

# The extraction of work from hidden degrees of freedom

Corresponding Author: Professor Clemens Bechinger

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

This article presents an interesting theoretical and experimental study of protocols which can produce work using the memory of non Markovian dynamics. The idea is original and the results, which may have important applications, could be published in Nature Physics after that the authors will take into account the following questions. Although the article is in average well written there are several sections that are rather hard to understand and they need to be clarified. Specifically:

The main idea is that in order to produce work from memory a single measurement is not enough and a delayed measure is needed. This second delayed measure is actually the key step because it gives information about the dynamical correlations that can actually be used to produce work. Figures 1 and 2 and associated equations explain rather well the two measurements methods.

The analysis in terms of information, summarized in Fig.3 is not very clear.

1) Is  $\tilde{I}$  used only to check the difference between a random sampling and a delayed sampling ? Indeed from eq.2 and eq.5 I need only  $I$  to bound the work. What is the point of introducing  $\tilde{I}$ ?

2) The description of the method used to estimate  $\tilde{I}$  is very short and Supplementary Section 3 does not help. In the Figure in SupMat the  $\delta t$  is not indicated and in the phrase "the remainder of the  $|1\rangle$  ensemble,  $|01\rangle$ ;  $|01\rangle$  should not be  $|10\rangle$ ". Probably there is a problem of wording. But I am quite confused of figure b. If I understand correctly : you select all the points in which the first reading gave 1. Starting from these points you sample again either after a fixed  $\delta t$  or a randomly chosen  $\delta t$ . From the results of the second measure you compute  $I$  and  $\tilde{I}$  respectively. Is this correct ? If the answer is yes the Figure in SupMat and text are rather misleading especially Figure b.

3) The relationship between Figure 3 and the rest of the article is very confused. A paragraph which makes a clear connection between the results of Fig.3 and those of Fig.4 is needed. Furthermore, as it is explained in methods, I agree that in Fig.4 only  $\bar{V}(t)$  is needed to estimate the work produced. However it would be useful to have a plot with both  $W(t)$  and  $I(t)$  to check the quality of the bound.

4) The reason of the work difference between the single and double well is not clearly explained. It took a lot of time to understand this after finding the phrase "To extract additional energy stored in hidden modes, one must transiently impede particle relaxation while the bath remains out of equilibrium.". This phrase must be emphasized because the fact of producing work is not just a matter of memory effect but also of the chosen protocol.

5) I do not understand the phrase "To increase the number of detectable events within the limited experimental data set, we allowed  $\delta t$  to take values within the interval  $[0, 200 \text{ s}]$  and aggregated the corresponding events to improve statistical significance.". The protocol last 1.2s and the delay 200s. Does it mean that you drive the system in a NESS ?

6) minor problem. it is never written that the acronym SEM stands for Standard Error of the Mean. It is not necessarily known.

7) Finally concerning non-Markov demon there are recent articles on harmonic oscillators which will be worth to quote: Archambault et al. PRL 135, 147101 (2025); Archambault et al, EPL 148 41002 (2024).

(Remarks on code availability)

## Reviewer #2

### (Remarks to the Author)

This manuscript shows how to improve work extraction from a stochastic system by exploiting its non-Markovian character, using a two-time measurement protocol. I find this to be an exciting result, highlighting that memory effects are not just a nuisance complicating the equations of motion, but a source of information about hidden bath degrees of freedom. The paper is very clear and well written, and I recommend publication. I have two questions for the authors to consider:

1. The authors consider a specific artificial measurement protocol, in which the system is mapped onto two discrete states, below threshold  $|0\rangle$  and above threshold  $|1\rangle$ . One could increase the number of states or instead report on the actual values of the coordinate at the times of the measurements. Likewise, one could consider 3-time measurements. Would such more elaborate measurement protocols further improve work extraction? What are the limits on how much can be gained?
2. Figure 2 reports on conditional relaxation curves obtained from experimental measurements. It would be helpful if one could get some theoretical insight into those. For example, since the bath has a single relaxation timescale, could the dynamics  $x(t)$  be modeled as a generalized Langevin equation with an exponential memory kernel (or, equivalently, could the hidden degrees of freedom be described by a single degree of freedom coupled harmonically to  $x$ )? If so, it may be possible to have a simple theory for the data appearing in Figs. 2-3

### (Remarks on code availability)

## Reviewer #3

### (Remarks to the Author)

The present manuscript clearly presents, by experiments and by theory, how to extract thermodynamic work from degrees of freedom that are not directly observable. Besides some advances on the theoretical framework, it truly provides (for the first time) a full experimental confirmation of such extraction of work.

Non-Markovian effects, in particular hidden/unobserved degrees of freedom and their role in thermodynamics, are gaining more and more traction and are attracting increasingly more attention in various fields (simply judging based on the most recent perspective articles and back-to-back special issues in not one but in fact several journals). The results in the present manuscript are likely to significantly advance not only the field of stochastic thermodynamics, but also statistical physics, biophysics, and even beyond. The results are original and convincing up to some points that I would ask the authors to address in the revision (see below).

Information stored in hidden degrees (including non-Markovian effects) is a challenging but very important and timely field of research, which (given the intense efforts in the field) is expected to soon transgress from conceptual challenges to technological applications. Since the results in the manuscript instructively and clearly present a useful case for addressing this hidden information, they concurrently make a strong case for further studies in similar and related directions. Overall, I recommend accepting the paper in Nature Physics under the condition that the following (critical) remarks are satisfactorily answered.

(1) There seems to be one issue with the energetic considerations. Below Eq. (7), it is mentioned that this form is due to "rigid translations not giving rise to changes in equilibrium free energy and Shannon entropy". While the first point is correct as shown in Eq. (16), above Eq. (18) it is only mentioned (not shown) that the Shannon entropy is "invariant under rigid potential transformations". In the given dynamical setting, I disagree with this statement. Shifting this potential with a certain velocity, the shape of the probability density (and thus the entropy) will generally change (e.g., consider already the flat equilibrium density of a Brownian particle in a box that becomes tilted when applying a constant drift). The only exception where the shape and entropy indeed stays invariant is the parabolic potential. This is, however, not a general statement and I expect this to fail for the double-well example in Fig. 4d. Since Fig. 4d shows one of the main results (exceeding the limit in Eq. (7)) while the validity of Eq. (7) is here unclear due to the potential mistake in the energetic considerations, this issue urgently requires to be resolved. I would urge to also be more precise with the underlying assumptions, the respective implications of said assumptions, and the generality of the claims.

(2) Although the main message is conveyed, the presentation of Figure 1 is not very professional (somehow looks like it may have been generated by an LLM). It is not clear whether Fig. 1a is a sketch or a true trajectory (there is no label on the time-axis). Furthermore, the colored densities are sometimes (more or less) renormalized after conditioning and sometimes not. Moreover, there are further graphical artifacts and misaligned balls and lines. Please fix this.

(3) Some (parts) of the claims about the open question of "whether and how such concealed memory can be harnessed as a thermodynamic resource" (see abstract) were already answered in Ref. [32] where in theory it was shown that memory effects help to extract energy from information. Also a (theoretical) example was given. Since Ref. [32] addressed quantum systems, only theory, and offers less information-theoretic insights, I do not think that this seriously impedes the novelty of the present work. Anyway, the authors should highlight the actual novelty more precisely in their manuscript and in particular explain, to which extent the presented strategy of harnessing in this work overlaps with the theoretical example of Ref. [32] and in which aspects it is different.

(Remarks on code availability)

The raw data is in matlab format. Since I am not using matlab (nor do I have access to it), I am unfortunately not able to review the code. However, the analyses are (at least technically speaking) quite straightforward and as such I see no reason to expect any issues in this respect.

Reviewer #4

(Remarks to the Author)

This manuscript studies the optical trapping of a colloidal particle in a viscoelastic, non-Markovian bath. It shows that delayed repeated measurements can retrieve information stored in hidden environmental degrees of freedom and convert it into extractable work. This work combines a two-measurement protocol, state-conditioned relaxation analysis, a “history-blind” control construction, and work-extraction experiments in single-well and double-well potentials. The central claims are that: (i) hidden environmental memory can be operationally detected through delayed measurements; and (ii) this hidden memory can serve as a thermodynamic resource for work extraction. The work is interesting, technically solid, and potentially significant. At the same time, I have several major concerns about interpretation, mechanism, and positioning relative to prior literature. The manuscript could be publishable in Nature Physics after substantial revision, but its present form leaves important issues unresolved.

(1) Figure 2. I can understand the relaxation of states  $|11\rangle$  and  $|00\rangle$  relatively easily. The nonmonotonic relaxation of  $|10\rangle$  is both surprising and interesting. What is less intuitive, however, is the contrast between  $|10\rangle$  and  $|01\rangle$ , which does not show analogous behavior. This state specificity is not well explained in the current presentation. At a coarse level,  $|01\rangle$  and  $|10\rangle$  are two history-switching branches, so one might expect some degree of symmetry between them. The manuscript does not sufficiently explain why only one of them shows the clear backflow signature. Further elaboration of the physical origin of this observed asymmetry would strengthen the mechanistic depth of this work.

(2) While the experimental results are convincing, several of the interpretive claims would be more convincing if the manuscript could include a minimal non-Markovian model, for example, a generalized Langevin description or an equivalent hidden-variable model. This is especially important for understanding the branch-resolved asymmetry in Figures 2 and 3, the emergence of the  $|10\rangle$ -specific backflow signature, and the enhanced work extraction in the double-well case in Figure 4. At present, the manuscript provides a plausible physical narrative, but it remains more phenomenological than predictive. It would be interesting to see which ingredients of the particle-bath coupling are needed to reproduce the main features reported here, even at a semi-quantitative level.

(3) The novelty of the manuscript depends strongly on how it is positioned relative to prior work on non-Markovian baths and work recovery. In the main text, this paper does not cite closely related prior work, such as Nature Communications 11, 1360 (2020) and Nature Communications 16, 10114 (2025). In particular, the latter is a recent paper from the same research group showing a viscoelastic bath can store injected energy and later return part of it as useful work in colloidal systems. I strongly recommend that the authors revise the introduction and discussion to clarify more explicitly what is genuinely new here. Otherwise, the novelty of this work is undermined, and it may create the impression that the manuscript is overstating originality.

(4) Since the viscoelastic relaxation time of the bath is expected to depend on the micellar concentration, can the authors clarify whether the delay corresponding to the maximum integrated order parameter shifts with concentration? This would help establish how directly the observed information backflow is controlled by the bath relaxation timescale.

(5) Supplementary figure 2: can the authors show the fitting equations and the values of both short and long relaxation times?

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors answered very precisely to my questions and they clarified my doubts. The results are quite interesting and merit to be published.

I have a minor comment that the authors should take into account prior of publication.

After eq.7 and eq.16 is written that  $\varrho$  is a state of the system used to define entropy. This is not fully correct when the system is described by continuous variables as it is the case for the system studied in this article. The value of the entropy depends on the bin size used to define the states because strictly speaking one should write  $\log(P(x) dx)$  as we are dealing with probability density and not probabilities. Notice that in ref.1 they correctly use the sum in order to define the entropy. This does not affect the results because only the entropy differences are used but it is a matter of precise definition. Shannon information is infinite when the bin size goes to zero.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

The authors have addressed the comments from the reviewers appropriately. I recommend publication of this manuscript in its present form

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

The authors have consistently responded to all my questions and comments, and have revised their manuscript accordingly. I am happy to recommend publication of the revised manuscript in Nature Physics, this is now a really nice and strong piece.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)

I am satisfied with the authors' response and revision. Therefore, I recommend its publication on Nature Physics.

(Remarks on code availability)

**Open Access** This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

## Reviewer #1 (Remarks to the Author):

This article presents an interesting theoretical and experimental study of protocols which can produce work using the memory of non Markovian dynamics. The idea is original and the results, which may have important applications, could be published in Nature Physics after that the authors will take into account the following questions.

**Reply:** We thank the referee for the positive assessment and the interest in our work.

Although the article is in average well written there are several sections that are rather hard to understand and they need to be clarified. Specifically: The main idea is that in order to produce work from memory a single measurement is not enough and a delayed measure is needed. This second delayed measure is actually the key step because it gives information about the dynamical correlations that can actually be used to produce work. Figures 1 and 2 and associated equations explain rather well the two measurements methods.

**The analysis in terms of information, summarized in Fig.3 is not very clear.**

1) Is  $\tilde{I}$  used only to check the difference between a random sampling and a delayed sampling ? Indeed from eq.2 and eq.5 I need only  $I$  to bound the work. What is the point of introducing  $\tilde{I}$ ?

**Reply:** We are thankful for this important question, which allows us to clarify the role of  $\tilde{I}$ .

The referee is correct that the fundamental bounds on extractable work are given by the KL divergence and its average (the mutual information), entering Eqs. (2),(5), as well as Eqs. (12),(13) for non-Markovian systems. While being of fundamental importance, these bounds are however not experimentally accessible, as they explicitly involve the hidden variables  $y$ . This is the central reason for introducing  $\tilde{I}$ .

The quantity  $\tilde{I}$  provides an experimentally accessible reference, constructed from a history-blind ensemble, which has the same initial condition for  $x$  at  $t=0$ , but contains no information about the earlier measurement. It thus allows us to *isolate the role of memory*. Importantly:

(i) The quantity  $\tilde{I}_{|ij)}(x)$  is connected to work extraction. It captures the available information, and thus the corresponding work bound according to Eq.(2), in the considered non-Markovian as well as for the corresponding Markovian system, if we had the same starting configuration for  $x$ , but we had done only a single measurement, and hence no information about the memory. By removing trivial differences in extractable work due to changes in initial distribution,  $I_{|ij)}(x, t) - \tilde{I}_{|ij)}(x, t)$  isolates the additional information about  $x$  acquired by the double-measurement that is genuinely due to memory.

(ii) More importantly,  $\tilde{I}_{|ij)}(x)$  directly detects what we refer to as “information backflow”. Namely, for Markovian systems the future time evolution only depends on the instantaneous system state, so that  $I_{|ij)}(x, t)$  and  $\tilde{I}_{|ij)}(x, t)$  evolve identically for all  $t$ . For non-Markovian processes,  $I_{|ij)}(x, 0) = \tilde{I}_{|ij)}(x, 0)$ , by construction, but both quantities may differ at later times.

Non-zero values  $I_{|ij\rangle}(x, t) - \tilde{I}_{|ij\rangle}(x, t)$  for  $t > 0$  then capture any extra information about  $x$  that becomes available due to the influence of hidden degrees of freedom.  $\tilde{I}_{|ij\rangle}(x)$  thus provides an operational measure of information backflow about  $x$  from the environment.

We experimentally observed that tracking when this information backflow is largest gives valuable insights into the timescales of memory, and thus, can inform extraction strategies.

**Changes:** We have revised the manuscript by adding multiple text pieces to pages 4+7, to make both the connection to work extraction and information backflow (and its timescales) more explicit.

2) The description of the method used to estimate  $\tilde{I}$  is very short and Supplementary Section 3 does not help. In the Figure in SupMat the  $\delta t$  is not indicated and in the phrase "the remainder of the  $|1\rangle$  ensemble,  $|01\rangle$ ";  $|01\rangle$  should not be " $|10\rangle$ ". Probably there is a problem of wording. But I am quite confused of figure b. If I understand correctly : you select all the points in which the first reading gave 1. Starting from these points you sample again either after a fixed  $\delta t$  or a randomly chosen  $\delta t$ . From the results of the second measure you compute  $I$  and  $\tilde{I}$  respectively. Is this correct ?

If the answer is yes the Figure in SupMat and text are rather misleading especially Figure b.

**Reply:** We agree with the referee that the construction of the control ensemble and the definition of  $\tilde{I}$  were not explained clearly enough.

To start with, SI Fig.5 corresponds to a sketch of the distribution, not real data. We have updated the text and caption accordingly. In this Figure, we wanted to illustrate how the distribution  $\tilde{\rho}_{|11\rangle}$  is built from  $\rho_{|1\rangle}$ . From the question the referee asked, we understand that we failed to properly convey that all subsampling is done for a given measurement by conditioning in the past, and not in the future (as suggested by the referee).

Indeed, the reference state always corresponds to the second (not the first!) measurement at time  $t = 0$ . We then split the state proportionally to the contribution of the previous measurement at  $t = -\delta t$ . In this example,  $|1\rangle$  is then split into  $|01\rangle$  (the blue dots) and  $|11\rangle$  (the orange diamonds).

Starting from this point, we then take random trajectories that reached  $|1\rangle$  at time  $t = 0$  (which then includes both  $|11\rangle$  and  $|01\rangle$ ), but randomly sample them so that this new distribution  $\tilde{\rho}_{|11\rangle}$  matches  $\rho_{|11\rangle}$ :  $\tilde{\rho}_{|11\rangle}(t = 0) = \rho_{|11\rangle}(t = 0)$ . In other words,  $\tilde{\rho}_{|11\rangle}$  has the exact same initial shape as  $\rho_{|11\rangle}$ , but contains random trajectories from  $|1\rangle$  (so both  $|01\rangle$  and  $|11\rangle$ ), and is therefore history blind.

Finally, from  $\rho_{|11\rangle}$ , we build  $I_{|11\rangle}$ , and from  $\tilde{\rho}_{|11\rangle}$ , we build  $\tilde{I}_{|11\rangle}$  (and similarly for every state).

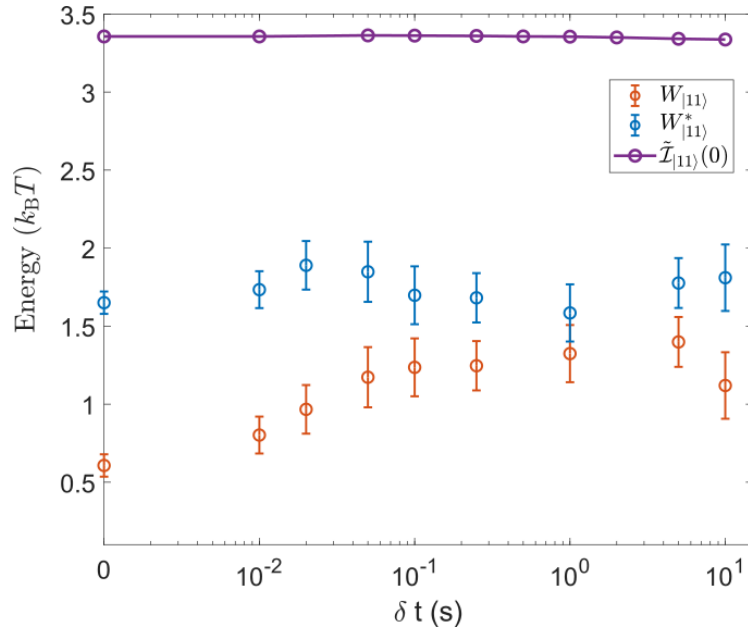
**Changes:** We have updated the figure caption and have revised the SI accordingly (see SI Note 3 and SI Fig.5).

3) The relationship between Figure 3 and the rest of the article is very confused. A paragraph which makes a clear connection between the results of Fig.3 and those of Fig.4 is needed. Furthermore, as it is explained in methods, I agree that in Fig.4 only  $\bar{V}(t)$  is needed to estimate the work produced. However it would be useful to have a plot with both  $W(t)$  and  $I(t)$  to check the quality of the bound.

**Reply:** We thank the reviewer for this constructive suggestion, which helped us clarify the connection between the non-Markovian information measure in Fig. 3 and the extracted work shown in Fig. 4.

In Fig. 3a–c, we quantify the excess information obtained through the double-measurement, as a function of the delay time  $\delta t$ , and use it to construct a non-Markovian order parameter that captures the information backflow from the hidden degrees of freedom. According to Eq. 2, such measure evaluated at time 0 sets the corresponding Markovian bound to work extraction, if one would have the same initial condition for  $x$ , but only performed a single measurement. Note that the information–theoretic bounds [Eq. 1+2] relate the total extracted work  $\max\{W(t \rightarrow \infty)\}$  to the *initial* information content of the post-measurement state, rather than to its full time evolution. Consequently, there is no rigorous relation between the time traces of  $W$  and  $I$ . While time-traces  $I(t)$  and  $\tilde{I}(t)$  characterize the relaxation of information,  $W(t)$  we show time-resolved results from a specific finite-time protocol.

Moreover, since we use rigid potential shift extraction protocols, we decided to not compare directly to the informational bound ( $I_m(0)$ ), which turns out to be very loose, but use a tighter bound ( $W_m^*$ ) specific to rigid potential shift extraction protocols as shown in the figure below for the case of  $m = |11\rangle$ . This energetic bound we however construct based on the informational bound (Eq. 8). In the main text (p. 5f, Eq. 7-9) and in the methods (p. 10), we added paragraphs to better explain how the informational and energetic bounds are connected. The initial information  $I_{|11\rangle}(0)$  is largely insensitive to  $\delta t$  because both measurements condition the observable coordinate  $x$  to remain above the threshold  $x_{th}$ , corresponding to a high-energy state in both measurements. As a result, the spatial distribution of  $x$  changes only weakly with  $\delta t$ . In contrast, the hidden degrees of freedom can occupy different configurations that depend strongly on  $\delta t$ , and their influence becomes visible only through the subsequent evolution of  $I_{|11\rangle}(t)$ . We emphasize that the main goal of our work is not to present accurate bounds on extractable work, but to unravel under which conditions work can in principle be extracted from memory, and to demonstrate that this is readily possible in experiments, even in a minimal example system.



**Comparison between extracted work and thermodynamic bounds :** We compare the maximum work we were able to experimentally extract with our rigid potential-shift protocols,  $W_{|11\rangle}$  (red), to the energetic bound  $W_{|11\rangle}^*$  (blue) derived for the corresponding Markovian system with rigid potential-shift protocols, and the more general information-theoretic bound  $I_{|11\rangle}(0)$  (violet) of the corresponding Markovian system, both shown as a function of  $\delta t$ . While the bounds depend only on the observable degree of freedom  $x$  and therefore vary weakly with  $\delta t$ , the extracted work exhibits a more pronounced dependence on  $\delta t$ , reflecting the contribution of memory effects. Since the energetic bound is specified to the case of rigid potential-shift protocols only, it is tighter than the informational bound (which applies to all types of potential manipulations).

Beyond the fundamental connection to work extraction bounds, we empirically found that the informational measures are in fact very useful to predict which states are most promising for work extraction, and on which timescale information backflow occurs. As we now explicitly state in the manuscript: The states yielding the highest information backflow also yield the highest extractable work from memory. And the time-dependent behavior of the informational order parameter resembles the state-dependent ordering of maximum work extracted from the corresponding states shown in Fig. 4a. This establishes a heuristic link between information backflow and energetic gain. To further strengthen this connection, Fig. 4b has been extended to display the extracted work for all four measurement-conditioned states, making the experimentally observed correspondence to Fig. 3 more obvious.

**Changes:** We have introduced a new version of Fig. 4b which shows the extracted work for all potential states, and extended the discussion to make the connection with the order parameter  $\phi$ .

4) The reason for the work difference between the single and double well is not clearly explained. It took a lot of time to understand this after finding the phrase "To extract additional energy stored in hidden modes, one must transiently impede particle relaxation

while the bath remains out of equilibrium". This phrase must be emphasized because the fact of producing work is not just a matter of memory effect but also of the chosen protocol.

**Reply:** The referee's comment made us aware that we did not clearly explain the origin of the difference in extracted work between the single-well and double-well potentials in the original manuscript.

It is indeed a key point that enhanced work extraction is not solely a consequence of memory, but also depends on how the protocol exploits it. In particular, extracting energy stored in hidden degrees of freedom requires that the timescale of the work extraction protocol is matched to the relaxation dynamics of the environmental memory. When this condition is fulfilled, energy stored in the environment can be very efficiently transferred back to the particle and subsequently harvested. In the double-well case, the presence of the barrier facilitates such matching and can therefore lead to larger extracted work than in a parabolic well.

In response to the referee's comment and related remarks by other referees, we have repeated our experiments on work extraction but now only using a single-well (harmonic) potential, where work extraction from hDOF can be clearly demonstrated. By doing so, we avoid the specific complexity of the double-well potential which may be distracting from our main message. The revised presentation highlights now work extraction from hDOF in a minimal setting and thereby sharpens the conceptual message of the two-measurement protocol.

**Changes:** We have removed the double-well from the manuscript altogether, and instead have focused on extracting work from the other states. In particular, we now show that  $|10\rangle$  can also largely beat the Markovian bound, which made the double-well example redundant.

5) I do not understand the phrase " To increase the number of detectable events within the limited experimental data set, we allowed  $\delta t$  to take values within the interval  $[0, 200 \text{ s}]$  and aggregated the corresponding events to improve statistical significance. ". The protocol last 1.2s and the delay 200s. Does it mean that you drive the system in a NESS ?

**Reply:** In the double-well potential, transitions between the two wells (e.g.,  $|01\rangle$  or  $|10\rangle$ ) require barrier-crossing events, which are relatively rare. Restricting the analysis to a fixed delay time  $\delta t$  would therefore require exceedingly long measurements to obtain sufficient statistics. To overcome this limitation, we aggregated events corresponding to measurement pairs with  $\delta t \in [0, 200 \text{ s}]$  which improves the statistical sampling of such rare events.

The delay  $\delta t$  refers only to the temporal separation between two measurements used to identify suitable initial conditions. It is not related to the duration of the work extraction protocol. Each feedback protocol is triggered only after such an event is detected and has a duration of  $t_f = 1.2 \text{ s}$ . The individual realizations are independent, and the system is not continuously driven during the interval  $\delta t$ . Therefore, the system is not operated in a non-equilibrium steady state.

We acknowledge that this adds a layer of complexity that was somewhat distracting from our main message. As mentioned above and to avoid complications that arise due to the rich

dynamics within double-well potentials, including the issue of the aforementioned choice of varying  $\delta t$ , we performed additional work extraction experiments using the double well and decided to only present these in the revised version. Here,  $\delta t$  is simply a fixed constant, allowing indeed a more straightforward interpretation of the results.

**Changes:** Same as above, the double-well was confusing and redundant, so we removed it from the manuscript, and used  $|10\rangle$  as an example instead.

6) minor problem. it is never written that the acronym SEM stands for Standard Error of the Mean. It is not necessarily known.

**Changes:** In the revised version, we now spell out the acronym SEM upon its first use.

7) Finally concerning non-Markov demon there are recent articles on harmonic oscillators which will be worth to quote: Archambault et al. PRL 135, 147101 (2025); Archambault et al, EPL 148 41002 (2024).

**Reply:** We thank the referee for the valuable hint. We fully agree that this work needs to be cited since it theoretically addresses several issues related to our work. We apologize for having overlooked these studies and have gladly included them at the end of the introduction.

**Changes:** Both references were added to the revised version.

## Reviewer #2 (Remarks to the Author):

This manuscript shows how to improve work extraction from a stochastic system by exploiting its non-Markovian character, using a two-time measurement protocol. I find this to be an exciting result, highlighting that memory effects are not just a nuisance complicating the equations of motion, but a source of information about hidden bath degrees of freedom.

**Reply:** We appreciate the encouraging comment and are glad that the referee shares our excitement for this shift in perspective, which was precisely our motivation behind the work.

The paper is very clear and well written, and I recommend publication. I have two questions for the authors to consider:

1. The authors consider a specific artificial measurement protocol, in which the system is mapped onto two discrete states, below threshold  $|0\rangle$  and above threshold  $|1\rangle$ . One could increase the number of states or instead report on the actual values of the coordinate at the times of the measurements. Likewise, one could consider 3-time measurements. Would such more elaborate measurement protocols further improve work extraction? What are the limits on how much can be gained?

**Reply:** We thank the referee for this insightful question, which we have gladly added to our discussion. Indeed, more refined measurements (increased spatial resolution or additional time points) can enhance work extraction. This occurs for two reasons: (1) repeated measurements reduce measurement uncertainty (error), and (2) repeated measurements reveal (more of the) correlations with hidden degrees of freedom, thereby providing indirect access to additional nonequilibrium free energy.

However, these improvements are indeed bounded. According to the generalized second law of information thermodynamics, the total extractable work (independent of the measurement scheme) is fundamentally limited by the conditional nonequilibrium free energy of the system at the onset of control,

$$W_m \leq k_B T I_m(Z) = F_{\text{neq}}[p(z|m)] - F_{\text{eq}},$$

where  $m$  denotes a specific measurement outcome  $m$  and  $Z$  the complete system state (including hidden degrees of freedom). To get the corresponding bound for the full measurement record one averages the above equation over  $m$ .

Thus, even in the limit of infinitely many (ideal) measurements, one can at most fully access the free energy already contained in all degrees of freedom in the (thermally excited) initial system state, but not exceed it, which would lead to saturation upon adding more measurements.

**Changes:** We have included the above equation to the revised manuscript (now Eq. 7), and added to p.2: *“To investigate work extraction beyond the Markovian limit, we focus in this work on the minimal measurement scheme capable of accessing memory, namely performing two consecutive measurements only, and using binary resolution (“left/right”). As we demonstrate, this minimal setup already suffices to extract work from memory.”*; as well as to the conclusion (p.7): *“In this work, we have demonstrated work extraction*

*beyond the Markovian limit, focusing on the minimum measurement scheme to access memory—two binary measurements—and an extraction protocol restricted to rigid potential shifts. By using more flexible and optimized extraction protocols, or by increasing the number of measurements or their spatial resolution to increase the accessible information, the extractable work could further be enhanced; up to the fundamental limit set of work extraction is set by the total excess free energy of the (thermally excited) system state at the onset of control (7).”*

2. Figure 2 reports on conditional relaxation curves obtained from experimental measurements. It would be helpful if one could get some theoretical insight into those. For example, since the bath has a single relaxation timescale, could the dynamics  $x(t)$  be modeled as a generalized Langevin equation with an exponential memory kernel (or, equivalently, could the hidden degrees of freedom be described by a single degree of freedom coupled harmonically to  $x$ )? If so, it may be possible to have a simple theory for the data appearing in Figs. 2-3

**Reply:** Yes, the referee is fully correct in suggesting a simple minimal model. Indeed, the colloidal dynamics  $x(t)$  can be described by a generalized Langevin equation with an exponential memory kernel, which is formally equivalent to introducing a hidden degree of freedom represented by an elastically coupled fictitious bath particle.

Initially, we deliberately did not include such a model in the main text, as a central point of our work is that the extraction of work from hidden degrees of freedom does not rely on detailed knowledge of their microscopic nature, but is obtained directly from experimental data in a model-independent manner.

In response to the referee’s request, we have nevertheless added this minimal model to the Supplementary Information (Note 4), as it indeed serves as an intuitive illustration of the observed dynamics. Specifically, the model corresponds to a Maxwell-type viscoelastic bath, in which the colloidal particle (with friction coefficient  $\gamma$ ) is coupled via a harmonic spring of stiffness  $\kappa_b$  to a second, hidden bath particle (with friction  $\gamma_b$ ). The coupled overdamped Langevin equations for the (observable) colloidal and the (hidden) bath particle explicitly encode the exchange of elastic energy and, when combined, are equivalent to a single non-Markovian Langevin equation with an exponentially decaying memory kernel. This representation makes transparent how viscoelastic memory arises from the delayed relaxation of the bath.

Within this framework, the model naturally captures the key physical mechanism underlying our observations: after a fluctuation-induced displacement of the colloidal particle, the bath particle cannot instantaneously follow due to its larger effective friction, leading to a transient elastic force that depends on the system’s history. This delayed relaxation gives rise to state-dependent effective forces and thus to the markedly different conditional relaxation curves.

In particular, the model qualitatively reproduces the state-dependent relaxation behavior shown in Fig. 2, which arises from the coupled relaxation of the colloidal particle and the hidden bath degree of freedom. Since the relative position of the bath particle with respect to

the threshold depends sensitively on the initial state, the resulting relaxation dynamics can differ strongly between conditions.

**Changes:** We have added a micro-mechanical bath-particle model to the SI (Note 4). We have also included corresponding simulation data (SI Fig. 3), demonstrating good agreement with the experimental results shown in Figs. 2 and 3 as requested by the referee. In addition, we also show an example of how the extracted work compares to our experimental values for all stats and for  $\delta t = 1\text{s}$  (see SI Fig. 4b).

### **Reviewer #3 (Remarks to the Author):**

The present manuscript clearly presents, by experiments and by theory, how to extract thermodynamic work from degrees of freedom that are not directly observable. Besides some advances on the theoretical framework, it truly provides (for the first time) a full experimental confirmation of such extraction of work. Non-Markovian effects, in particular hidden/unobserved degrees of freedom and their role in thermodynamics, are gaining more and more traction and are attracting increasingly more attention in various fields (simply judging based on the most recent perspective articles and back-to-back special issues in not one but in fact several journals). The results in the present manuscript are likely to significantly advance not only the field of stochastic thermodynamics, but also statistical physics, biophysics, and even beyond. The results are original and convincing up to some points that I would ask the authors to address in the revision (see below).

**Reply:** We thank the referee for the positive assessment. We agree that the central contribution of this work lies in the experimental proof-of-principle demonstration, and we appreciate the referee's recognition of the importance of such experimental validation.

Information stored in hidden degrees (including non-Markovian effects) is a challenging but very important and timely field of research, which (given the intense efforts in the field) is expected to soon transgress from conceptual challenges to technological applications. Since the results in the manuscript instructively and clearly present a useful case for addressing this hidden information, they concurrently make a strong case for further studies in similar and related directions. Overall, I recommend accepting the paper in Nature Physics under the condition that the following (critical) remarks are satisfactorily answered.

(1) There seems to be one issue with the energetic considerations. Below Eq. (7), it is mentioned that this form is due to "rigid translations not giving rise to changes in equilibrium free energy and Shannon entropy". While the first point is correct as shown in Eq. (16), above Eq. (18) it is only mentioned (not shown) that the Shannon entropy is "invariant under rigid potential transformations". In the given dynamical setting, I disagree with this statement. Shifting this potential with a certain velocity, the shape of the probability density (and thus the entropy) will generally change (e.g., consider already the flat equilibrium density of a Brownian particle in a box that becomes tilted when applying a constant drift). The only exception where the shape and entropy indeed stays invariant is the parabolic potential. This is, however, not a general statement and I expect this to fail for the double-well example in Fig. 4d. Since Fig. 4d shows one of the main results (exceeding the limit in Eq. (7)) while the validity of Eq. (7) is here unclear due to the potential mistake in the energetic considerations, this issue urgently requires to be resolved. I would urge to also be more precise with the underlying assumptions, the respective implications of said assumptions, and the generality of the claims.

**Reply:** We thank the referee for this important comment. We agree that the point is subtle and important and benefits from clarification, which we provide below.

During finite-time protocols that affect the potential, including those in which the potential is only shifted, the probability density generally evolves and its Shannon entropy thus indeed dynamically changes. However, this dynamical evolution does actually not enter the fundamental work extraction bound (14) considered here. Rather, the maximal extractable work is determined solely by the difference between the nonequilibrium free energy of the post-measurement state at the onset of control (i.e. at  $t=0$ ) and the equilibrium free energy of the final potential after full relaxation (Eq. 14) (at  $t \rightarrow \infty$ ). This is why this dynamical change does not enter.

An important conceptual step is then to go from the fundamental bound to the optimal work extraction strategy. As the referee correctly notes, rigid transformations do not affect the resulting *equilibrium* Shannon entropy, irrespective of whether the potential is harmonic or not. Maximizing the extracted work thus amounts to performing a control that effectively minimizes how much of the nonequilibrium free energy of the initial state can be exploited. The optimization over  $\Lambda$  is thus performed at the level of the initial distribution, not along the dynamical evolution of  $\rho$ . Thereby it is important to note that any relaxation of the distribution prior to or during the shift generally reduces the available free energy, since  $F_{\text{neq}}$  is non-increasing under relaxation. To see this, we can use the general connection to the KL-Divergence:  $k_B T D_{\text{KL}}(\rho(t) || \rho_{\text{eq}}) = F_{\text{neq}}[\rho(t); V] - F_{\text{eq}}$  and note that the KL-Divergence is monotonically decreasing under Markovian relaxation, implying that relaxation is generally not beneficial to work extraction in this setting.

To summarize, the validity of Eq. (7) (or Eq. 17 in the appendix) relies only on the invariance of the *equilibrium* free energy under rigid translations, but it does not rely on any assumption about the entropy during the dynamical evolution. Therefore, the result is not restricted to harmonic potentials and applies equally to nontrivial landscapes such as the double-well potential shown in Fig. 4d. We hope this resolves the apparent contradiction raised by the referee.

**Changes:** To provide a better explanation of the energetic bound for rigid-potential translations, we have significantly extended the discussion in the method. In particular, we now give a hopefully more vivid explanation by decomposing the dynamics into infinitesimal time steps of shifting and intermediate relaxation (see page 10).

More importantly, we found that the Markovian limit can already be surpassed in a simple harmonic potential by conditioning on the  $|10\rangle$  state, making the double-well geometry unnecessary for demonstrating work extraction from memory. We therefore removed this part from the manuscript and focused exclusively on the single-well system. This substantially simplifies the presentation by avoiding the additional complexity of double-well dynamics, which is not essential to the demonstration of work extraction from memory and, as highlighted by the referee reports, tended to distract from the main results.

In the method section (p10), we also have added a short discussion using the optimal control framework (for the specific example of harmonic potential), which provides the same energetic bound from an entirely different route.

(2) Although the main message is conveyed, the presentation of Figure 1 is not very professional (somehow looks like it may have been generated by an LLM). It is not clear whether Fig. 1a is a sketch or a true trajectory (there is no label on the time-axis).

Furthermore, the colored densities are sometimes (more or less) renormalized after conditioning and sometimes not. Moreover, there are further graphical artifacts and misaligned balls and lines. Please fix this.

**Reply:** We thank the reviewer for pointing out these issues and agree that Figure 1 required improvement.

The trajectory shown is experimental data displayed with a reduced acquisition rate to improve readability and limit visual noise; this is now explicitly stated in the caption. We also corrected the inconsistent normalization of the conditional densities and added a clearly labelled time axis to panel (a). The apparent misalignment of the spheres is intentional: they indicate the mean particle positions associated with the displayed probability distributions, corresponding to the end points of the measurements shown in Fig. 2b. We apologize that this clarification was absent in the original manuscript and have now included it in the revised caption.

**Changes:** We have revised the figure and caption accordingly.

(3) Some (parts) of the claims about the open question of "whether and how such concealed memory can be harnessed as a thermodynamic resource" (see abstract) were already answered in Ref. [32] where in theory it was shown that memory effects help to extract energy from information. Also a (theoretical) example was given. Since Ref. [32] addressed quantum systems, only theory, and offers less information-theoretic insights, I do not think that this seriously impedes the novelty of the present work. Anyway, the authors should highlight the actual novelty more precisely in their manuscript and in particular explain, to which extent the presented strategy of harnessing in this work overlaps with the theoretical example of Ref. [32] and in which aspects it is different.

**Reply:** We fully agree that Ref. [32] (Poulsen et al. PRE (2022) (now Ref. [36]) demonstrates at the theoretical level that non-Markovian memory can enhance the performance of a Maxwell demon, and it is indeed instructive to compare the underlying strategies in Poulsen et al. and our study. Both works demonstrate that information backflow from the environment can be exploited as a resource. In the Poulsen paper, this is achieved by appropriately timing the demon operation relative to the system–bath dynamics, thereby taking advantage of enhanced predictability induced by non-Markovian memory. In our work, a closely related role is played by the delay time  $\delta t$  between measurements, which selects regimes where memory effects and information backflow are maximal.

At the same time, the operational approach differs in an essential way. Poulsen et al. consider a specific, fully controlled quantum model in which the demon protocol is synchronized with the underlying coherent dynamics. In contrast, our double-measurement strategy does not rely on microscopic knowledge or control of the environment. Instead, it provides an experimental and model-independent protocol that accesses and exploits hidden environmental memory through measurement outcomes alone. In this sense, rather than synchronizing with the dynamics, our approach selects memory-bearing environmental states via conditioning on two-time measurements.

Furthermore, our work introduces a direct experimental quantification of information backflow, by comparing measurement-conditioned ensembles to a history-blind reference.

This allows us to isolate the contribution of hidden degrees of freedom independently of initial-state effects and to link it quantitatively to enhanced work extraction.

**Changes:** We have clarified this relation and the distinction at the end of the introduction. The revised manuscript now better highlights how our approach complements and extends the theoretical ideas put forward in Poulsen et al.

Reviewer #3 (Remarks on code availability):

The raw data is in matlab format. Since I am not using matlab (nor do I have access to it), I am unfortunately not able to review the code. However, the analyses are (at least technically speaking) quite straightforward and as such I see no reason to expect any issues in this respect.

**Reply:** We apologize for only providing data in Matlab format, and understand the difficulty that can arise for non-Matlab users. In fact, Python users can use the `scipy.io.loadmat` library to read our data.

**Changes:** We added this information in the read-me of the dataset.

#### Reviewer #4 (Remarks to the Author):

This manuscript studies the optical trapping of a colloidal particle in a viscoelastic, non-Markovian bath. It shows that delayed repeated measurements can retrieve information stored in hidden environmental degrees of freedom and convert it into extractable work. This work combines a two-measurement protocol, state-conditioned relaxation analysis, a “history-blind” control construction, and work-extraction experiments in single-well and double-well potentials. The central claims are that: (i) hidden environmental memory can be operationally detected through delayed measurements; and (ii) this hidden memory can serve as a thermodynamic resource for work extraction. The work is interesting, technically solid, and potentially significant. At the same time, I have several major concerns about interpretation, mechanism, and positioning relative to prior literature. The manuscript could be publishable in Nature Physics after substantial revision, but its present form leaves important issues unresolved.

**Reply:** We thank the referee for the positive assessment and appreciate their recognition of the interest and technical soundness of our work. We are also grateful for the comments and hints which we have taken into account in a substantial revision.

(1) Figure 2. I can understand the relaxation of states  $|11\rangle$  and  $|00\rangle$  relatively easily. The nonmonotonic relaxation of  $|10\rangle$  is both surprising and interesting. What is less intuitive, however, is the contrast between  $|10\rangle$  and  $|01\rangle$ , which does not show analogous behavior. This state specificity is not well explained in the current presentation. At a coarse level,  $|01\rangle$  and  $|10\rangle$  are two history-switching branches, so one might expect some degree of symmetry between them. The manuscript does not sufficiently explain why only one of them shows the clear backflow signature. Further elaboration of the physical origin of this observed asymmetry would strengthen the mechanistic depth of this work.

**Reply:** Indeed, the relaxation behaviour of  $|10\rangle$  and  $|01\rangle$  states are quite surprising and reflect a non-trivial impact of hDOF. Following the suggestion of the referee (see next question) we have added a simple micromechanical model which considers the hDOF in our system by introducing a fictitious (i.e. hidden) bath particle which is connected by a harmonic spring to the visible colloid (see SI Note 4). Notably, the friction coefficient of the bath particle is typically larger than that of the colloid which reproduces the time-delayed colloidal response to an external force being characteristic in viscoelastic fluids. As a side remark we note, that such model yields an exponential memory kernel in agreement with experimental observations, e.g. the diffusive motion or the overcoming over a potential barrier in a viscoelastic fluid (see e.g. Ginot et al. PRL 128, 028001 (2022)).

Within such model, the non-monotonic relaxation of the  $|10\rangle$  state is understood as follows: Opposed to equilibrium where the spring between the colloid and the bath particle is fully relaxed, this does no longer apply after a binary measurement. After a  $|10\rangle$  state, the colloid has moved within  $\delta t$  from above ( $|1\rangle$ ) to below ( $|0\rangle$ ) the threshold. However, the bath particle is lagging behind, and is still (on average) above the threshold. Thereby the spring to the bath particle is expanded which leads to an elastic force pulling the colloid towards the bath particle (i.e. up within the potential). At the same time, however, the optical trap is forcing the colloid down towards the potential minimum. This competition eventually explains

the non-monotonic relaxation seen in Fig.2a which is also corroborated by our numerical simulations which are now added to the SI (see SI Fig. 3).

Overall  $|10\rangle$  and  $|01\rangle$  still show some degree of (anti)symmetry.  $|10\rangle$  is a low energy state which has positive feedback from the hDOF (due to the memory of the  $|1\rangle$ ), and relaxes slower than expected (and has a non-monotonic relaxation).  $|01\rangle$  is the opposite, as it starts in a high energy state, but relaxes faster than expected due to the memory of the initial  $|0\rangle$  encoded in the hDOF.

(2) While the experimental results are convincing, several of the interpretive claims would be more convincing if the manuscript could include a minimal non-Markovian model, for example, a generalized Langevin description or an equivalent hidden-variable model. This is especially important for understanding the branch-resolved asymmetry in Figures 2 and 3, the emergence of the  $|10\rangle$ -specific backflow signature, and the enhanced work extraction in the double-well case in Figure 4. At present, the manuscript provides a plausible physical narrative, but it remains more phenomenological than predictive. It would be interesting to see which ingredients of the particle-bath coupling are needed to reproduce the main features reported here, even at a semi-quantitative level.

**Reply:** Following the request of the referee (see also referee 2), we have included a minimal model (established already in, e.g., Ginot et al. PRL, 128, 028001 (2022)), which indeed helps to understand the branch-resolved asymmetry demonstrated experimentally. Such model, which is now explicitly discussed in the Supplementary Information (see Note 4) describes the hidden degree of freedom within the viscoelastic bath using a fictitious (i.e. experimentally invisible) “bath particle” that is elastically coupled to the (visible) colloidal particle. Because the friction coefficient of the bath particle is larger than that of the colloid, this reproduces the observed time-delayed motion of a colloid in our micellar solution. Indeed such a model corresponds to a generalized non-Markovian Langevin equation with the hidden position of the bath particle corresponding to the hDOF (and generating the memory kernel and coloured noise, when integrated out).

As for the  $|10\rangle$  case, mentioned by the referee, this means that during the motion of the colloid from above to below the threshold within  $\delta t$ , energy is dumped into the non-Markovian surrounding by loading the spring to the bath particle. This energy will be released during particle relaxation but also during work extraction and explains the results shown in Figs. 2, 3, and 4.

However, while we fully agree that the model is helpful to gain a more intuitive understanding of our results, we aim to emphasize that assuming a concrete model is not essential to conduct the work extraction experiments, and our analytical calculations are largely model-independent. To emphasize the generality of our results, we decided to present the numerical simulations only in the supplemental notes ( SI notes 4, 5 and 6) and refer to this in the main text.

**Changes:** We have added a simple micro-mechanical model to the SI (notes 4-6), which help rationalize the experimental observations.

(3) The novelty of the manuscript depends strongly on how it is positioned relative to prior work on non-Markovian baths and work recovery. In the main text, this paper does not cite closely related prior work, such as Nature Communications 11, 1360 (2020) and Nature Communications 16, 10114 (2025). In particular, the latter is a recent paper from the same research group showing a viscoelastic bath can store injected energy and later return part of it as useful work in colloidal systems. I strongly recommend that the authors revise the introduction and discussion to clarify more explicitly what is genuinely new here. Otherwise, the novelty of this work is undermined, and it may create the impression that the manuscript is overstating originality.

**Reply:** We thank the referee for this comment, and used it to improve the discussion, in particular regarding how our work relates to and differs from the cited studies.

First, the work by Debiossac *et al.* (Nat. Commun. 11, 1360 (2020)) investigates non-Markovian effects arising from a finite time delay in an externally applied feedback loop. In this framework, memory is effectively introduced by the controller, and, in particular, is associated with a breaking of detailed balance (since the fluctuation dissipation relation of the underlying Langevin equation is broken due to the feedback force), meaning this form of memory acts itself as a thermodynamic driving that puts the system intrinsically out of equilibrium. Debiossac *et al.* presents a thermodynamic description to account for the associated entropy pumping. In contrast, our work focuses on intrinsic environmental memory, which originates from hidden degrees of freedom of the bath, and does not break detailed balance (our system reaches thermal equilibrium). The environment naturally retains correlations with the system's past and no external feedback is required to generate memory. Our approach therefore addresses a conceptually different question, namely, how such intrinsic (equilibrium) memory can be accessed and exploited as a thermodynamic resource.

Second, the recent work by Ginot & Bechinger (Nat. Commun. 16, 10114 (2025)) demonstrates energy recuperation in a viscoelastic bath under non-equilibrium conditions. In such work energy has been first injected into the environment during a driven protocol and, subsequently, partially recovered as useful work. This mechanism relies on energy storage and delayed release, i.e. prior to extraction the bath has to be actively driven out of equilibrium. Later one regains part of the injected energy due to its slow relaxation dynamics.

In contrast, our work operates in a fundamentally different regime. Here, we consider an initially equilibrium system and show that net work can be extracted purely by harnessing information contained in *equilibrium* time-correlated fluctuations. No prior energy injection into the bath is required. In this sense, the two approaches are conceptually distinct. Our work establishes environmental memory itself as a thermodynamic resource, thereby clearly going beyond Ref. (2025).

Despite these conceptual differences, we fully agree that both mentioned works are related and relevant in the broader context of non-Markovian thermodynamics. We are grateful for the suggestion which helps to put in context and discern our work from previous approaches.

**Changes:** We have now included and discussed both citations in the introduction to more

clearly position our contribution.

(4) Since the viscoelastic relaxation time of the bath is expected to depend on the micellar concentration, can the authors clarify whether the delay corresponding to the maximum integrated order parameter shifts with concentration? This would help establish how directly the observed information backflow is controlled by the bath relaxation timescale.

**Reply:** The referee is right to expect the maximum of the order parameter to depend on the material properties of the viscoelastic fluid (specifically, the maximum depends on the overall relaxation time of the colloidal particle given by the relaxation time of the bath and the relaxation within the optical trap). Unfortunately, however, variations in the micellar solution only weakly change the relaxation dynamics. This has been demonstrated in Caspers et al. J. Chem. Phys. **158**, 024901 (2023) which reports the concentration-dependent relaxation times of the used micellar solution which only varies very little in the range of 5-9mM. Outside that range, experiments are difficult to perform due to either little viscoelastic behaviour or because of the viscosity which strongly increases with concentration and very much slows down the particle dynamics.

(5) Supplementary figure 2: can the authors show the fitting equations and the values of both short and long relaxation times?

**Reply:** We agree that these equations were missing from the SI, and thank the referee for pointing this out.

We have included the double exponential equation which has been used fitting our data in SI Fig.1 together with the derived fitting parameters (see SI Note 2). The fitting curve corresponds to the analytical solution of a recoil curve obtained for a two-bath particle model (see Caspers et al. NJP **24**, 123013 (2022)). To reduce the number of fitting parameters for the numerical simulations (see above), we used an effective single-bath particle model derived from the double exponential fit parameters and now explained in the SI (Note 4).

**Changes:** We have updated the SI and their figures accordingly.

**Reviewer #1 (Remarks to the Author):**

The authors answered very precisely to my questions and they clarified my doubts. The results are quite interesting and merit to be published.

Reply : We thank the Reviewer for the positive evaluation of our work and for recommending its publication.

I have a minor comment that the authors should take into account prior of publication. After eq.7 and eq.16 is written that  $\rho$  is a state of the system uses to define entropy. This is not fully correct when the system is described by continuous variables as it is the case for the system studied in this article. The value of the entropy depends on the bin size used to define the states because strictly speaking one should write  $\log(P(x) dx)$  as we are dealing with probability density and not probabilities. Notice that in ref.1 they correctly use the sum in order to define the entropy. This does not affect the results because only the entropy differences are used but it is a matter of precise definition. Shannon information is infinite when the bin size goes to zero.

Reply : We also appreciate the Reviewer for pointing out the subtle issue regarding the definition of entropy for continuous variables. We have replaced the equation for entropy with a sum as suggested.

**Reviewer #2 (Remarks to the Author):**

The authors have addressed the comments from the reviewers appropriately. I recommend publication of this manuscript in its present form

Reply : We thank the Reviewer for the careful reading of the manuscript and for recommending its publication.

**Reviewer #3 (Remarks to the Author):**

The authors have consistently responded to all my questions and comments, and have revised their manuscript accordingly. I am happy to recommend publication of the revised manuscript in Nature Physics, this is now a really nice and strong piece.

Reply : We sincerely thank the Reviewer for the positive assessment of our revised manuscript and for appreciating our revisions.

**Reviewer #4 (Remarks to the Author):**

I am satisfied with the authors' response and revision. Therefore, I recommend its publication on Nature Physics. We thank the Reviewer for the careful evaluation of our revised manuscript and for recommending its publication.

Reply : We thank the Reviewer for the careful evaluation of our revised manuscript and for recommending its publication.