The authors consider stochastic processes in Markov networks and, in particular, the statistics of the fluctuating, time-integrated current across an individual edge of a network. For a statistical characterisation of these currents, this work focuses on the distribution of the infimum of the current, against an overall positive trend. Using methods from the theory of martingales, it is shown that this distribution is always a geometric one. The strength of this work is that it provides an operationally accessible way to evaluate the effective affinity of a single edge of a network, of which all other edges are invisible. This unlocks the possibility to apply bounds on the overall rate of entropy production that have been derived before (by one of the present authors, in Refs. [15,18]) and where this effective affinity enters as well. Unlike in this previous work, it is now possible to evaluate the effective affinity from only typical fluctuations, circumventing the unpractical procedure involving the limit of exponentially unlikely fluctuations. The distribution of the infimum and the resulting bound on the entropy production (compared to established bounds) are illustrated numerically for a well-known model for a molecular motor. Remarkably, this bound on the entropy production can be much tighter than the established bounds.

The article is well written, with the right attention to detail, and the results are indeed relevant. I appreciate the splitting of the work into two "companion papers", along with Ref. [1]. This manuscript is self-contained (and I understand the other one is so as well), and the equivalence of the results is discussed in Sec. 2.2. Attempting to sketch both derivations (which are complementary) in one manuscript would have overwhelmed the reader.

A few points remained unclear to me and there are a number of minor points that need to be fixed. I expect that the authors will be able to address these points and change the manuscript adequately, such that the paper will be ready for acceptance.

Our Reply: We thank the Referee for writing a detailed report and providing us with comments to improve the paper.

(a) *In the introduction, I'd find it important to mention what class of physical systems is being considered, to provide some context for the first few paragraphs. The bound mentioned in the first paragraph is fairly universal, yet systems with inertia or coherent quantum dynamics can break it.*

Our Reply:

The result (14) is valid generically for stationary, Markov jump processes. No physical assumptions are required for this result.

On the other hand, the bound on the rate of dissipation, equation (21), relies on the formula (58) for the rate of dissipation $\dot{s}$, which is valid in classical stochastic processes for which all variables have even parity under time reversal. This assumption is valid when inertia can be neglected and when the process is not under the influence of a magnetic field.

Following the Referee's suggestion, we have clarified this in the Introduction, for example:

"A second reason why extreme values of currents are interesting is because they can be used to estimate the average rate $\dot{s}$ in stationary processes consisting of variables that have even parity under time reversal, which is the case when inertia is negligible and external forces are not governed by magnetic fields, see Refs. [9,14,17]. "

and

"In stationary processes with variables that have even parity under time reversal, the derived results for the extreme value statistics of edge currents can be interpreted in terms of an effective thermodynamic picture of a marginal observer that only sees the edge current J, ignorant of the existence of other currents in the system."

(b) *It would be good to discuss in more detail the relation between the "effective affinity" and the "true affinity" (whatever that is). When are the two the same? I suspect that $a^*(x \to y)$ is more related to the overall cycle affinity (in case of a unicyclic system) than to the true edge-affinity $a(x \to y)$, even though the notation would suggest some relation to the latter.*

Our Reply: For unicyclic systems, the effective affinity equals the "true", macroscopic affinity, while the edge affinity is smaller, as we show below. On the other hand, for multicyclic systems it is difficult to relate effective affinities to true affinities, as the current through a single edge contains contributions from several macroscopic currents.

Let us come back to the unicyclic case, in particular, let us consider the Markov process

$$\partial_t p(x;t) = p(x+1;t)k_{x+1\to x} + p(x-1;t)k_{x-1\to x} - (k_{x\to x+1} + k_{x\to x-1})p(x;t) \tag{1}$$

with $x \in \mathcal{X} = \{1, 2, \ldots, \ell\}$ and where it is understood that $\ell + 1 = 1$ and $0 = \ell$.
Let us first discuss the expressions for the microscopic affinities. At stationarity,

$$\overline{j}_{x-1\to x} = \overline{j}_{x\to x+1} = \overline{j}, \tag{2}$$

and therefore,

$$\dot{s} = a\overline{j} \tag{3}$$

with the "true" affinity

$$a = \sum_{x=1}^{\ell} a_{x\to x+1} \tag{4}$$

that is a sum of all microscopic affinities. Hence, in general the microscopic affinities contribute a small part of the total affinity. For example, when $k_{x\to x+1} = k_+$ for all x, and $k_{x\to x-1} = k_-$ for all x, then

$$a_{x\to x+1} = \ln \frac{k_+}{k_-} \tag{5}$$

for all x, and

$$a = \ell \ln \frac{k_+}{k_-}. \tag{6}$$

Let us now consider the effective affinity. The effective affinity is given by

$$a^*_{x\to x+1} = \ln \frac{p^{x,x+1}_{\rm ss}(x) k_{x\to x+1}}{p^{x,x+1}_{\rm ss}(x+1) k_{x+1\to x}}. \tag{7}$$

In the present case,

$$p^{x,x+1}_{\rm ss}(x) = p_0 \prod_{y=1;y\neq x}^{\ell} \frac{k_{y\to y+1}}{k_{y+1\to y}} \tag{8}$$

and

$$p^{x,x+1}_{\rm ss}(x+1) = p_0, \tag{9}$$

where p_0 is a normalisation constant. Therefore,

$$a^*_{x\to x+1} = \sum_{y=1}^{\ell} a_{y\to y+1} = a. \tag{10}$$

To clarify this in the manuscript, we have added the following sentence in the paragraph after equation (69):

"*Although, in general, microscopic affinities are not simply related to the "true" macroscopic affinities, for unicyclic systems it holds that the effective affinity equals the macroscopic affinity, while the microscopic affinity captures in this case a possibly small portion of the macroscopic affinity (see also Appendix C).* "

In addition, we have added Appendix C to the manuscript that clarifies the meaning of this sentence.

(c) *p.7: "the system attains the stalling state, which according to the marginal observer is indistinguishable from equilibrium [1]". It may be distinguishable if the marginal observer does the right analysis, e.g., detect coherent oscillations in the autocorrelation function of the current. Also, I couldn't find the term "marginal observer" in [1], so I couldn't find what it has to say about this.*

Our Reply: This phrase was indeed confusing. We have removed it from the manuscript.

(d) *In Sec. 3, the concept of a current being "proportional" to the entropy needs to be specified better. Is it exactly proportional or up to finite boundary terms?*

Our Reply: In the example considered, the current has to be exactly proportional to the total stochastic entropy production, as otherwise $e^{-cJ_{x\to y}}$ is not a martingale. We have put the equation $S(t) = cJ_{x\to y}(t)$ in display to make this more clear.

(e) *I don't see how to get from (44) to (46); how can the current and the factor 1/2 be identified? It seems to me that we need to have $l(u \to v) = k(v \to u)$ (which is the case in the following applications of this formula).*

Our Reply: The Referee is correct. We require the additional condition

$$\frac{\ell_{u\to v}}{k_{u\to v}} = \frac{k_{v\to u}}{\ell_{v\to u}}, \tag{11}$$

which is the case when we use this formula in section 5. We have clarified this accordingly in the paper.

(f) *Sec. 6.4 (and also mentioned elsewhere in this manuscript): I doubt that it is possible to have "all microscopic affinities zero except one". When the affinity $\ln((p_u k(u- > v))/(p_v k(v- > u))$ is zero, then also the current $(p_u k(u- > v)) - (p_v k(v- > u))$ is zero. But, because of the Kirchhoff rules for stationary currents, it is not possible to have a current through one edge, without a backflow through some other edge(s). In Sec. 7.2.1 it looks like the affinity is only the log-ratio of transition rates (without the contribution from the stationary distribution). In that case, it may be possible to have only one "affinity" equal to zero. But I'm not sure whether the argument from 6.4 then still applies.*

Our Reply: The Referee is correct, and we thank the Referee for spotting this mistake.

What we meant to say is that the process is driven out of equilibrium by a single edge. In other words, the Markov process $(X, \mathbb{P}^{x,y})$ obtained by removing the transitions from x to y and backwards from the Markov transition graph satisfies detailed balance. We have rephrased the corresponding text in Section 6.4 accordingly.

(g) *In Sec. 7, I'd welcome some more discussion. Beyond confirming the geometric distribution, what other information can be extracted from this kind of measurement? For example, is it possible to decide whether the observed current stems from a single or from several edges? In the latter case, is there a strong deviation from the geometric distribution?*

Our Reply: Yes if the current is not coming from a single edge, then the distribution

$$p_-(\ell) = \mathbb{P}_{\rm ss}\left(J^{\rm inf}_{x\to y} = -\ell\right) \tag{12}$$

is not geometric. However, for large enough values of ℓ, it holds that

$$p_-(\ell) = e^{-a^*\ell(1+o_\ell(1))}, \tag{13}$$

where $o_\ell(1)$ represents an arbitrary function converging to zero when ℓ is large enough, and a^* a constant value that depends on the current J under consideration. Note that because of $o_\ell(1)$ in the exponent, $p_-(\ell)$ approximates a geometric distribution for large values of ℓ, but, in general, won't be geometric for small values of ℓ.

We can test whether $p_-(\ell)$ is geometric by determining the modified infimum ratio

$$\hat{\hat{s}}_{\rm inf}(\ell) = \ln\frac{p_-(\ell)}{p_-(\ell+1)}. \tag{14}$$

If $\hat{\hat{s}}_{\rm inf}(\ell)$ is constant, then $p_-(\ell)$ is geometric.

In the case of an edge current, $J = J_{x\to y}$,

$$\hat{\hat{s}}_{\rm inf}(\ell) = a^*_{x\to y} \tag{15}$$

for all values of ℓ, except for $\ell = 0$.

For other currents, $\hat{\hat{s}}_{\rm inf}(\ell)$ approaches a constant value for large enough values of ℓ, but is nonconstant for small ℓ. We illustrate this on a simple model in the new Figure 6 of the manuscript, and discuss this problem in the Appendix H, which is a new addition to the manuscript.

(h) *It is not clear to me how the bound derived from the infimum statistics can be so much stronger than the TUR. In which cases can the bound be saturated? (without the TUR saturating at the same time?) In Ref. [15], it is stated (below Eq. (19) there) that the equivalent bound on the splitting probability follows from the TUR. So how can it be stronger? In the caption of Fig. 5, it is stated that the TUR is evaluated at 100s, which I assume is long enough to capture the long-time limit. Yet, the finite-time TUR can (in some cases) be stronger. How would the TUR evaluated at the optimal time interval compete with the bound from the infimum statistics? I think this would be a "fair game", because also the latter takes into account fluctuations on finite time scales.*

Our Reply:

This question has three parts.

(a) First part: "*It is not clear to me how the bound derived from the infimum statistics can be so much stronger than the TUR. In which cases can the bound be saturated? (without the TUR saturating at the same time?)*:

Indeed, it is quite remarkable that the bound with infimum statistics is so much stronger than the TUR. For general currents, it is currently not well understood what governs the deviation from the equality. However, there is one particular case that is well understood, which is when the current is proportional to the entropy production, i.e., $J = cS$. As has been shown with martingale theory, see [I. Neri, Universal tradeoff relation between speed, uncertainty, and dissipation in nonequilibrium stationary states, SciPost Phys. 12, 139 (2022)], the infimum statistics bound is saturated when $J = cS$.

(b) Second part: "*In Ref. [15], it is stated (below Eq. (19) there) that the equivalent bound on the splitting probability follows from the TUR. So how can it be stronger?*":

This has been explained in detail in Section 6.3 of the paper ["I. Neri, Universal tradeoff relation between speed, uncertainty, and dissipation in nonequilibrium stationary states, SciPost Phys. 12, 139 (2022)."], which is the Ref.[15] that the Referee refers to.

Let us first note that below Eq.(19) of Ref.[15] it is **not** stated that the bound on the splitting probability follows from the TUR. Instead, it is stated that the bound on the splitting probability follows from the following parabolic bound on the large deviation function

$$\mathcal{J}(z) \leq \frac{\dot{s}}{4}\left(z/\overline{j} - 1\right)^2, \tag{16}$$

which is a general result that was conjectured in the paper [Pietzonka, Patrick, Andre C. Barato, and Udo Seifert, "Universal bounds on current fluctuations.", Physical Review E 93.5 (2016): 052145.] and proven in [Gingrich, T. R., Horowitz, J. M., Perunov, N., England, J. L. (2016). Dissipation bounds all steady-state current fluctuations. Physical review letters, 116(12), 120601.]. Note that the Eq. (16) implies the TUR for values $z \approx 0$, denoting the typical fluctuations around the stationary current $\overline{j}$, but the inequality is more general as it also applies to atypical fluctuations for $z \neq 0$.

As explained in Section 6.3 of Ref.[15], when $J = S$, the derivation of the inequality

$$\dot{s} \geq \frac{\ell_+}{\ell_-}\frac{|\log p_-|}{\langle T\rangle} \tag{17}$$

relies on evaluating the bound Eq. (16) at $z = -\dot{s}$. In this case the bound Eq. (16) is satisfied as an equality. Indeed, the Gallavotti-Cohen fluctuation relation implies that

$$\mathcal{J}(z) - \mathcal{J}(-z) = -z \tag{18}$$

and hence for $z = -\dot{s}$ it holds that

$$\mathcal{J}(-\dot{s}) = \dot{s}. \tag{19}$$

(c) Regarding the last part: *Yet, the finite-time TUR can (in some cases) be stronger. How would the TUR evaluated at the optimal time interval compete with the bound from the infimum statistics? I think this would be a "fair game", because also the latter takes into account fluctuations on finite time scales.*

Indeed, in general, the quality of the TUR varies as a function of time t. However, in the present case of a Markov jump process driven far away from thermal equilibrium, t plays a minor role in the quality of

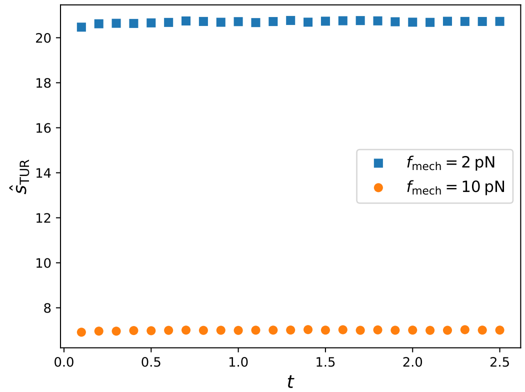

Figure 1: The thermodynamic uncertainty ratio $\hat{s}_{\mathrm{TUR}}$ as a function of t for the same parameters as in Figure 5 of the the submitted Manuscript and for two values of f_{mech}, as indicated in the legend.

the TUR. In Figure 1 of the present document, we show $\hat{s}_{\mathrm{TUR}}$ as a function of t for the same parameters as in Figure 5 of the Paper and for two values of f_{mech}. As can be seen, the $\hat{s}_{\mathrm{TUR}}$ is in the present case insensitive to the time interval t. We added the comment "*other t give similar results*" in the caption of Figure 5.

Minor points:

1. *abstract: "current" $\to$ "integrated current" (the more familiar concept of a current is that of a rate, where the extreme value would be ill-defined)*

   Our Reply: We agree with the Referee and have replaced current by integrated current at the beginning of the manuscript. We also mention below Eq.(11):

   "*Notice that, for convenience, we often speak of edge currents, tout court, and it should be understood that we consider empirical, integrated currents.* "

2. *abstract: "estimate" $\to$ "bound from below" (If I'm not mistaken, the bound can still be way off. For an "estimate" I would expect a guess with a lower and upper confidence interval)*

   Our Reply: Yes, but let us emphasize that in statistics an estimator can be biased. Of course, ideally an estimator is unbiased.

   The problem with the wording "bound from below" is that the average entropy production is a nonnegative number, and hence you can always bound it from below.

   To find a reasonable compromise, we say now: "*can estimate a finite fraction of the average entropy production rate*", which we believe gets the point across.

3. *"largest excursion against its average flow" sounds like $max(|J- < J > |)$, which would not make sense. I think this can be made more precise with a better word choice (or dropped altogether, the explanation before and the Figure already make the concept clear).*

   Our Reply: We have changed the sentence into "*in the direction that opposes the average flow*", which is more precise

4. *In which sense are (1) and (2) "analogous"? Can (1) be derived from (2)?*

   Our Reply: Yes that is correct, equation (1) can be derived from equation (2). This follows from

   $$\langle S_{\mathrm{inf}} \rangle = \int_{-\infty}^{0} d\ell [1 - \mathbb{P}_{\mathrm{ss}} \left(S^{\mathrm{inf}} \leq \ell \right)] \geq \int_{-\infty}^{0} d\ell [1 - e^{\ell}] = -1. \tag{20}$$

5. *"Interestingly, far from equilibrium, s_{inf} captures ..." What is the meaning of s_{inf} without the argument l? Is the limit to infinity implied, or is l arbitrary?*

Our Reply: Now, to be more precise, we specify that ℓ is large, namely, the text reads now:

"*$\hat{s}_{\rm inf}(\ell)$, for large values of ℓ,...*"

6. *The notation $s_{\rm inf}(l)$ in (9), (10) is somewhat confusing, since this quantity does not depend on l.*

   Our Reply: The modified infimum ratio is defined as

$$\hat{\hat{s}}_{\rm inf}(\ell) := \overline{j} \ln \frac{p_-(\ell)}{p_-(\ell+1)}. \tag{21}$$

   For $\ell \in \mathbb{N}$ it holds that

$$\hat{\hat{s}}_{\rm inf}(\ell) = \overline{j} a^*, \tag{22}$$

   but for $\ell = 0$, we have that

$$\hat{\hat{s}}_{\rm inf}(0) = \overline{j} \left[\ln \left(\frac{p_{\rm esc}(x_0)}{1 - p_{\rm esc}(x_0)} \right) - \ln \left(1 - e^{-a^*} \right) \right]. \tag{23}$$

   It should be noted that, a priori, it is not obvious that $\hat{\hat{s}}_{\rm inf}(\ell)$ is independent of ℓ when $\ell \in \mathbb{N}$, hence we think it makes sense to keep the dependence on ℓ in the definition of $\hat{\hat{s}}_{\rm inf}$. Also, given the fact that $\hat{\hat{s}}_{\rm inf}(0) \neq \hat{\hat{s}}_{\rm inf}(1) = \hat{\hat{s}}_{\rm inf}(2) = \ldots$, the ℓ independence in strictly not true for all $\ell \in \mathbb{N} \cup \{0\}$.

   To make this more clear, we changed the equality sign in (9) by a definition sign "*:=*".

7. *Below (16): I'm confused about the difference between "geometric" and "exponential" distribution. According to wikipedia, they (only?) differ by whether the support is discrete or continuous.*

   Our Reply: Yes correct. That is why the geometric distribution converges to an exponential distribution in the limit when $a^*_{x \to y} \approx 0$, as then the characteristic scale of the geometric distribution $e^{-\ell a^*_{x \to y}}$ is large and the discreteness of the distribution becomes irrelevant.

8. *Eq. (19): The $\to \infty$ is typeset as subscript (please check this throughout, I've also seen this elsewhere, e.g. (125), (126)).*

   Our Reply: We thank the Referee for spotting these typo's, and we have corrected them.

9. *Is the $f-$in Eq. (23) just $1 - p^{\rm esc}$? If so, say so.*

   Our Reply: Yes that is correct. We have added this clarification in the formula.

10. *Sec. 3: "Such an observer thinks that the observed current J is proportional to entropy production." Just because (s)he is naive, or is that simplest assumption indeed consistent with all observations?*

    Our Reply: We have clarified this as follows:

    Unaware of the existence of other currents in the system, the observer thinks...

11. *Eq. (31), typo $a \to c$?*

    Our Reply: Indeed! Thanks for spotting this. We have corrected the typo.

12. *Typo: Full stop at end of sentence including (33).*

    Our Reply: Many thanks for spotting this typo as well. It has been corrected.

13. *Beginning of Sec. 4.1: A few examples for "events" may be helpful ("event" sounds rather time-local, but I believe sigma can depend on several times)*

Our Reply: we have added the phrase:

"*and the set of measurable events contains, amongst others, the sets* $\{\omega \in \Omega : X(t,\omega) = x\}$ *for all* $t \geq 0$ *and* $x \in \mathcal{X}$. "

14. *typo: "devined"*

Our Reply: Fixed.

15. *Around (37), (38): Difference between Markov process and Markov chain?*

Our Reply: Yes, there is some inconsistency in terminology here. Now, we always speak of Markov jump process.

16. *I never understood why it's called Radon-Nikodym "derivative" (and not "ratio"). For a derivative I'd expect the notion of a small difference. What is the meaning of "dQ" and "dP" in (44)?*

Our Reply: The Radon-Nikodym derivative can indeed be seen as a derivative. Recall that $\mathbb{P}$ and $\mathbb{Q}$ are probability measures that assign probabilities to sets Φ of trajectories. In other words, $\mathbb{P}[\Phi]$ is the probability to observe a trajectory in the set Φ of trajectories. Then, the ratio

$$\frac{d\mathbb{P}[X_0^t]}{d\mathbb{Q}[X_0^t]} \tag{24}$$

can be considered to be

$$\frac{d\mathbb{P}[X_0^t]}{d\mathbb{Q}[X_0^t]} = \lim_{\mathbb{P}[\Phi]\to 0} \frac{\mathbb{P}[\Phi]}{\mathbb{Q}[\Phi]}, \tag{25}$$

where Φ represents a decreasing sequence of sets containing the trajectory X_0^t, i.e., $X_0^t \in \Phi$, and the Radon-Nikodym derivative is obtained in the limit when $\mathbb{P}[\Phi] \to 0$.

For example, let us consider the simpler case where $\mathbb{P}$ and $\mathbb{Q}$ are two probability measures defined on the Lebesgue measurable subsets of the real line $\mathbb{R}$, such that

$$\mathbb{P}[(a,b)] = \mathbb{P}\left[X \in (a,b)\right] \tag{26}$$

and correspondingly for $\mathbb{Q}$. Let us consider the Radon-Nikodym derivative

$$\frac{d\mathbb{P}[x]}{d\mathbb{Q}[x]} \tag{27}$$

with $x \in \mathbb{R}$ an element of the real line. Then, this ratio can be obtained from the limit

$$\frac{d\mathbb{P}[x]}{d\mathbb{Q}[x]} = \lim_{\epsilon\to 0} \frac{\mathbb{P}[(x-\epsilon, x+\epsilon]]}{\mathbb{Q}[(x-\epsilon, x+\epsilon]]} \tag{28}$$

Defining the cumulative distributions

$$g_{\mathbb{P}}(x) = \mathbb{P}\left[X \leq x\right] \tag{29}$$

and

$$g_{\mathbb{Q}}(x) = \mathbb{Q}\left[X \leq x\right] \tag{30}$$

we can also write this as

$$\frac{d\mathbb{P}[x]}{d\mathbb{Q}[x]} = \lim_{\epsilon\to 0} \frac{g_{\mathbb{P}}(x+\epsilon) - g_{\mathbb{P}}(x-\epsilon)}{g_{\mathbb{Q}}(x+\epsilon) - g_{\mathbb{Q}}(x-\epsilon)}. \tag{31}$$

If $g_{\mathbb{P}}(x)$ and $g_{\mathbb{Q}}(x)$ are differentiable towards x, then the probability density functions

$$p_X(x) = \frac{dg_{\mathbb{P}}(x)}{dx} \tag{32}$$

and

$$q_X(x) = \frac{dg_{\mathbb{Q}}(x)}{dx} \tag{33}$$

exist, and

$$\frac{d\mathbb{P}[x]}{d\mathbb{Q}[x]} = \frac{p_X(x)}{q_X(x)}. \tag{34}$$

However, Radon-Nikodym derivative can be defined generally for probability measures defined on measurable spaces, and this includes trajectory spaces of stochastic processes, which do not have a Lebesgue measure.

17. *Below (52): "than" → "then"*

    Our Reply: Thanks! Fixed.

18. *Before (54), "$|M(t)| < c$". Does c need to be independent of T?*

    Our Reply: Yes it should be. The quantity c is a constant, which means that it is a number independent of X and t, and hence it is also statistically independent of T. For example, if

    $$T = \inf\{t \geq 0 : |M(t)| \leq c\} \tag{35}$$

    then the condition is satisfied.

    We clarified now:

    "*and there exists a constant $c \in \mathbb{R}^+$ such that $|M(t)| < c$ for all $t \leq T$*"

19. *"which is nota bene different from $p^{xy}_{\rm ss}$": worth pointing out that it is also different from $p_{\rm ss}$*

    Our Reply: Sure, we have pointed this out as well.

20. *Before (58): Typo $P_{q_{ss}} \to P_{p_{ss}}$?*

    Our Reply: Yes, correct! We have fixed this.

21. *In (60): is $R_{q_{\rm ss}} = Q_{\rm ss} \circ \Theta$?*

    Our Reply: Yes, that is correct. We have added this comment after Eq.(67).

22. *"Lastly, introducing the effective microscopic affinity": is this the same quantity as in Ref. [30]?*

    Our Reply:It is, and we added the reference for clarity.

23. *In what sense is the meaning of the effective affinity "kinematic" (as opposed to the true efficiency?)*

    Our Reply: It is kinematic in the sense that

    $$a^*_{x\to y}\overline{j}_{x\to y} > 0 \tag{36}$$

    and thus the direction in which the current flows is dictated by $a^*_{x\to y}$.

24. *p. 15: Notation switches from $J_{x->y}(T)$ to $J(T)$. Make consistent or at least introduce the abbreviation.*

Our Reply: Thank you for spotting this typo, which we have corrected.

25. *I think "generic" should be replaced by "general" in Sec. 6.2. Surely, the initial condition x, which is not "generic", is included in this general case.*

Our Reply: We have followed the Referee's suggestion.

26. *Between (81) and (82), the sentence starting with "Therefore ..." appears out of context with the preceding sentence ("Notice ...").*

Our Reply:We have rephrased this as follows:

"*Notice that for general initial conditions, Eq. (67) holds for values $\ell_- \neq 0$ and $\ell_+ \neq 0$, but not when either of the two thresholds is zero, and therefore $\ell_- \neq 0$ in Eq. (80). From Eqs. (80) and (81) follows that the probability mass function of $J^{\rm inf}_{x\to y}$, which for $\ell \in \mathbb{N}$ is given by...*"

27. *Below (85): I believe this should read "for a *near* stalled current", and a* should be approximately 0, otherwise $p_e sc$ and j would be zero exactly.*

Our Reply: Correct. We have rephrased the sentence accordingly.

28. *Before (95): It is not clear how the parameter dependence "does not contribute to dissipation". The affinity doesn't depend on the parameter, but the kinetics depends on it, hence the dissipation rate (if this can be equated to "dissipation") indirectly depends on the parameter as well.*

Our Reply: Yes correct. We have modified the text as follows:

"*However, since $\chi_{ij} = \chi_{ji}$, it holds that*

$$\frac{k_{i\to j}(f_{\rm mech})}{k_{j\to i}(f_{\rm mech})} = \frac{k_{i\to j}(0)}{k_{j\to i}(0)} \tag{37}$$

for $(i, j) \notin \{(2, 5), (5, 2)\}$ and all values of $f_{\rm mech}$, and consequently $f_{\rm mech}$ provides a nonzero contribution to the microscopic affinity $a_{2\to 5}$ only. "

29. *Caption Fig. 3: "lines"→"line" (I just see one line)*

Our Reply: Thanks for pointing this out. We have corrected it.

30. *Sec. 7.2.3: typo "stalled stated"*

Our Reply: Indeed, we have fixed it.

31. *p. 24: "However, when the statistics of the current J contain strong non-Markovian effects and when J is not proportional to the entropy production S, than $s_{\rm KL}$ provides a poor estimate ..." With strong non-Markovian effects, can $s_{\rm KL}$ even be measured? (Before, it is stated that non-Markovian effects are ignored).*

Our Reply: Yes absolutely. The definition of $\hat{s}_{\rm KL}$ is

$$\hat{s}_{\rm KL} = \sum_{j\in\mathcal{J}} \dot{n}_j \ln \frac{\dot{n}_j}{\dot{n}_{-j}} \tag{38}$$

which has also a definite value when J is nonMarkovian. However, in this case $\hat{s}_{\rm KL}$ will be significantly lower than $\dot{s}$ as the nonMarkovian correlations contain information about the arrow of time.

32. *I am not sure whether I understand the distinction between s hat and s double-hat. If (121) "applies" only to a single edge, should't it be a double-hat? I understand that in any case, the estimated entropy production is that of the overall network, not just a single edge.*

Our Reply: The $\hat{s}_{\rm inf}$ is defined in Eq.(3) [note that now we specify := to indicate that this is a definition] and $\hat{\hat{s}}_{\rm inf}$ is defined in Eq.(9). These are in general different quantities.

However, for edge currents they are related by Eq.(10). In other words,

$$\hat{\hat{s}}_{\rm inf}(\ell) = \lim_{\ell' \to \infty} \hat{s}_{\rm inf} \tag{39}$$

for all $\ell \in \mathbb{N}$ (which excludes zero as $\hat{\hat{s}}_{\rm inf}(0) \neq \lim_{\ell' \to \infty} \hat{s}_{\rm inf}$).

Coming to Eq.(127) [before (121)]. In this case we have the closely related first-passage ratio $\hat{s}_{\rm FPR}$ as defined in (123). The infimum ratio $\hat{s}_{\rm inf}$ is a specific case of the first-passage ratio, in the sense that

$$\hat{s}_{\rm inf}(\ell_-) = \lim_{\ell_+ \to \infty} s_{\rm FPR}(\ell_-, \ell_+). \tag{40}$$

Equation (127) states that

$$\hat{s}_{\rm FPR} = a^*_{x \to y} \bar{j}_{x \to y}(1 + o_{\ell_{\rm min}}(1)), \tag{41}$$

in other words, $\hat{s}_{\rm FPR} = a^*_{x \to y} \bar{j}_{x \to y}$ in the asymptotic limit that both thresholds diverge.