in the cross state. Similarly, it will set the UFOS to the bar state when it detects a photon in mode B. As before, the UFOS remain in bar for all other detection events

In the partitioned box analogy, Fig. 1 **(c)** shows the MD gaining information about particles on both sides of the partition, by measuring a particle on only one side. Here the colour denotes correlated pairs of particles. For every red (blue) particle on the left there is a corresponding blue (red) particle on the right moving in a mirrored direction. Thus, if the MD finds a particle impinging on the aperture from one side, it measures (and destroys) that particle leaving its twin to arrive alone from the other side.

Notice that now the MD's logic is reversed compared to the uncorrelated thermal states. Previously, finding a particle on one side implied that there were more particles on that side. Now the opposite is true. This also introduces a new experimental issue: the power of the MD critically depends on the ratio between correlated photons (which experimentally correspond to coincidence events) and uncorrelated photons (by which we mean photons that have lost their partner due to experimental loss). In practice that means that the coupling efficiency ϵ^2, defined as the ratio between the singles rate the coincidence rate, is essential to create a powerful demon.

In Appendix C, we derive the power of the MD using correlated thermal states in the regime of low photon numbers in the presence of experimental loss. We obtain

$$\frac{\Delta N}{N} = 2\epsilon^2 R^2 (1 - R^2). \qquad (2)$$

This equation resembles Eq. (1). Here, however, the power of the MD does not depend on $\bar{n}$, since the correlations are the driving force of the power

of the demon in this regime. In fact the correlations counteract the thermal bunching effect. Moreover, notice that the power of the MD is scaled by the coupling efficiency ϵ^2. For a perfect source $\epsilon^2 = 1$. Interestingly, we can experimentally probe this, by simply computing the power of the MD per photon pair (by normalising ΔN by the total number of coincidences C). As we show in Appendix C, this results in:

$$\frac{\Delta N}{C} = 2R^2 (1 - R^2). \qquad (3)$$

Anti-correlated thermal states. Finally, we consider the operation of the MD with anti-correlated N00N states (here we set $N = 2$) [50]. We refer to this state as anti-correlated because whenever there is a photon in mode In_A, there are none in mode In_B (see Fig. 1h). In Ref. [31], a specific two-mode state was proposed with two important properties. First, for every N-photon event, the two modes are in a N00N state — where there are N photons mode A and zero in mode B, and vice-versa. Second, tracing out either mode, leaves the remaining mode in a thermal state. Experimentally, we do not create this state, but we note that some experimentally-realised N00N state sources may come close to these conditions [51, 52]. Instead, here we perform Hong-Ou-Mandel (HOM) interference between the two photons from the SPDC source to create a 2-photon N00N state. The higher-order modes will not be in a N00N state; rather, they will be in so-called Holland-Burnett states [53]. Nevertheless, as we show in Appendix C for low-photon numbers, the demon effect is driven by the two-photon component and in this regime our source behaves almost exactly as the state proposed in [31]. Ideally, our state has only the vacuum component and the two-photon terms $(|2,0\rangle_{\text{In}_A,\text{In}_B} + |0,2\rangle_{\text{In}_A,\text{In}_B})/\sqrt{2}$. Starting from this state, in Appendix C we derive the power of the MD in the presence of loss (coupling efficiency ϵ^2) and for imperfect Hong-Ou-Mandel visibility (v^2), finding

$$\frac{\Delta N}{N} = 2\epsilon^2 (2v^2 - 1) R^2 (1 - R^2). \qquad (4)$$

This differs from operation with correlated states (Eq. (2)) by a factor depending on the visibility $2v^2 - 1$. Here, the key resource are the anti-correlations: In Fig. 1 **(g)** photons come from the source in pairs, denoted by red and blue.
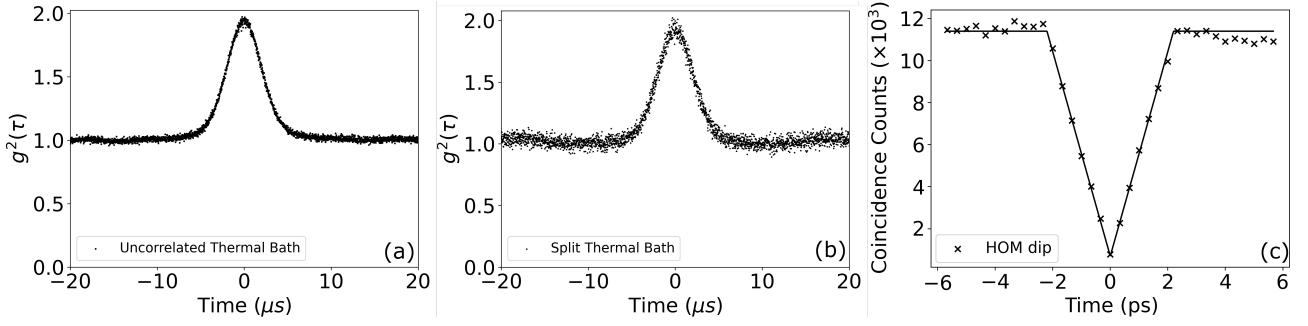
Figure 3: **Source characterisation.** Representative measured $g^{(2)}(\tau)$ of the thermal beam **(a)** and one mode of the split thermal beam **(b)**. Both display $g^{(2)}(\tau = 0)$ very close to 2. Panel **(c)** shows the Hong-Ou-Mandel dip (visibility $v^2 = 0.87 \pm 0.02$) between the signal and idler down-converted photons with act as our resource for the anti-correlated states.

If the MD detects, say, exactly one photon at Dem_A, then it knows that a second photon is left in mode A, and there are no photons in mode B. In analogy, Fig. 1 **(d)** shows the MD gaining information about both particles by measuring only one of them: if the MD detects, say, the red particle then it knows that another one is also arriving from the same side. Here, as with the quantum correlated states above, the power of MD depends on the coupling efficiency, but also on the visibility (i.e. the quality of the N00N state source). Again, if we instead normalise our results by the average coincidence rate C, we get closer to the ideal performance

$$\frac{\Delta N}{C} = 2(2v^2 - 1)R^2(1 - R^2), \qquad (5)$$

but this does not remove the visibility dependence.

Finally, we remark that in the regime of low photon numbers and for optimal visibility $v^2 = 1$, operation with correlated and anti-correlated states would provide the same power of the MD. However, for anti-correlated states, the demon switches particles in the same manner as for the thermal bunching. This is why at higher temperatures (photon numbers) Ref. [31] found that the anti-correlated states greatly out perform the correlated states.

3   Experiment

Our photonic MD experiment, shown in Fig. 2, is built around a 2x2 UFOS (BATi 2x2 Nanona fibre switch). This optical switch can route light from two input modes into two output modes with a variable splitting ratio. The response time

of the UFOS is below 60 ns, with a maximal duty-cycle of 1 MHz, and a cross-channel isolation greater than 20 dB for any polarisation, see Ref. [33] for more details. The reflectivity R of the variable beamsplitter, implemented by a TDC, sets the amount of information gained by the MD about the input state. After the BS, photons go through a 152 m fibre delay line [1] before entering the UFOS, where the path they take (towards either detector D_A or D_B) is conditionally controlled by the MD. All detectors are superconducting nanowire single-photon detectors (SNSPD) from PhotonSpot Inc. (deadtime 50 ns, average system detection efficiency 90%, and timing jitter of $\approx$ 150 ps). Detection events are recorded and processed by a commercial time tagging module (UQDevices Logic16 TTM). A photon detection event triggers a wave-function generator (Keysight 33500B Series) to generate a 2.5 μs TTL pulse that will make the UFOS switch states.

Before presenting our experimental results, we will explain our measurement procedure which is used for the four different initial states. Our target metric is ΔN, which we then normalise by either the total singles count N or the coincidence count C. However, this is highly sensitive to variations in the imbalance between photon numbers at the input, and thus also to losses in the setup, which can (and do) differ slightly between the upper A modes and B lower modes. In order to account for these unbalanced losses, we first balance the count rates of the two arms

---

[1]The extra fibre relaxes the constraints on the adjustment of electronic delays, pulse size, and other parameters that are more easily controlled than the physical length of optical fibres and BNC cable.

when the UFOS is fixed in the bar state and then measure the counts at the output. This ensures that $\Delta N_{\text{bar}} = N_A - N_B = 0$, where N_A (N_B) refers to the total number of photons detected in mode D_A (D_B). Second, we measure the intrinsic imbalance of the setup by fixing the UFOS to the cross state, and preventing the MD from controlling the path by feed-forward. In this configuration we measure $\Delta N_{\text{cross}} = N_A - N_B$, which represents the maximum imbalance the MD could 'accidentally' achieve without properly using information from either thermal bunching or the correlations between the states. We then measure the photon number difference with the MD being active, i.e. we turn on the feed-forward logic (which differs depending on the input state), yielding $\Delta N_{\text{FF}} = N_A - N_B$. Given these measurements, we construct the unnormalised power of the MD as

$$\Delta N := \Delta N_{\text{FF}} - \Delta N_{\text{cross}}. \qquad (6)$$

Finally, we need to normalise this by either the number of input single photons N or the number of incident photon pairs C. To estimate the N, we simply measure the total number of counts at the end of the experiment at the detectors in modes D_A and D_B for each reflectivity of the TDC. We then divide these rates by the reflectivity to obtain the singles rates arriving at the beamsplitter, and average over the counts in modes A and B. To estimate the incoming coincidence counts, we perform a similar protocol. However, due to limitations in the counting logic, we only measured the coincidence rates between detectors D_A and Dem_A (C_{D_A, Dem_A}) and detectors D_A and Dem_B (C_{D_A, Dem_B}) for a reflectivity of 50%. From these measurements we compute the input coincidence rates for both the correlated and anti-correlated input states (note that this calculation differs slightly for the N00N state and SPDC state).

Uncorrelated thermal state. — To generate uncorrelated thermal baths we make two independent thermal states of light, using a rotating ground glass wheel (see Appendix A). We verify the thermal nature of the source by measuring $g^{(2)}(\tau = 0)$, which is typically ≈ 1.95. A representative measurement is presented in Fig. 3 **(a)**. Since $g^{(2)}(\tau = 0) > 1$ is only a necessary condition for a thermal state, we further verify that without the rotating ground glass $g^{(2)}(\tau = 0) \approx 1$. With these measurements, and

knowledge of our experiment, we can be confident that we have generated a thermal state. By fitting to the $g^{(2)}(\tau)$, assuming a Gaussian thermal beam of the form $g^{(2)}(\tau) = 1 + e^{-\pi(\tau/\tau_c)^2}$ [46], we obtain a coherence time of $\tau_c = 5.42~\mu$s.

We then proceed as described above, setting $\Delta N_{\text{bar}} = 0$ and measuring ΔN_{cross} and ΔN_{FF} as we tune the MD's reflectivity from 0% to 50%. For every reflectivity we reconstruct the power of the MD, defined above. The result is plotted in Fig. 4 as red stars. For these measurements, we set the average incident photon number in modes In_A and In_B to $27,800 \pm 500$ cps. We fit the experimental data with a single parameter, the average photon number per mode $\overline{n}$. Given the 5.42μs coherence time of the pseudo thermal light and the photon rate, we expect $\overline{n} = 0.151 \pm 0.002$ (the error bars are the standard deviation of the measured singles rates for different reflectivities). However, we find a much better fit with a value of $\overline{n} = 0.05$.

The discrepancy in $\overline{n}$ is likely due to experimental limitations in our switching logic. We set the switching time to 2.5μs, which is less than the photon coherence time. However, setting it higher limits our maximum count rate, since the wave-function generator cannot respond between time it is triggered and the output pulse is being generated. Because of our reduced switching time, some bunched events may be missed, thus limiting the power of the MD. We use this value in our theoretical calculation shown in Fig. 4 **(a)** (solid red curve). Note, we also plot this data in Fig. 4 **(b)**. Although it is experimentally possible for us to increase the average photon number, the long coherence time of our thermal distribution would then take us out of the regime of low photon numbers, which is the basis of the fair comparison of the power of the MD with other input states.

Split thermal state. — We produce this input state by simply splitting one of the thermal states at a balanced beamsplitter after the ground glass wheel. We first verify that this still results in a thermal state in each mode by measuring the $g^{(2)}(\tau)$ after the beamsplitter, a representative $g^{(2)}(\tau)$ is plotted in Fig. 3 **(b)**, displaying a $g^{(2)}(\tau = 0)$ close to 2.

After this, we balance the count rates and measure the power of the MD as before. The result is plotted in Fig. 4 **(a)** as the violet diamonds. For
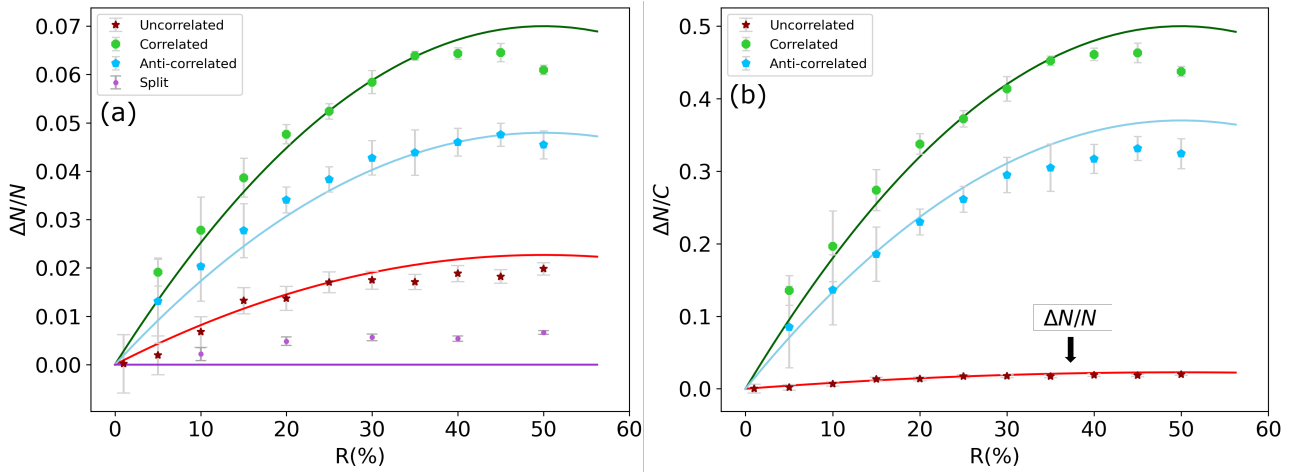
Figure 4: **Power of Maxwell's Demon. (a)**, The photon number difference, ΔN that the Demon can achieve per incident photon: Eq. (6) for various correlations. Solid curves are the theory for: uncorrelated state (1) with $\overline{n} = 0.05$ (red); split thermal state $\Delta N/N = 0$ (violet); correlated state (2) with $\epsilon^2 = 0.14$ (turquoise); anti-correlated state (4) with $\epsilon^2 = 0.14$ and $v^2 = 0.87$ (green). **(b)**, The same ΔN data, but here the correlated and anti-correlated data are normalised by the correlation rate to compute the demon's power per incident pair. Note, the uncorrelated data plotted in red in this panel is normalised by the singles counts, and is included here for comparison. The solid curves are the theory for: uncorrelated state (1) with $\overline{n} = 0.05$ normalised with coincidence rate (red); correlated state (3) (turquoise); anti-correlated state (5) with $v^2 = 0.87$ (green). The experimental errors are the standard deviation of the measurement set.

these data, $\overline{n} \approx 0.7$ in each spatial mode at the input. Although for this $\overline{n}$, our low-photon approximation no longer applies, Ref. [31] showed that our result is independent of $\overline{n}$. Experimentally, this corresponds to incident singles rates in modes In_A and In_B of $293,000 \pm 3000$ cps. The resulting experimentally measured $\Delta N/N$ is very close to zero, as expected. The small deviation from zero is likely caused by slightly imbalanced losses in the setup, such that although ΔN_{OFF} is close to 0 the real photon number before the UFOS is not exactly balanced.

Correlated thermal states. — We produce correlated thermal states by SPDC. In this source (described in Appendix A) the coherence time is ≈ 2 ps (which we verify by a Hong-Ou-Mandel experiment in Fig. 3 **(c)**). This is much smaller than both the minimum resolution of our time tagger (which is 156 ps), and the timing jitter of our detectors (≈ 200 ps). Hence, we cannot measure the $g^{(2)}(\tau = 0)$ as the width of the expected peak is below the minimum resolution of our detection system. However, it is now well established that in such sources the signal (idler) beam is indeed thermal when the idler (signal) beam is traced out [54, 55].

For this state, we operate with a balanced average photon rate of $168,000 \pm 1000$ cps in modes

In_A and In_B. Although this is a higher count rate than for the uncorrelated states, the shorter 2 ps coherence time means the average photon number per mode is much smaller $\overline{n} \approx 10^{-7}$. We measure the power of the MD per incident photon by varying R (green circles in Fig. 4 **(a)**). In this case, the theoretical curve (Eq. (2)) depends only on the coupling efficiency ϵ^2. This is calculated by measuring the coincidence rate between the detectors and correcting for the reflectivity of each BS loss and dividing by the singles rate, yielding $\epsilon^2 = 0.14$. The resulting solid green curve fits the data very well.

In Fig. 4 **(b)**, we also plot the power of the MD per incident photon pair, by normalising ΔN by the coincidence rate (green circles). This removes the effect of almost all experimental imperfections, and our experimentally measured power of the MD almost reaches the theoretical maximum value of 0.5 (see Eq. (3)). Moreover, the experimentally measured data is in good agreement with the full theoretical curve (solid green line).

Anti-correlated thermal states. — We generate the input 2-photon N00N state by HOM interference between the two photons of our SPDC source. The details are described in Appendix A. We estimate the quality of the N00N state by
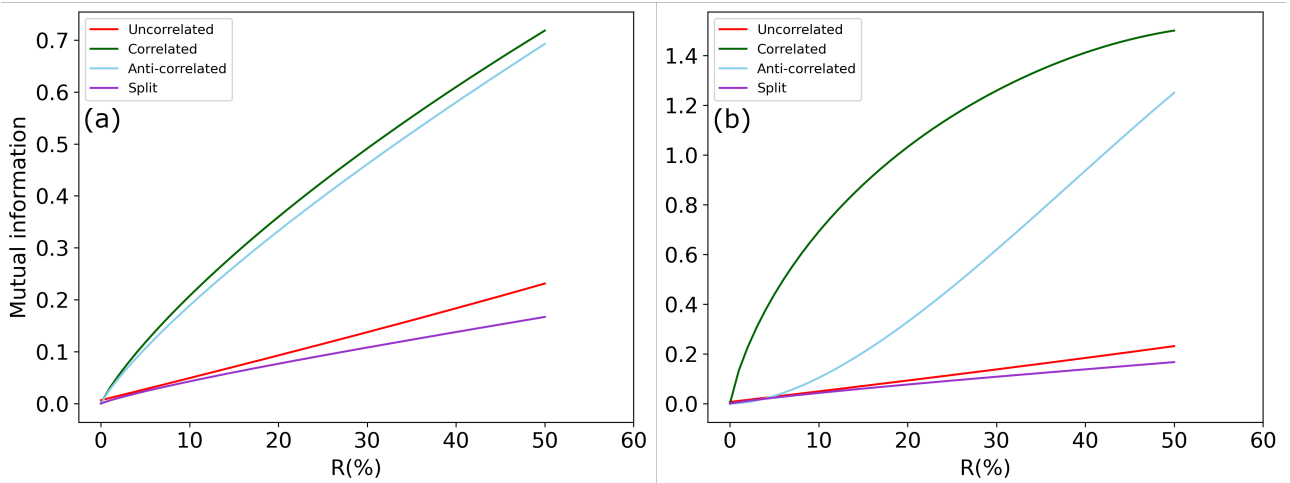
Figure 5: **MD's Information.** The mutual information between the MD's measurement outcomes and the photon number a) for our experimental parameters, and b) for the ideal situation. The parameters used to evaluate these curves are the same as those used in the fits to the data presented in Fig. 4. In all cases, as the MD's beamsplitters reflectively increases, it gains more information about the photon number. Moreover, for all of the parameter regimes investigated in this work the correlated and anti-correlated states out perform the uncorrelated thermal states.

scanning the relative delay and measuring the visibility of the HOM dip — we measure $v^2 = 0.87 \pm 0.02$. To ensure that our state is classically anti-correlated in the Fock basis, we observe the absence of correlations (coincidences) at a time delay of zero (Fig. 3 **(c)**), which implies that the two photons incident on the fiber beamsplitter in Fig. 2 **(c)** bunch and exit in the same mode. As mentioned above, this does not result in thermal states in modes A and B, but we show in Appendix C that the behaviour is almost identical in the low-photon number regime.

For this state, we operate with balanced singles rates of $141,000 \pm 3000$ cps, and, again, $\overline{n} \approx 10^{-7}$. Proceeding as before, we obtain the power of the MD and plot the result normalised by singles as turquoise pentagons in Fig. 4 **(a)**. For the corresponding theory curve (Eq. (4)), we use our measured HOM visibility of $v^2 = 0.87$ and take $\epsilon^2 = 0.14$ (as above). In Fig. 4 **(b)**, we normalise ΔN by the measured coincidence rate. The theory curve, Eq. (5), only requires the visibility which we take as above, and plot as the solid turquoise line. Although ideally the correlated and anti-correlated states behave identically for low photon numbers, the anti-correlated states suffer an additional imperfection arising from the HOM interference.

4 Discussion

Firstly, we note that the agreement between the experimental data and the theory is very good for all states, with a slight deviation at high reflectivity. This deviation comes from our simplified theory, which does not include the MD's detection efficiency. Indeed, in the theory of [31] it can be observed that changes in the relative detection efficiency between detectors Dem_A and Dem_B can shift the dependence of ΔN on reflectivity. Moreover, we stress that only one data set, that of the uncorrelated thermal states normalised by singles, required a fitting parameter. All other curves presented in Fig. 4 agree with experimental data using only independently measured experimental parameters, such as the coupling efficiency and HOM visibility. This demonstrates the great potential of relatively simple optical experiments to investigate information processing scenarios characteristic of thermodynamics using pseudo-thermal beams, which moreover give access to a wealth of exotic correlations readily obtained with photons.

We have experimentally demonstrated that the power of the MD with correlated and anticorrelated baths (which we produce with entangled quantum states) vastly outperforms operation with classical states (uncorrelated and split thermal states), as predicted in [31]. For example, when considering the power of the MD

per photon (i.e. the data normalised to the incident single-photon rate, plotted in Fig. 4 **(a)**), the power of the MD is a factor of ≈ 5 higher with correlated baths than with uncorrelated baths (green curve compared to the red curve in panel **(a)**. Moreover, the enhancement is even more striking when considering the power of the MD per photon pair, plotted in Fig. 4 **(b)**. This analysis serves to essentially filter out the single photons that have lost their partner, and allows us to reach the theoretical limits for the correlated baths. In particular, although operating with uncorrelated thermal states does yield a finite power, this is outperformed by a factor 16 when operating with anti-correlations (blue curve compared to the red curve in panel **(b)**), and by a factor 28 when operating with correlated states (green curve compared to the red curve in panel **(b)**). Here we are comparing to the power of the MD with uncorrelated thermal baths normalised by the singles rate (plotted in red in both panels **(a)** and **(b)**), since the thermal state is robust to loss.

The MD's ability to extract work is limited by the amount of information it can extract. While this was originally formalized by Landauer's erasure principle [2, 17], applying this concept in a quantitative manner to our system would require a generalized Jarzynski equality such as that derived in [56]. Nonetheless, the mutual information I between the MD's measurement outcomes in modes Dem_A and Dem_B and the photon number in mode A and mode B can be used to qualitatively explain the enhancement of the power of the MD in the presence of the correlated input states. In Appendix D, we derive I for each input state, and we plot the results in Fig. 5. Note that MD has 4 measurement outcomes so the maximum possible entropy is 2. In panel **(a)** we have used the coupling efficiency and N00N state visibility used to fit to our experimental data normalized by the singles, while in panel **(b)** we used the parameters used to fit to the data normalized by the coincidence rate. In both panels $\bar{n} = 0.05$ for the uncorrelated and split thermal cases. These plots demonstrate that, for the both the correlated and anti-correlated states, the MD has access to more information, which is consistent with our observation that it can generate a larger ΔN for these data. Interestingly, for the split thermal state the MD only has marginally less informa-

tion than it does for the uncorrelated states. It is hence possible that, for an operation that is more complex than simply the swapping the modes, the MD could still operate with this state.

In our analysis we have not considered the absolute efficiency of the MD, but rather focused on the enhancement provided by correlations in the baths. In the original thought experiment, the cost of both the demon's measurement of the particle's velocity and its cost to open the aperture was assumed to be negligible. This was later formalized by Landauer, who showed that if a reversible process is used for the measurement and sorting, then the cost (i.e. entropy increase) from this step is zero. He then famously showed that at some point the MD must erase its measurement results, which will increase the entropy [2, 17]. In our experiment, we used irreversible measurements, and hence we are not in the idealized regime. In principle, however, a photonic platform could reach this regime. For example, the photon that the demon subtracts could be used to trigger a unitary (and hence reversible) controlled-path gate [57] that operates on multiparticle input states. Nevertheless, the measurement and switching in our experiment do not inject any external energy into the baths. Note that the MD's beamsplitters do extract energy from the baths, but since this extraction is balanced in Mode A and Mode B this is not the source of the power of the MD. Thus the enhancement we observed here would also be present in the ideal scenario.

Our work is valid for low photon numbers (*i.e.* low temperature of the thermal beams). For higher photon numbers, we expect the anti-correlated baths to perform the best, since the anti-correlations act in the same 'direction' as the thermal correlations. Thus a natural extension of our work would be to create a superposition of N00N states for various N [31]. To fully take advantage of this effect, one would need to create such states with a higher $\bar{n}$, which would likely require the use of a pulsed laser when generating the photon pairs by SPDC.

To conclude, we have implemented the first photonic Maxwell's Demon without post-processing, and verified that operating with correlations enhances the power of the MD to process the information it gains and generate a temperature difference (which enables the extrac-

tion of work) through classical feed-forward by over an order of magnitude. Our work lays the foundations for future investigations of the interrelation between quantum information processing and thermodynamics at the microscopic scale. For example our work calls for a theoretical derivation of the fluctuation relations [56] that could describe the 'experimental information' obtained by the MD in each round of our feed-forward experiment. Finally, although we used entangled states to generate correlations, our MD protocol only required classical correlations. It would be interesting to develop MD protocols that exploit the quantum entanglement that is readily available in our system.

Data Availability

All of the data that are necessary to replicate, verify, falsify and/or reuse this research is available online at [58].

Acknowledgements

References

[1] J. Clerk. Maxwell. *Theory of heat*. Longman, 1871.

[2] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, 1961. DOI: 10.1147/rd.53.0183.

[3] Lidia del Rio, Lea Kraemer, and Renato Renner. Resource theories of knowledge. *arXiv:1511.08818 [cond-mat, physics:math-ph, physics:quant-ph]*, November 2015. DOI: 10.48550/arXiv.1511.08818.

[4] John Goold, Marcus Huber, Arnau Riera, Lídia del Rio, and Paul Skrzypczyk. The role of quantum information in thermodynamics—a topical review. *Journal of Physics A: Mathematical and Theoretical*, 49(14): 143001, April 2016. DOI: 10.1088/1751-8113/49/14/143001.

[5] Jonathan Oppenheim, Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Thermodynamical approach to quantifying quantum correlations. *Phys. Rev. Lett.*, 89:180402, Oct 2002. DOI: 10.1103/PhysRevLett.89.180402.

[6] Wojciech Hubert Zurek. Quantum discord and maxwell's demons. *Phys. Rev. A*, 67: 012320, Jan 2003. DOI: 10.1103/PhysRevA.67.012320.

[7] Koji Maruyama, Fumiaki Morikoshi, and Vlatko Vedral. Thermodynamical detection of entanglement by maxwell's demons. *Phys. Rev. A*, 71:012108, Jan 2005. DOI: 10.1103/PhysRevA.71.012108.

[8] William Thomson. Kinetic theory of the dissipation of energy. *Nature*, 9(232):441–444, Apr 1874. DOI: 10.1038/009441c0.

[9] Harvey S. Leff and Andrew F. Rex, editors. *Maxwell's Demon: Entropy, Information, Computing*. Princeton University Press, 1990.

[10] M. B. Plenio and V. Vitelli. The physics of forgetting: Landauer's erasure principle and information theory. *Contemporary Physics*, 42(1):25–60, January 2001. DOI: 10.1080/00107510010018916.

[11] Koji Maruyama, Franco Nori, and Vlatko Vedral. Colloquium: The physics of Maxwell's demon and information. *Reviews of Modern Physics*, 81(1):1–23, January 2009. DOI: 10.1103/RevModPhys.81.1.

[12] Juan M. R. Parrondo, Jordan M. Horowitz, and Takahiro Sagawa. Thermodynamics of information. *Nature Physics*, 11(2):131–139, February 2015. DOI: 10.1038/nphys3230.

[13] Andrew Rex. Maxwell's Demon—A Historical Review. *Entropy*, 19(6):240, June 2017. DOI: 10.3390/e19060240.

[14] Konstantin Beyer, Kimmo Luoma, and Walter T. Strunz. Steering Heat Engines: A Truly Quantum Maxwell Demon. *Physical Review Letters*, 123(25):250606, December 2019. DOI: 10.1103/PhysRevLett.123.250606.

[15] Meng-Jun Hu, Xiao-Min Hu, and Yong-Sheng Zhang. Maxwell Demon and Quantum Non-locality. *arXiv:2105.05656 [quant-ph]*, May 2021. DOI: 10.48550/arXiv.2105.05656.

[16] L. Szilard. über die entropieverminderung in einem thermodynamischen system bei eingriffen intelligenter wesen. *Zeitschrift für Physik*, 53(11):840–856, Nov 1929. DOI: 10.1007/BF01341281.

[17] Charles H. Bennett. The thermodynamics of computation—a review. *International Journal of Theoretical Physics*, 21(12):905–940, December 1982. DOI: 10.1007/BF02084158.

[18] Shoichi Toyabe, Takahiro Sagawa, Masahito Ueda, Eiro Muneyuki, and Masaki Sano. Experimental demonstration of information-to-energy conversion and validation of the generalized jarzynski equality. *Nature Physics*, 6(12):988–992, Dec 2010. DOI: 10.1038/nphys1821.

[19] Antoine Bérut, Artak Arakelyan, Artyom Petrosyan, Sergio Ciliberto, Raoul Dillenschneider, and Eric Lutz. Experimental verification of landauer's principle linking information and thermodynamics. *Nature*, 483(7388):187–189, Mar 2012. DOI: 10.1038/nature10872.

[20] J. V. Koski, V. F. Maisi, T. Sagawa, and J. P. Pekola. Experimental observation of the role of mutual information in the nonequilibrium dynamics of a maxwell demon. *Phys. Rev. Lett.*, 113:030601, Jul 2014. DOI: 10.1103/PhysRevLett.113.030601.

[21] Maxime Debiossac, David Grass, Jose Joaquin Alonso, Eric Lutz, and Nikolai Kiesel. Thermodynamics of continuous non-markovian feedback control. *Nature Communications*, 11(1):1360, Mar 2020. DOI: 10.1038/s41467-020-15148-5.

[22] Stella Seah, Stefan Nimmrichter, and Valerio Scarani. Maxwell's lesser demon: A quantum engine driven by pointer measurements. *Phys. Rev. Lett.*, 124:100603, Mar 2020. DOI: 10.1103/PhysRevLett.124.100603.

[23] Tamir Admon, Saar Rahav, and Yael Roichman. Experimental realization of an information machine with tunable temporal correlations. *Phys. Rev. Lett.*, 121:180601, Nov 2018. DOI: 10.1103/PhysRevLett.121.180601.

[24] Baldo-Luis Najera-Santos, Patrice A. Camati, Valentin Métillon, Michel Brune, Jean-Michel Raimond, Alexia Auffèves, and Igor Dotsenko. Autonomous maxwell's demon in a cavity qed system. *Phys. Rev. Research*, 2:032025, Jul 2020. DOI: 10.1103/PhysRevResearch.2.032025.

[25] Aishwarya Kumar, Tsung-Yao Wu, Felipe Giraldo, and David S. Weiss. Sorting ultracold atoms in a three-dimensional optical lattice in a realization of maxwell's demon. *Nature*, 561(7721):83–87, Sep 2018. DOI: 10.1038/s41586-018-0458-7.

[26] Nathanaël Cottet, Sébastien Jezouin, Landry Bretheau, Philippe Campagne-Ibarcq, Quentin Ficheux, Janet Anders, Alexia Auffèves, Rémi Azouit, Pierre Rouchon, and Benjamin Huard. Observing a quantum maxwell demon at work. *Proceedings of the National Academy of Sciences*, 114(29):7561–7564, 2017. DOI: 10.1073/pnas.1704827114.

[27] W.-B. Wang, X.-Y. Chang, F. Wang, P.-Y. Hou, Y.-Y. Huang, W.-G. Zhang, X.-L. Ouyang, X.-Z. Huang, Z.-Y. Zhang, H.-Y. Wang, L. He, and L.-M. Duan. Realization of quantum maxwell's demon with solid-state spins *. *Chinese Physics Letters*, 35(4):040301, apr 2018. DOI: 10.1088/0256-307x/35/4/040301.

[28] Y. Masuyama, K. Funo, Y. Murashita, A. Noguchi, S. Kono, Y. Tabuchi, R. Yamazaki, M. Ueda, and Y. Nakamura. Information-to-work conversion by maxwell's demon in a superconducting circuit quantum electrodynamical system.

Nature Communications, 9(1):1291, Mar 2018. DOI: 10.1038/s41467-018-03686-y.

[29] G Engelhardt and G Schaller. Maxwell's demon in the quantum-zeno regime and beyond. *New Journal of Physics*, 20(2):023011, feb 2018. DOI: 10.1088/1367-2630/aaa38d.

[30] Takahiro Sagawa and Masahito Ueda. Second law of thermodynamics with discrete quantum feedback control. *Phys. Rev. Lett.*, 100:080403, Feb 2008. DOI: 10.1103/PhysRevLett.100.080403.

[31] Angeline Shu, Jibo Dai, and Valerio Scarani. Power of an optical maxwell's demon in the presence of photon-number correlations. *Phys. Rev. A*, 95:022123, Feb 2017. DOI: 10.1103/PhysRevA.95.022123.

[32] Mihai D. Vidrighin, Oscar Dahlsten, Marco Barbieri, M. S. Kim, Vlatko Vedral, and Ian A. Walmsley. Photonic maxwell's demon. *Phys. Rev. Lett.*, 116:050401, Feb 2016. DOI: 10.1103/PhysRevLett.116.050401.

[33] Guilherme Luiz Zanin, Maxime J Jacquet, Michele Spagnolo, Peter Schiansky, Irati Alonso Calafell, Lee A Rozema, and Philip Walther. Fiber-compatible photonic feed-forward with 99% fidelity. *Opt. Express*, 29(3):3425–3437, Feb 2021. DOI: 10.1364/OE.409867.

[34] R. Dorner, S. R. Clark, L. Heaney, R. Fazio, J. Goold, and V. Vedral. Extracting quantum work statistics and fluctuation theorems by single-qubit interferometry. *Phys. Rev. Lett.*, 110:230601, Jun 2013. DOI: 10.1103/PhysRevLett.110.230601.

[35] Augusto J. Roncaglia, Federico Cerisola, and Juan Pablo Paz. Work measurement as a generalized quantum measurement. *Phys. Rev. Lett.*, 113:250601, Dec 2014. DOI: 10.1103/PhysRevLett.113.250601.

[36] Tiago B. Batalhão, Alexandre M. Souza, Laura Mazzola, Ruben Auccaise, Roberto S. Sarthour, Ivan S. Oliveira, John Goold, Gabriele De Chiara, Mauro Paternostro, and Roberto M. Serra. Experimental reconstruction of work distribution and study of fluctuation relations in a closed quantum system. *Phys. Rev. Lett.*, 113:140601, Oct 2014. DOI: 10.1103/PhysRevLett.113.140601.

[37] Shuoming An, Jing-Ning Zhang, Mark Um, Dingshun Lv, Yao Lu, Junhua Zhang, Zhang-Qi Yin, H. T. Quan, and Kihwan Kim. Experimental test of the quantum jarzynski equality with a trapped-ion system. *Nature Physics*, 11(2):193–199, Feb 2015. DOI: 10.1038/nphys3197.

[38] J. P. S. Peterson, R. S. Sarthour, A. M. Souza, I. S. Oliveira, J. Goold, K. Modi, D. O. Soares-Pinto, and L. C. Céleri. Experimental demonstration of information to energy conversion in a quantum system at the landauer limit. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 472(2188):20150813, 2016. DOI: 10.1098/rspa.2015.0813.

[39] Johannes Roßnagel, Samuel T. Dawkins, Karl N. Tolazzi, Obinna Abah, Eric Lutz, Ferdinand Schmidt-Kaler, and Kilian Singer. A single-atom heat engine. *Science*, 352(6283):325–329, 2016. DOI: 10.1126/science.aad6320.

[40] Luca Mancino, Marco Sbroscia, Ilaria Gianani, Emanuele Roccia, and Marco Barbieri. Quantum simulation of single-qubit thermometry using linear optics. *Phys. Rev. Lett.*, 118:130502, Mar 2017. DOI: 10.1103/PhysRevLett.118.130502.

[41] Zhenxing Zhang, Tenghui Wang, Liang Xiang, Zhilong Jia, Peng Duan, Weizhou Cai, Ze Zhan, Zhiwen Zong, Jianlan Wu, Luyan Sun, and et al. Experimental demonstration of work fluctuations along a shortcut to adiabaticity with a superconducting xmon qubit. *New Journal of Physics*, 20:085001, Aug 2018. DOI: 10.1088/1367-2630/aad4e7.

[42] T. P. Xiong, L. L. Yan, F. Zhou, K. Rehan, D. F. Liang, L. Chen, W. L. Yang, Z. H. Ma, M. Feng, and V. Vedral. Experimental verification of a jarzynski-related information-theoretic equality by a single trapped ion. *Phys. Rev. Lett.*, 120:010601, Jan 2018. DOI: 10.1103/PhysRevLett.120.010601.

[43] Chang-Kang Hu, Alan C. Santos, Jin-Ming Cui, Yun-Feng Huang, Diogo O. Soares-Pinto, Marcelo S. Sarandy, Chuan-Feng Li, and Guang-Can Guo. Quantum thermodynamics in adiabatic open systems and its trapped-ion experimental realization. *npj Quantum Information*, 6(1):73, Aug 2020. DOI: 10.1038/s41534-020-00300-2.

[44] P. H. Souto Ribeiro, T. Häffner, G. L. Zanin, N. Rubiano da Silva, R. Medeiros de Araújo, W. C. Soares, R. J. de Assis, L. C. Cé-

leri, and A. Forbes. Experimental study of the generalized jarzynski fluctuation relation using entangled photons. *Phys. Rev. A*, 101:052113, May 2020. DOI: 10.1103/PhysRevA.101.052113.

[45] Pieter Kok, Hwang Lee, and Jonathan P. Dowling. Creation of large-photon-number path entanglement conditioned on photodetection. *Phys. Rev. A*, 65:052104, Apr 2002. DOI: 10.1103/PhysRevA.65.052104.

[46] R. Loudon. *The Quantum Theory of Light*. OUP Oxford, 2000. ISBN 9780191589782.

[47] Masahito Ueda, Nobuyuki Imoto, and Tetsuo Ogawa. Quantum theory for continuous photodetection processes. *Phys. Rev. A*, 41: 3891–3904, Apr 1990. DOI: 10.1103/PhysRevA.41.3891.

[48] Yu. I. Bogdanov, K. G. Katamadze, G. V. Avosopiants, L. V. Belinsky, N. A. Bogdanova, A. A. Kalinkin, and S. P. Kulik. Multiphoton subtracted thermal states: Description, preparation, and reconstruction. *Phys. Rev. A*, 96:063803, Dec 2017. DOI: 10.1103/PhysRevA.96.063803.

[49] K. G. Katamadze, G. V. Avosopiants, Yu. I. Bogdanov, and S. P. Kulik. How quantum is the "quantum vampire" effect?: testing with thermal light. *Optica*, 5(6):723–726, Jun 2018. DOI: 10.1364/OPTICA.5.000723.

[50] Pieter Kok, Hwang Lee, and Jonathan P. Dowling. Creation of large-photon-number path entanglement conditioned on photodetection. *Phys. Rev. A*, 65:052104, Apr 2002. DOI: 10.1103/PhysRevA.65.052104.

[51] Itai Afek, Oron Ambar, and Yaron Silberberg. High-noon states by mixing quantum and classical light. *Science*, 328(5980):879–881, 2010. DOI: 10.1126/science.1188172.

[52] Lee A Rozema, James D Bateman, Dylan H Mahler, Ryo Okamoto, Amir Feizpour, Alex Hayat, and Aephraim M Steinberg. Scalable spatial superresolution using entangled photons. *Physical review letters*, 112(22):223602, 2014. DOI: 10.1103/PhysRevLett.112.223602.

[53] MJ Holland and K Burnett. Interferometric detection of optical phase shifts at the heisenberg limit. *Physical review letters*, 71(9):1355, 1993. DOI: 10.1103/PhysRevLett.71.1355.

[54] B. Yurke and M. Potasek. Obtainment of thermal noise from a pure quantum state. *Phys. Rev. A*, 36:3464–3466, Oct 1987. DOI: 10.1103/PhysRevA.36.3464.

[55] Bibiane Blauensteiner, Isabelle Herbauts, Stefano Bettelli, Andreas Poppe, and Hannes Hübel. Photon bunching in parametric down-conversion with continuous-wave excitation. *Phys. Rev. A*, 79: 063846, Jun 2009. DOI: 10.1103/PhysRevA.79.063846.

[56] Takahiro Sagawa and Masahito Ueda. Generalized Jarzynski Equality under Nonequilibrium Feedback Control. *Physical Review Letters*, 104(9):090602, March 2010. DOI: 10.1103/PhysRevLett.104.090602.

[57] Xiao-Qi Zhou, Timothy C. Ralph, Pruet Kalasuwan, Mian Zhang, Alberto Peruzzo, Benjamin P. Lanyon, and Jeremy L. O'Brien. Adding control to arbitrary unknown quantum operations. *Nature Communications*, 2(1):413, Aug 2011. DOI: 10.1038/ncomms1392.

[58] Guilherme L. Zanin, Michael Antesberger, Maxime J. Jacquet, Paulo H. Souto Ribeiro, Lee A. Rozema, and Philip Walther. Raw data for the manuscript 'Enhanced Photonic Maxwell's Demon with Correlated Baths', July 2021. URL https://doi.org/10.5281/zenodo.5113016.

[59] W. Martienssen and E. Spiller. Coherence and fluctuations in light beams. *American Journal of Physics*, 32(12):919–926, 1964. DOI: 10.1119/1.1970023.

[60] F. T. Arecchi. Measurement of the statistical distribution of gaussian and laser sources. *Phys. Rev. Lett.*, 15:912–916, Dec 1965. DOI: 10.1103/PhysRevLett.15.912.

[61] F.T. Arecchi, E. Gatti, and A. Sona. Time distribution of photons from coherent and gaussian sources. *Physics Letters*, 20(1):27–29, 1966. DOI: 10.1016/0031-9163(66)91034-1.

[62] R. Hanbury Brown and R. Q. Twiss. Correlation between photons in two coherent beams of light. *Nature*, 177(4497):27–29, Jan 1956. DOI: 10.1038/177027a0.

[63] R. Q. Tiss, A. G. Little, and R. Hanbury Brown. Correlation between photons, in coherent beams of light, detected by a coincidence counting technique. *Na-*

ture, 180(4581):324–326, Aug 1957. DOI: 10.1038/180324a0.

[64] Rui-Bo Jin, Ryosuke Shimizu, Kentaro Wakui, Mikio Fujiwara, Taro Yamashita, Shigehito Miki, Hirotaka Terai, Zhen Wang, and Masahide Sasaki. Pulsed sagnac polarization-entangled photon source with a ppktp crystal at telecom wavelength. *Opt. Express*, 22(10):11498–11507, May 2014. DOI: 10.1364/OE.22.011498.

A Experimental setup

In this section, we detail the arrangement of the experimental setup and provide all technical information needed to reproduce it.

A.1 Source of uncorrelated thermal light

We produce two single-mode pseudo-thermal states with the Arecchi's wheel [59–61] setup shown in Fig. 2 **(a)**. A laser beam of λ =1550 nm is divided into two spatial modes by a fibre beamsplitter (FBS). These modes are focused (lens L1, f=75mm) on two separate spots on the spinning ground glass (grit size 220). The rotation frequency of the step-motor (SM) is 20Hz. A bare fibre after the ground glass collects thermal light in each mode. We measure the second-order correlation $g^{(2)}(\tau)$ by the Brown and Twiss method [62, 63], and obtain $g^{(2)}(\tau) \approx 1.95$ (to be compared with $g^{(2)}(\tau) = 2$ for thermal light) and a coherence time $\tau \approx 5.42\ \mu s$.

A.2 Source of correlated thermal states

A 775 nm, CW beam is focused (lens L2, f=400mm) on a 30mm long ppKTP crystal in a Sagnac-loop configuration [64], yielding the singlet state $|\Psi\rangle^-$. In Fig. 2 **(b)**, a dichroic mirror (DM) reflects 775 nm light and transmits 1550 nm. We use standard mirrors (M), quarter- and half-waveplates (QWP, HWP) and polarising beamsplitters (PBS). The output light is sent to fibre couplers (FC).

A.3 Generation of anti-correlated thermal states

The Hong-Ou-Mandel (HOM) interferometer presented in Fig. 2 **(c)** takes photons from the SPDC source (Fig. 2 **(b)**) and produces anti-correlated photons by spatially and temporally overlapping the modes with a delay line (DL) and a FBS, respectively.

A.4 Operation of the photonic MD
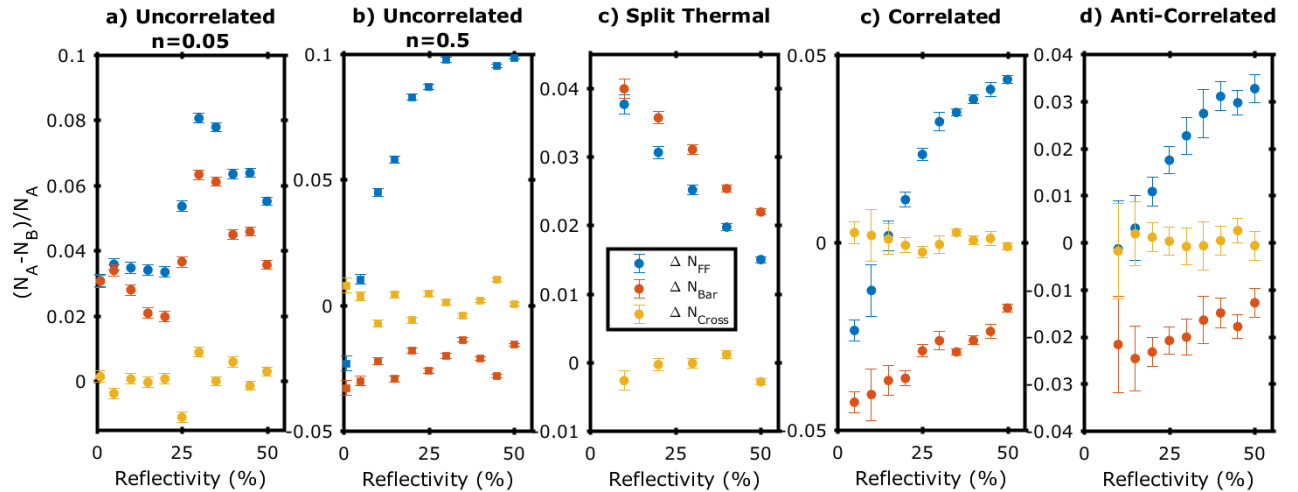
Figure 6: Individual Terms to compute the power of the MD. Here we plot the measured photon differences normalized by the input photon number when the UFOS is set to bar ΔN_{bar} (orange), cross ΔN_{cross} (red), and when the feed forward is active ΔN_{FF} (blue). Data are shown for the uncorrelated thermal state with $\overline{n} = 0.05$ (Panel a) and $\overline{n} = 0.5$ (Panel b), split thermal state (Panel c), correlated state (Panel d), and the anti-correlated state (Panel e).

The active MD is shown in Fig. 2 **(d)**. The input statistics are fed in the setup simply by connecting the fibres from either of the sources (Fig. 2 **(a)**, **(a)**+BS, **(b)**, **(b)**+**(c)**). At the TDC's, the reflected arm goes to the Demon detectors ($\text{Dem}_{A,B}$), the transmitted arm goes to a fibre delay (FD) and then

goes to the UFOS, whose configuration ('cross' or 'bar' state) is decided conditionally on the MD's information. We can adapt the switching logic according to the statistics of the MD's states. After the UFOS, photons are sent to detectors $D_{A,B}$ and the output statistics are recorded with the time tagging module (TTM) The photons' polarisation state is maintained with fibre paddles (PD) before and after the UFOS.

B   Switching logic for the feed-forward

Here we detail the switching logic used in the feed-forward control of the UFOS state. When this logic is applied, the difference (power of the MD) $\Delta N = N_A - N_B$ is improved by using the MD's information($\mathrm{Dem}_{(A,B)}$).

Dem(A,B)	UFOS state
Dem(0,0)	bar
Dem(1,1)	bar
Dem(1,0)	bar
Dem(0,1)	cross

Table 1: **Logic for uncorrelated thermal states, split thermal states and anti-correlated states.** Dem(A,B) the MD's detectors – 0 (1) represents *no click* (*click*). Bar state – input 1 (2) goes to output 1 (2); cross state – input 1 (2) goes to output 2 (1).

Dem(A,B)	UFOS state
Dem(0,0)	bar
Dem(1,1)	bar
Dem(1,0)	cross
Dem(0,1)	bar

Table 2: **Logic for correlated thermal states.** Details as in Tab. 1.

C   Derivation of the power of the MD

Here we will present our simple calculations for our photonic Maxwell's Demon in the low-photon number regime. In all cases, we define this low-photon regime to be events with two or less photons. Note that although some intermediate equations in the following derivations contain contributions from higher-order terms, all of our final results depend only on one- and two-photon detection events.

C.1   Uncorrelated Thermal States

A thermal state of light is described by:

$$\rho_{\mathrm{Th}} = \sum_{n=0}^{\infty} P_n \left| n \right\rangle \left\langle n \right|, \tag{7}$$

where

$$P(n) = \frac{\overline{n}^n}{(1 + \overline{n})^{n+1}}. \tag{8}$$

here $\overline{n}$ is the average photon number per mode. Sending this state to a beamsplitter (with vacuum incident in the second input port) that reflects the light into mode Dem_A, and transmits it into mode

A (as sketched in Fig. 1e, where mode A refers to the mode before and after the beamsplitter), results in the following probabilities to detect n photons in mode Dem_A and m photons in mode A [46]:

$$P(m,n) = \frac{(n+m)!}{n!m!} \frac{\overline{n}^{n+m}}{(1+\overline{n})^{n+m+1}} R^{2n}(1-R^2)^m, \tag{9}$$

where R is the reflection *amplitude* of the beamsplitter. Then after the beamsplitter, we will approximate the state to be:

$$\sigma_{\text{Dem}_A,A} = P(0,0) \left|0,0\right\rangle \left\langle 0,0\right|_{\text{Dem}_A,A} + P(1,0) \left|0,1\right\rangle \left\langle 0,1\right|_{\text{Dem}_A,A} + \tag{10}$$
$$P(0,1) \left|1,0\right\rangle \left\langle 1,0\right|_{\text{Dem}_A,A} + P(1,1) \left|1,1\right\rangle \left\langle 1,1\right|_{\text{Dem}_A,A} + \dots$$

In this expression we have ignored all off-diagonal terms, and terms with more than two photons. Furthermore, we have left out the terms $P(2,0) \left|0,2\right\rangle \left\langle 0,2\right|_{\text{Dem}_A,A}$ and $P(0,2) \left|2,0\right\rangle \left\langle 2,0\right|_{\text{Dem}_A,A}$ since $P(2,0) << P(1,0)$ for low photon numbers. This condition is valid when the probability of populating terms with photon numbers of three or higher is negligible compared to the one and two-photon contributions. In practice this holds for $\overline{n} \lesssim 0.2$. Then for our experiment we have one such state in modes A and Dem_A, and one in modes B and Dem_B. Thus the total input state is $\sigma_{\text{Dem}_A,A} \otimes \sigma_{\text{Dem}_B,B}$, with $\sigma_{i,j}$ defined above. The overall expression is straight-forward but rather lengthy, so we do not include it here.

The next step is to include the action of the demon. For uncorrelated thermal states, when the demon detects a photon in mode Dem_A (Dem_B), the photon number in mode A, (B) is increased. Since the demon's goal is to increase the temperature (photon number) of mode A, this means that when it detects a photon in Dem_A it does nothing. On the other hand, when it detects a photon in mode Dem_B there are more photons in mode B, so it should swap modes A and B. If the demon detects a photon in both mode Dem_A and mode Dem_B it gains no information; hence, it does nothing. To represent this mathematically we simply make the following substitutions:

$$\left|n,m,0,1\right\rangle_{A,B,\text{Dem}_A,\text{Dem}_B} \to \left|m,n,0,1\right\rangle_{D_A,D_B,\text{Dem}_A,\text{Dem}_B} \tag{11}$$

Since we do not use number resolving detectors, we do not swap modes A and B in situations such as $\left|n,m,1,2\right\rangle$, as this will register as a detection event at both Dem_A and Dem_B. With this operation, the low-photon number state after the demon is approximated as

$$\begin{aligned}
\rho_{\text{out}} = \quad & P^2(0,0) \left|0000\right\rangle \left\langle 0000\right| + P(0,0)P(0,1) \left|0001\right\rangle \left\langle 0001\right| + P(0,0)P(1,0) \left|0100\right\rangle \left\langle 0100\right| + \\
& P(0,0)P(1,1) \left|1001\right\rangle \left\langle 1001\right| + P(0,1)P(0,0) \left|0010\right\rangle \left\langle 0010\right| + P^2(0,1) \left|0011\right\rangle \left\langle 0011\right| + \\
& P(0,1)P(1,0) \left|0110\right\rangle \left\langle 0110\right| + P(0,1)P(1,1) \left|0111\right\rangle \left\langle 0111\right| + \\
& P(1,0)P(0,0) \left|1000\right\rangle \left\langle 1000\right| + P(1,0)P(0,1) \left|0101\right\rangle \left\langle 0101\right| + P^2(1,0) \left|1100\right\rangle \left\langle 1100\right| + \\
& P(1,0)P(1,1) \left|1101\right\rangle \left\langle 1101\right| + P(1,1)P(0,0) \left|1010\right\rangle \left\langle 1010\right| + P(1,1)P(0,1) \left|1011\right\rangle \left\langle 1011\right| + \\
& P(1,1)P(1,0) \left|1110\right\rangle \left\langle 1110\right| + P^2(1,1) \left|1111\right\rangle \left\langle 1111\right| + \dots
\end{aligned} \tag{12}$$

In Eq. 12, the mode labels are left off, but they run D_A, D_B, Dem_A, Dem_B; i.e. $\left|ijkl\right\rangle$ is labelled as $\left|ijkl\right\rangle_{D_A,D_B,\text{Dem}_A,\text{Dem}_B}$. In our experiments, we always measure the singles rates in modes D_A and D_B to observe the power of the demon. From Eq. 12, we can extract the probability for a detector in mode D_A or mode D_B to click.

$$P_A = 2P(0,0)P(1,1) + 2P(1,0)P(1,1) + P(0,0)P(1,0) + P(0,1)P(1,1) + P^2(1,0) + P^2(1,1) \tag{13}$$

$$P_B = P(0,0)P(1,0) + 2P(1,0)P(0,1) + 2P(1,0)P(1,1) + P(0,1)P(1,1) + P^2(1,0) + P^2(1,1). \tag{14}$$

Now experimentally, we run this setup many times, observing average singles rates at both detectors. In general, these rates will be given by ΓP_A and ΓP_B, where Γ is the total number of repetitions of the experiment. Thus the measured photon number difference is given by $\Delta N = \Gamma(P_A - P_B)$. Finally, using our expressions for P_A, P_B we arrive at:

$$\Delta N = (2P(0,0)P(1,1) - 2P(0,1)P(1,0))\Gamma$$
$$= \frac{2\overline{n}^2}{(1-\overline{n})^4}R^2(1-R^2)\Gamma, \tag{15}$$

where we have substituted $P(n,m)$ from Eq. 9 to arrive at the last line. As a last step, we construct our figure of merit for the power of the MD $\Delta N/N$. Here N is the number of input photons in mode $\mathrm{In_A}$, which we constrain to be equal to the number of input photons in mode $\mathrm{In_B}$. This figure or merit can be interpreted as the demon power per photon. In the low-photon limit, the probability for one photon to be in mode $\mathrm{In_A}$ is given by $P(1) = \overline{n}/(1+\overline{n})$. Then the number of single-photon events detected per second is $N = \Gamma P(1)$, and thus

$$\frac{\Delta N}{N} = \frac{2\overline{n}}{(1-\overline{n})^2}R^2(1-R^2). \tag{16}$$

Thus we see that, even in the low photon number limit, the demon effect depends on the effective temperature of the thermal state via $\overline{n}$. As we will see in the next sections, this is not the case for correlated and anti-correlated states in this regime.

C.2   Split Thermal States

In this case, the input to the experiment is a single thermal state, which has been split at a beamsplitter. Thus, a thermal state is still input into both mode A and B; however, correlations will now exist between the two modes. Roughly speaking, now when the demon detects a photon at (say) detector Dem_A, it can no longer know if the bunching effect will be stronger in mode A or B.

In more detail, the state in modes $\mathrm{In_A}$ and $\mathrm{In_B}$ is now:

$$\rho_{\text{split}} = \sum_{n,m=0}^{\infty} P(n,m) |n,m\rangle \langle n,m|_{\mathrm{In_A,In_B}}, \tag{17}$$

where $P(n,m)$ is defined above, with $R = 1/\sqrt{2}$. As above, we will keep only up to the two-photon terms and consider only the diagonal elements:

$$\rho_{\text{split}} \approx P(0,0) |0,0\rangle \langle 0,0|_{\mathrm{In_A,In_B}} + P(0,1) |0,1\rangle \langle 0,1|_{\mathrm{In_A,In_B}} + P(1,0) |1,0\rangle \langle 1,0|_{\mathrm{In_A,In_B}} +$$
$$P(2,0) |2,0\rangle \langle 2,0|_{\mathrm{In_A,In_B}} + P(0,2) |0,2\rangle \langle 0,2|_{\mathrm{In_A,In_B}} + P(1,1) |1,1\rangle \langle 1,1|_{\mathrm{In_A,In_B}} + \dots \tag{18}$$

Next, we will add in the two beamsplitters. The beamsplitters enact the following transformations:

$$|0,1\rangle \langle 0,1|_{\mathrm{In_A,In_B}} \rightarrow (1-R^2) |0100\rangle \langle 0100|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B} + R^2 |0001\rangle \langle 0001|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B}$$

$$|1,0\rangle \langle 1,0|_{\mathrm{In_A,In_B}} \rightarrow (1-R^2) |1000\rangle \langle 1000|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B} + R^2 |0010\rangle \langle 0010|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B}$$

$$|1,1\rangle \langle 1,1|_{\mathrm{In_A,In_B}} \rightarrow (1-R^2)^2 |1100\rangle \langle 1100|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B} + R^2(1-R^2) |1001\rangle \langle 1001|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B}$$
$$+ R^2(1-R^2) |0110\rangle \langle 0110|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B} + R^4 |0011\rangle \langle 0011|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B}$$

$$|0,2\rangle \langle 0,2|_{\mathrm{In_A,In_B}} \rightarrow (1-R^2)^2 |0200\rangle \langle 0200|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B} + 2R^2(1-R^2) |0101\rangle \langle 0101|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B}$$
$$+ R^4 |0002\rangle \langle 0002|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B}$$

$$|2,0\rangle \langle 2,0|_{\mathrm{In_A,In_B}} \rightarrow (1-R^2)^2 |2000\rangle \langle 2000|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B} + 2R^2(1-R^2) |1010\rangle \langle 1010|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B}$$
$$+ R^4 |0020\rangle \langle 0020|_{A,B,\mathrm{Dem}_A,\mathrm{Dem}_B}. \tag{19}$$

Again, here we have not included the off-diagonal terms. This will not affect our final result since we never recombine two modes, i.e. we never interfere modes after splitting them. The demon will now

attempt to increase the photon number at the detector in mode D_A by swapping modes A and B on detecting a photon at detector Dem_B, as in Eq. 11. However, it was pointed out in Ref. [31] that the result is the same if the demon instead swaps the modes upon detection of a photon at detector Dem_A. Substituting in these transformations, and using the fact that $P(n,m) = P(m,n)$ when $R = 1/\sqrt{2}$ we arrive at the following expression for the output state:

$$
\begin{aligned}
\rho_{\text{out}} = \quad & P(0,0)\,|0000\rangle\langle0000| + P(0,1)\left[(1-R^2)(|0100\rangle\langle0100| + |1000\rangle\langle1000|) + \right. \\
& R^2(|0001\rangle\langle0001| + |0010\rangle\langle0010|)\right] + P(1,1)\left[(1-R^2)^2\,|1100\rangle\langle1100| + \right. \\
& R^2(1-R^2)(|0101\rangle\langle0101| + |0110\rangle\langle0110|) + R^4\,|0011\rangle\langle0011|\right] + \\
& P(0,2)\left[(1-R^2)^2(|2000\rangle\langle2000| + |0200\rangle\langle0200|) + R^4(|0020\rangle\langle0020| + |0002\rangle\langle0002|) \right. \\
& \left. + 2R^2(1-R^2)(|1010\rangle\langle1010| + |1001\rangle\langle1001|)\right],
\end{aligned}
\tag{20}
$$

where the mode labels are D_A, D_B, Dem_A, Dem_B.

From here, we compute the probability for detector D_A and D_B to click:

$$
P_A = P(0,1)(1-R^2) + P(1,1)(1-R^2)^2 + P(0,2)(1-R^2)^2 + 4P(0,2)R^2(1-R^2) \tag{21}
$$
$$
P_B = P(0,1)(1-R^2) + P(1,1)(1-R^2)^2 + P(0,2)(1-R^2)^2 + 2P(1,1)R^2(1-R^2). \tag{22}
$$

Then, we calculate the photon number difference $\Delta N = \Gamma(P_A - P_B)$.

$$
P_A - P_B = 2R^2(1-R^2)(2P(0,2) - P(1,1)) = 0, \tag{23}
$$

where 0 is obtained after substituting in the explicit expressions for $P(0,2)$ and $P(1,1)$ with $R = 1/\sqrt{2}$. Hence,

$$
\frac{\Delta N}{N} = 0, \tag{24}
$$

for all values of the reflectivity of the demon's beamsplitter. Moreover, in Ref. [31] it was shown that this result holds in general, not just in the low-photon number regime.

C.3 Correlated Thermal States

To generate our correlated thermal states, we use a type-II SPDC source. In such a source, photons are always emitted in pairs. However, when the idler (signal) beam is traced out, the idler (signal) beam is left in a thermal state [46]. The general input state is given by

$$
|\psi\rangle = \frac{1}{\cosh s}\sum_{n=0}^{\infty}(\tanh s)^n\,|n,n\rangle_{\text{In}_A,\text{In}_B}, \tag{25}
$$

but for our purposes we will only consider up to the two photon term

$$
|\psi\rangle \approx \frac{1}{1-\frac{s^2}{2}}\,|0,0\rangle_{\text{In}_A,\text{In}_B} + \frac{s}{1-\frac{s^2}{2}}\,|1,1\rangle_{\text{In}_A,\text{In}_B} + \dots. \tag{26}
$$

Here, s is the so-called squeezing parameter related to the brightness of the SPDC source, and thus to the temperature of the thermal states in each mode: $\sinh^2(s) = \overline{n}$. Since the SPDC source is pumped with a continuous-wave laser, the terms of order higher than 2 ($|2,2\rangle + |3,3\rangle\dots$) in Eqs. (25) and (26) have low probability and can be neglected. I.e. we are in the 'low photon-number' regime.

For the uncorrelated thermal states, we did not need to consider loss, since a thermal state remains a thermal state upon losing a photon. However, here, the correlations are essential so loss will play a large roll in the power of the demon (the information need to extract work). We model loss as a beamsplitter, which takes $|1\rangle_A \to \epsilon\,|1,0\rangle_{\text{In}_A,l_A} + \sqrt{1-\epsilon^2}\,|0,1\rangle_{\text{In}_A,l_A}$. In words, the photon remains in the mode with probability ϵ^2.

Assuming the same loss in each mode, we have

$$|\psi\rangle \approx \frac{1}{1-\frac{s^2}{2}}|0,0,0,0\rangle_{\mathrm{In_A,In_B},l_A,l_B} + \frac{s\epsilon^2}{1-\frac{s^2}{2}}|1,1,0,0\rangle_{\mathrm{In_A,In_B},l_A,l_B}$$

$$+ \frac{s(1-\epsilon^2)}{1-\frac{1}{s^2}}|0011\rangle_{\mathrm{In_A,In_B},l_A,l_B} \frac{s\epsilon\sqrt{1-\epsilon^2}}{1-\frac{s^2}{2}}\left(|1,0,0,1\rangle_{\mathrm{In_A,In_B},l_A,l_B} + |0,1,1,0\rangle_{\mathrm{In_A,In_B},l_A,l_B}\right). \quad (27)$$

Next, we will add in the action of the demon's beamsplitters, which will implement the following operation on a photon in mode A: $|1,0\rangle_{\mathrm{In_A,E_A}} \to \sqrt{1-R^2}|1,0\rangle_{\mathrm{A,Dem_A}} + R|0,1\rangle_{\mathrm{A,Dem_A}}$, and the same on mode B, where E_i labels the empty input mode the beamsplitters.

$$|\psi\rangle \approx \left(\frac{1}{1-\frac{s^2}{2}}|0,0,0,0,0,0\rangle + \frac{s(1-\epsilon^2)}{1-\frac{1}{s^2}}|0,0,0,0,1,1\rangle \frac{s\epsilon\sqrt{1-\epsilon^2}}{1-\frac{s^2}{2}}R\left(|0,0,1,0,0,1\rangle + |0,0,0,1,1,0\rangle\right)\right.$$

$$+\frac{s\epsilon\sqrt{1-\epsilon^2}}{1-\frac{s^2}{2}}\sqrt{1-R^2}\left(|1,0,0,0,0,1\rangle + |0,1,0,0,1,0\rangle\right) +$$

$$\frac{s\epsilon^2}{1-\frac{s^2}{2}}R\sqrt{1-R^2}\left(|0,1,1,0,0,0\rangle + |1,0,0,1,0,0\rangle\right) +$$

$$\left.\frac{s\epsilon^2}{1-\frac{s^2}{2}}\left((1-R^2)|1,1,0,0,0,0\rangle + R^2|0,0,1,1,0,0\rangle\right)\right), \quad (28)$$

where we have dropped the mode labels but they are $|i,j,k,l,m,n\rangle_{\mathrm{A,B,Dem_A,Dem_A},l_A,l_B}$. Finally, the demon uses its optical switch when it measures a photon at detector Dem_A, taking $|i,j,1,0,0,0\rangle \to |j,i,1,0,0,0\rangle$. In the above equation, this only changes one term, $|0,1,1,0,0,0\rangle \to |1,0,1,0,0,0\rangle$. This results in the following state:

$$|\psi\rangle \approx \left(\frac{1}{1-\frac{s^2}{2}}|0,0,0,0,0,0\rangle + \frac{s(1-\epsilon^2)}{1-\frac{1}{s^2}}|0,0,0,0,1,1\rangle + \frac{s\epsilon\sqrt{1-\epsilon^2}}{1-\frac{s^2}{2}}R\left(|0,0,1,0,0,1\rangle + |0,0,0,1,1,0\rangle\right)\right.$$

$$+\frac{s\epsilon\sqrt{1-\epsilon^2}}{1-\frac{s^2}{2}}\sqrt{1-R^2}\left(|1,0,0,0,0,1\rangle + |0,1,0,0,1,0\rangle\right) +$$

$$\frac{s\epsilon^2}{1-\frac{s^2}{2}}R\sqrt{1-R^2}\left(|1,0,1,0,0,0\rangle + |1,0,0,1,0,0\rangle\right) +$$

$$\left.\frac{s\epsilon^2}{1-\frac{s^2}{2}}\left((1-R^2)|1,1,0,0,0,0\rangle + R^2|0,0,1,1,0,0\rangle\right)\right), \quad (29)$$

where the mode labels are now $|i,j,k,l,m,n\rangle_{\mathrm{D_A,D_B,Dem_A,Dem_A},l_A,l_B}$. Finally, from this we can compute the probability for detectors D_A and D_B to click:

$$P_A = \frac{s^2}{\left(1-\frac{s^2}{2}\right)^2}\left(\epsilon^2(1-\epsilon^2)(1-R^2) + \epsilon^4(1-R^2)^2 + 2\epsilon^4 R^2(1-R^2)\right) \quad (30)$$

$$P_B = \frac{s^2}{\left(1-\frac{s^2}{2}\right)^2}\left(\epsilon^2(1-\epsilon^2)(1-R^2) + \epsilon^4(1-R^2)^2\right). \quad (31)$$

Then, as before, our experimentally measured ΔN is given by $\Gamma(P_A - P_B)$, and thus

$$\Delta N = \frac{s^2}{\left(1-\frac{s^2}{2}\right)^2}2\epsilon^4 R^2(1-R^2)\Gamma. \quad (32)$$

Again, our figure of merit is the demon effect per photon. However, we now have two ways of computing this from our experimental data. First of all, we can proceed as we did with the uncorrelated

thermal light and normalise by the input singles rate. In this case, the probability for a photon to be incident in mode In_A (or in mode In_B) is $s^2/(1-s^2/2)^2\epsilon^2$, coming directly from Eq. 27. Thus, the incoming photon rate is $N = \Gamma s^2/(1-s^2/2)^2\epsilon^2$, and our figure of merit when normalising by singles is

$$\frac{\Delta N}{N} = 2\epsilon^2 R^2(1-R^2).\tag{33}$$

After a closer look at Eq. 27, one can see that ϵ^2 can be understood as the coupling efficiency. In other words, it is the ratio of the measured coincidence rate to the singles rate before the demon acts. We thus see that in this case the demon effect is scaled down by the coupling efficiency, which makes sense since a lower coupling efficiency degrades the correlations, making the demon less effective. The other point to note is that, in this regime, the demon is no longer driven by the thermal statistics of the light, but the correlations from the SPDC dominate. This is even more apparent when we instead compute the demon effect per photon pair.

Examining Eq. 27 again, we see that the probability for a correlated pair to be incident in mode In_A and In_B is given by $P_{AB} = s^2/(1-s^2/2)^2\epsilon^4$, and thus the coincidence rate is $C = \Gamma P_{AB}$, so

$$\frac{\Delta N}{C} = 2R^2(1-R^2).\tag{34}$$

Now the coupling efficiency drops of altogether, and we are left with the ideal demon effect for correlated thermal states. We also see that when the demon uses a 50:50 beamsplitter ($R = 1/\sqrt{2}$), then $\Delta N/C = 0.5$. This means that on average, the demon is able to create a photon number imbalance of 0.5 per photon pair.

C.4 Anti-Correlated states

In Ref. [31], a specific two-mode state was proposed with two important properties. First, for every N-photon event, the two modes are in a N00N state. Second, tracing out either mode, leaves the remaining mode in a thermal state. Here we do not create this state exactly, but we do note that some experimentally-realised N00N state sources may come close to these conditions [51, 52], as they do result in a superposition of N00N states for various photon numbers. Instead, here we perform Hong-Ou-Mandel interference between the two photons from our our SPDC source to create a 2-photon N00N state. The higher-order modes will not be in a N00N state; rather, they will be in so-called Holland-Burnett states [53]. The main qualitative difference between these Holland-Burnett states and the anti-correlated state of Ref. [31] is that the Holland-Burnett states contain only terms with even photon numbers, while those of Ref. [31] contain all photon numbers. Since experimentally we are in the low-photon number regime (where the probability to detect two photons is much larger than the probability to detect four photons) the main difference between our experimentally-generated state and that of Ref. [31] is the absence of the one-photon component. It is straight-forward to see that this term has no effect on the MD's power. The one-photon term is: $\frac{1}{\sqrt{2}}|1,0\rangle_{\text{In}_A,\text{In}_B} + \frac{1}{\sqrt{2}}|0,1\rangle_{\text{In}_A,\text{In}_B}$. After application of the demon's beamsplitters, the state is:

$$\frac{1}{\sqrt{2}}\Big[\sqrt{1-R^2}\,|1,0,0,0\rangle_{A,B,\text{Dem}_A,\text{Dem}_B} + R\,|0,0,1,0\rangle_{A,B,\text{Dem}_A,\text{Dem}_B}$$
$$+ \sqrt{1-R^2}\,|0,1,0,0\rangle_{A,B,\text{Dem}_A,\text{Dem}_B}\, R\,|0,0,0,1\rangle_{A,B,\text{Dem}_A,\text{Dem}_B}\Big].\tag{35}$$

Clearly, this state does not change regardless of the demon's action with the UFOS upon detection at Dem_A or Dem_B. Thus, the probability to detect a photon at D_A or D_B is equal; i.e. $P_A = P_B = 1-R^2/2$. Since ΔN is proportional to $P_A - P_B$, $\Delta N = 0$ for this state. Therefore, the absence of the one-photon component will not affect our measurement of ΔN.

To model our experimentally produced state, we begin with

$$|\psi\rangle = \frac{1}{1-\frac{s^2}{2}}|0,0\rangle_{\text{In}_A,\text{In}_B} + \frac{s}{1-\frac{s^2}{2}}\left(\frac{v}{\sqrt{2}}(|2,0\rangle_{\text{In}_A,\text{In}_B} + |0,2\rangle_{\text{In}_A,\text{In}_B}) + \sqrt{1-v^2}|1,1\rangle_{\text{In}_A,\text{In}_B}\right).\tag{36}$$

Here, again, s is the squeezing parameter. We have also modeled the imperfect HOM visibility by the parameter v, whereby the photons bunch into a N00N state with probability v^2.

We then proceed as before, modeling loss as $|1\rangle_{\text{In}_A} \to \epsilon |1,0\rangle_{\text{In}_A,\text{l}_A} + \sqrt{1-\epsilon^2} |0,1\rangle_{\text{In}_A,\text{l}_A}$, which acts as $|2\rangle_{\text{In}_A} \to \epsilon^2 |2,0\rangle_{\text{In}_A,\text{l}_A} + (1-\epsilon^2) |0,2\rangle_{\text{In}_A,\text{l}_A} + \sqrt{2}\epsilon\sqrt{1-\epsilon^2} |1,1\rangle_{\text{In}_A,\text{l}_A}$ on the two-photon terms. With the same definitions for mode B. Similarly, the demon's beamsplitters enact the following transformations: $|1,0\rangle_{\text{In}_A,\text{E}_A} \to \sqrt{1-R^2} |1,0\rangle_{A,\text{Dem}_A} + R |0,1\rangle_{A,\text{Dem}_A}$, and $|2,0\rangle_{\text{In}_A,\text{E}_A} \to (1-R^2)|2,0\rangle_{A,\text{Dem}_A} + R^2 |0,2\rangle_{A,\text{Dem}_A} + \sqrt{2}R\sqrt{1-R^2}|1,1\rangle_{A,\text{Dem}_A}$. In this case, the demon's strategy will be the same as for uncorrelated thermal states. This is because when the demon detects a photon in mode Dem_B, it is now more likely that there are more photons in mode B (unless both photons of the N00N state are reflected). Thus the demon will swap the modes. On the other hand, the demon will not do anything when a photon is detected in mode Dem_A. Working through all of these transformations, one arrives at the following state:

$$
\begin{aligned}
|\psi\rangle = {} & \frac{1}{1-\frac{s^2}{2}} |0,0,0,0,0,0\rangle + \frac{s}{1-\frac{s^2}{2}} \sqrt{1-v^2}\Big[\epsilon^2(1-R^2)|1,1,0,0,0,0\rangle + \epsilon^2 R^2 |0,0,1,1,0,0\rangle \\
& + \epsilon^2 R\sqrt{1-R^2}(|0,1,1,0,0,0\rangle + |0,1,0,1,0,0\rangle) \\
& + \epsilon\sqrt{1-\epsilon^2}\sqrt{1-R^2}(|1,0,0,0,0,1\rangle + |0,1,0,0,1,0\rangle) \\
& + \epsilon\sqrt{1-\epsilon^2}R(|0,0,1,0,0,1\rangle + |0,0,0,1,1,0\rangle) + (1-\epsilon^2)|0,0,0,0,1,1\rangle\Big] \\
& + \frac{s}{1-\frac{s^2}{2}}\frac{v}{\sqrt{2}}\Big[\epsilon^2(1-R^2)(|2,0,0,0,0,0\rangle + |0,2,0,0,0,0\rangle) + \epsilon^2 R^2(|0,0,2,0,0,0\rangle + |0,0,0,2,0,0\rangle) \\
& + \sqrt{2}\epsilon^2 R\sqrt{1-R^2}(|1,0,1,0,0,0\rangle + |1,0,0,1,0,0\rangle) \\
& + \sqrt{2}\epsilon\sqrt{1-\epsilon^2}\sqrt{1-R^2}(|1,0,0,0,1,0\rangle + |0,1,0,0,0,1\rangle) \\
& + \sqrt{2}\epsilon\sqrt{1-\epsilon^2}R(|0,0,1,0,1,0\rangle + |0,0,0,1,0,1\rangle) + (1-\epsilon^2)(|0,0,0,0,2,0\rangle + |0,0,0,0,0,2\rangle)\Big].
\end{aligned}
\tag{37}
$$

The mode labels are $|i,j,k,l,m,n\rangle_{\text{D}_A,\text{D}_B,\text{Dem}_A,\text{Dem}_B,\text{l}_A,\text{l}_B}$. It is then straightforward to compute the probability for the detectors in mode A and B to click then. For that we can obtain the photon number difference as before

$$
\Delta N = \frac{s^2}{\left(1-\frac{s^2}{2}\right)^2} 2(2v^2-1)\epsilon^4 R^2(1-R^2)\Gamma.
\tag{38}
$$

This expression is very similar to that of the SPDC light, simply with an additional factor of $(2v^2-1)$. So we see in the low photon-number regime, for an ideal source both the correlated and anti-correlated states behave identically. But in practice, the correlated state should perform better, since they do not require the additional experimental step of Hong-Ou-Mandel

As a last step, we again compute our two figures of merit, by normalising ΔN by both the singles and the coincidence counts, yielding

$$
\frac{\Delta N}{N} = 2\epsilon^2(2v^2-1)R^2(1-R^2)
\tag{39}
$$

and

$$
\frac{\Delta N}{C} = 2(2v^2-1)R^2(1-R^2).
\tag{40}
$$

Similar to the correlated states, we see that normalising by coincidences removes the effect of the imperfect coupling efficiency. However, the visibility factor remains.

D The Maxwell's Demon's Information

In order to quantify the information accessible to the MD we compute the mutual information between the MD's measurement outcomes (detector clicks at detectors Dem_A and Dem_B) and the photon numbers distribution in mode A and B after MD's two beamsplitters. As discussed in the main text, this quantity imposes an upper limit on the amount of work that the MD can extract. The mutual information between the detector outcomes and the photon number quantifies the common information between these systems, and, hence, how much the MD can learn about the photon number from its measurements.

Following Ref. [32], the mutual information between the demons measurement clicks (M_a and M_b) and the photon number in mode A and B after the beamsplitter (N_a and N_b) in our setup can be written as

$$I(M_a, M_b : N_a, N_b) = \sum_{m_A}\sum_{m_B}\sum_{n_A}\sum_{n_B} p(m_A, m_B, n_A, n_B)\left[log(p(m_A, m_B | n_A, n_B)) - log(p(m_A, m_B))\right].$$
(41)

In Eq. (D), $p(m_A, m_B, n_A, n_B)$ is the probability to detect m_A photons at detector Dem_A, m_B photons at detector Dem_B, while having n_A photons remaining in mode A and n_B photons remaining in mode B. Similarly, $p(m_A, m_B)$ is the probability for the demon to detect m_A and m_B photons at detectors Dem_A and Dem_B, respectively. We obtain $p(m_A, m_B)$ by marginalizing $p(m_A, m_B, n_A, n_B)$ over n_A and n_B. The joint probability $p(m_A, m_B, n_A, n_B)$ will depend on the input state, and can be taken almost directly from the output states computed in Appendix C. In particular, they are obtained from the coefficients of the various terms in Eq. (12) for the uncorrelated case, Eq. (20) for the split thermal state, Eq. (29) for the correlated state, and Eq. (37) for the anti-correlated state. However, the MD's operation given in Eq. (11) must first be undone. Furthermore, for the correlated and anti-correlated states, we ignore the vacuum component from the state preparation (i.e. we set $|\psi\rangle = |1,1\rangle_{A,B}$ in Eq. (26), and similarly in Eq. (36)) to reflect our normalization by the singles rate.

The remaining quantity in Eq. (D) is the conditional probability $p(m_A, m_B | n_A, n_B)$ which is the probability that the MD will detect m_A and m_B photons at detectors Dem_A and Dem_B, respectively, *given* that n_A and n_B photon are incident in modes In_A and In_B, respectively. Since the modes do not interact before the MD's beamsplitter this probability factorizes $p(m_A, m_B | n_A, n_B) = p(m_A | n_A) p(m_B | n_B)$. These individual probabilities for mode A and B can be readily computed by assuming a Fock state $|n\rangle$ is input to a beamsplitter of reflection amplitude R

$$p(m|n) = \sum_{k=0}^{n} \frac{n! \; \epsilon^{2k}(1-\epsilon^2)^{n-k} R^{2m}(1-R^2)^{k-m}}{m!(n-k)!(k-m)!},$$
(42)

where the coupling efficiency is ϵ^2 (as defined above).

Notice that the conditional probabilities are the same for every input state, as they are related just to the MD's apparatus. The effect of the input states on the mutual information arises through the joint probabilities coming from the calculations presented in Appendix C, and these introduce dependence the average photon number (for the uncorrelated and split thermal states) and the visibility (for the anti-correlated state).