nature portfolio

Peer Review File

Existence of multiple transitions of the critical state due to anesthetics

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The authors look at the changes induced by several anesthetic agents into the large scale brain dynamics. The main strength of the work is three fold: 1) the use of optogenetics techniques which allows the simultaneous recording of large brain areas, 2) the use of statistical techniques appropriate to evaluate the scaling of avalanches activity and 3) the use of a a wide variety of agents in the same experimental setup allowing for comparison not attempted before.

The author unveils multiple ways in which the mouse brain moves away from the awake critical dynamics. The results are very relevant to further our understanding the neural bases of consciousness and its disturbances by sleep, pharmacological agents or pathologies.

I have no concerns at all on the technical aspects of the results, they are described extremely well and they are very convincing. I would only recommend that the authors attempt to either compute the correlation length scaling with box size (as done recently by others (e.g. Martin et al [1] ; Suweis group[2]) or deposite the data in Github in a format that other colleagues can do it in the near future. The authors surely would agree that such calculations may reinforce and widening some of the conclusions based on the avalanches scaling.

Minor:
1-Could the authors mention if the observation of awake-like state is consistent with the REM state described for some but not all anesthetics agents.

2- In [3] fMRI voxels data was clustered (a different calculation from avalanches here) leading to a PDF scaling of 3/2. Any chance that similar computation can be attempted here?

3- how it is interpreted the value of 0.9 ~ 1.15 for the size avalanches exponent ?

Dante R Chialvo

Refs:

[1]Martin, D. A. et al. Box scaling as a proxy of finite size correlations. Sci. Rep. 11, 1–9 (2021).

[2] Mariani, B., Nicoletti, G., Bisio, M. et al. Disentangling the critical signatures of neural activity. Sci Rep 12, 10770 (2022). https://doi.org/10.1038/s41598-022-13686-0

[3]E Tagliazucchi, P Balenzuela, D Fraiman DR Chialvo. Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis- Frontiers in Physiology…, 2012 - frontiersin.org

Reviewer #2 (Remarks to the Author):

The study by Curic et al. addresses changes in cortical dynamics during anesthesia in transgenic mice using 2-photon imaging. Spontaneous activity was recorded over a large area of the brain and was analyzed for avalanche dynamics using three different types of anesthetics at 2 levels each – light and surgical plane. The authors focus on the change in avalanche size distributions with respect to the wake, quiet resting state and anesthesia. They find that size distributions change more when reaching the surgical plane in particularly with ketamine. The authors conclude that diverse changes in size distributions indicate significant individual variability and diverse underlying mechanisms.

Major critique:
1. The novelty of this study is unclear. Many studies have demonstrated significant and variable changes in cortical dynamics with increasing dosage of anesthetics using a variety of measures. The current study using avalanche size distribution statistics is in line with those general reports. Even the more distinct effect reported here for ketamine is in line with the common observation of burst suppression at surgical plane. It is unclear how the use of avalanche measures contributes significantly to our understanding of the effects of anesthesia on the brain.
2. The second problem with this study is the claim of different mechanisms underlying the effect of anesthetics on the awake, quiet state. The study would improve significantly if the authors indeed could identify or isolate at least one of those mechanisms. Unfortunately, currently those claims remain rather non-specific, are derived mainly from some of the known receptor targets (which on a global scale are rather difficult to pin down mechanistically), and are suggested to include potential differences in the brain regions involved.
3. There are several technical aspects not considered by the authors that weaken their study further. While the two-photon imaging field has somewhat gravitated to using the raw Df/f-signal, its complexity though makes interpretations of the underlying neuronal dynamics difficult. For example, absent spike deconvolution, the df/f decay is dominated by the decay in the fluoroprobe used reducing temporal resolution. Similarly, estimating f_0 from the average signal is highly unusual and can lead to false reconstructions of the underlying spike trains if non-stationarities are not controlled for. This latter point is particularly important given that the effects of some anesthetics are known to wear off quickly making it difficult to properly control the depths of anesthesia for the duration of the recording.

Reviewer #3 (Remarks to the Author):

Curic and coworkers explore the changes in critical behavior of the mouse cortex upon different anesthetic interventions, and show that light anesthesia largely preserves scale-free statistics, while surgical one changes the statistics depending on the anesthetic agent. The results are very interesting since they open new avenues in understanding the relation between chemical perturbations on a neuronal circuit and its dynamical state. The article is clear, very well written and organized. The experiments and analytical tools are convincing, and the results are of interest for a broad scientific community, from physics of complex systems to medicine.
However, before publication, I would like the authors to clarify some aspects of their work, as elaborated below.

Major concern:
1. The present study is somehow related to the previous work of the group (Yaghoubi et al., Sci Rep 2018), in which they studied criticality in primary neuronal cultures. There, the authors explored different critical exponents and their relation through scaling laws. A relation between critical exponents derived from the scaling laws provides validation to the analyses, but it was missing in the present study. Can the authors comment on this aspect? For instance, can they measure the duration of avalanches and extract other exponents?

Minor:
2. Experimental data is obtained through fluorescence calcium imaging on a large window of the mouse cortex, of about 9x9 mm2. Each pixel on the images may therefore contain the information of several neurons, many of them acting as “hidden layers” that affect the behavior of the accessed pixels but without direct monitoring. Can the authors explain if they have taken this background activity into account?
3. The authors use the genetically encoded calcium indicator jrGECO1a to monitor activity. Other sensors may differ in kinetics and brightness upon calcium binding and may therefore change the shape of the fluorescence traces. Can the authors comment on possible artifacts associated to the selected fluorescence probe?
4. How do the authors understand the fact that a disruption of the excitation-inhibition balance in light anesthetics lead to the roughly the same power law statistics? It is very surprising to me.
5. I found very interesting the analysis of avalanche initiation. In some cases, initiation is extremely focalized, as for ketamine 100 mg/kg. Maybe the authors could quantify the focalization of initiation by computing the Gini coefficient, as in e.g. Montalà-Flaquer et al., “Rich dynamics and functional organization on topographically designed neuronal networks in vitro”, iScience 2022. Additionally, do the authors have a model to understand why this extreme focalization occurs?

Sincerely yours,

Jordi Soriano
University of Barcelona

Reviewer Reply

Dear Editor,

We would like to thank the reviewers for their careful attention to detail and thoughtful suggestions for improving the manuscript. We believe this manuscript is much improved with this set of revisions and we are confident it will be of broad interest to the fields of physics and neurophysics, complex systems, and the broader anesthetics/medicine communities.

Below, we provide the referees' comments verbatim (burgundy, italicized) and a point-by-point description of the additional experiments, analyses, and in-text clarifications that have been added to the manuscript as a result of the reviewers' suggestions. Note that any numeric references to items such as figures, equations, tables, etc., are enumerated according to the current version of the manuscript or supplementary material. Separate documents with tracked changes for both the manuscript and supplementary material have been submitted as well.

Point by point replies

Reviewer 1:

The authors look at the changes induced by several anesthetic agents into the large scale brain dynamics. The main strength of the work is three fold: 1) the use of optogenetics techniques which allows the simultaneous recording of large brain areas, 2) the use of statistical techniques appropriate to evaluate the scaling of avalanches activity and 3) the use of a a wide variety of agents in the same experimental setup allowing for comparison not attempted before.

The author unveils multiple ways in which the mouse brain moves away from the awake critical dynamics. The results are very relevant to further our understanding the neural bases of consciousness and its disturbances by sleep, pharmacological agents or pathologies.

Comment I

I have no concerns at all on the technical aspects of the results, they are described extremely well and they are very convincing. I would only recommend that the authors attempt to either compute the correlation length scaling with box size (as done recently by others (e.g. Martin et al [1] ; Suweis group[2]) or deposite the data in Github in a format that other colleagues can do it in the near future. The authors surely would agree that such calculations may reinforce and widening some of the conclusions based on the avalanches scaling.

We entirely agree with the reviewer's point. We have tried this analysis and present it in Fig. R1 where we plot the characteristic length (as measured by the zero crossing of the connected correlation function, r_0) as a function of the window length, as well as the measured slope.

At criticality Martin *et al.* suggest r_0 should scale linearly with the linear dimension of the box (at least up to $L/2$, where L is the system size). We do indeed observe linear scaling up to a window length close to 100 pixels (Fig. R1(a)), which is consistent with our system size. Note that our system size is limited — a box of around 200×100 pixels — as we need to exclude the right hemisphere due to the high symmetry. In addition, this box contains masked-out pixels that do not correspond to the cortical surface, roughly about 10%.

Interestingly, we find similar results for all conditions, i.e., there are no significant differences between Quiet-wakefulness and the different drug conditions or deviations from linear scaling (Fig. R1(b)). While this could suggest that all high drug dose cases can be considered critical — in contrast to our findings based on avalanche statistics — it is important to realize that within the small range of box sizes we have available, all of the sub/super/critical cases basically look the same (e.g., Martin *et al.*, Fig. 2). Thus, we believe that the analysis of the correlation length scaling is not sensitive enough given our experimental limitations.

Instead, we have now included the analysis of avalanche durations in our revised manuscript to reinforce and widen some of our conclusions (see our Reply to Reviewer 3's Comment I).

Comment II

Minor: 1-Could the authors mention if the observation of awake-like state is consistent with the REM state described for some but not all anesthetics agents.

We considered this but concluded it was unlikely. While anesthetics are known to cause disruptions in REM/NREM (e.g., the time spent in either), the awake-like state spans the entirety of the recording, rather than switching. While getting an estimate for the duration of REM bouts in mice is not entirely

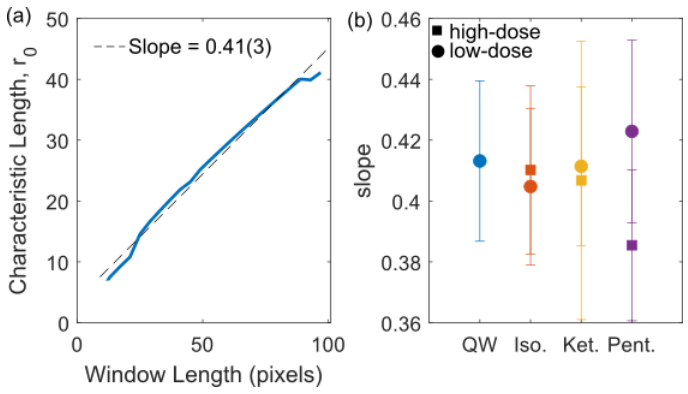

Figure R1: (a) Characteristic length determined by the first zero crossing of the connected correlation function, as a function of the box window length, averaged across all quiet-wakefulness recordings. (b) Measured slope for all conditions, not including the high-dose cases do not including awake-like recordings.

straightforward, a survey of publications regarding the topic seems to suggest an upper bound of about 2-3 minutes [Funk et al., 2016, McShane et al., 2012, Soltani et al., 2019]. Our recordings are 7.5 minutes and we did not observe indication of any sort of switching between REM and NREM in these 7.5 minutes across any of the awake-like recordings. Thus, if this was related to REM it would suggest the REM (or NREM) state stays for the duration of the recording, which is not, to our knowledge, consistent with any of the anesthetics we analyzed. We now explicitly mention in the manuscript that the awake-like states are stationary over the duration of our recordings.

Comment III

2- In [3] fMRI voxels data was clustered (a different calculation from avalanches here) leading to a PDF scaling of 3/2. Any chance that similar computation can be attempted here?

Yes, we have done this exact computation which is now included in the supplementary material. The method in [3], namely generating avalanche statistics from the point-processed signal (i.e., keeping only the time points when the signal crosses the threshold from below), is very similar to the method we had used previously to generate Fig. S8 of the supplementary material, which shows that point-processed signals exhibit the same avalanche size distribution as those in the main text, albeit with a different cutoff. The main difference was the clustering radius: Whereas in [3] co-activated pixels are clustered according to neighbouring pixels, we took a wider radius (determined by the partial correlation length) in order to account for long-range activation.

Figure R2a shows that if this neighbourhood is reduced to nearest neighbours, but the signal is *not* point-processed (solid line in Fig. R2), the critical exponent, estimated from avalanches concatenated across all QW recordings, is $\tau = 1.20(4)$ (if the wider neighbourhood is used, the corresponding exponent is 1.14(4)). Using point-processing in addition, the slope shifts to a measured exponent of $\tau = 1.41(6)$ (dashed line in Fig. R2). Yet, the range over which a power-law is a plausible fit is significantly reduced, with $s_{max}/s_{min} \approx 20$ (as opposed to ≈ 1000 in the non-point-processed case and in [3]). Such a reduced fitting range typically leads to larger systematic uncertainties. This together with the fact that the other three ways to define avalanches discussed here (as well as the one based on deconvolved calcium signals, see our Reply to Reviewer 2's Comment III and corresponding Fig. R4) give consistent results in the size exponent with $\tau \approx 1.2$ suggests that the latter value might be a more truthful characterization of the dynamics. Most importantly, however, regardless of which approach is used, whether the ones presented above in Fig. R2b, below in Fig. R4 or those presented in Fig. S8 and Fig. S10 of the supplementary material, our overall conclusions remain the same — power-law statistics in quiet-wakefulness are disrupted in a non-universal fashion by the different anesthetics.

Comment IV

3- how it is interpreted the value of $0.9 \sim 1.15$ for the size avalanches exponent ?

In light of reviewer 3's comment I, we have switched from reporting the critical exponents of the concatenated data to averages over individual recordings. Overall none of the conclusions changed, though we did observe that the exponents slightly shifted to higher values, albeit largely within the statistical uncertainties. For example, the QW size exponent went from $\tau = 1.14(4)$ to 1.2(1). Regardless, the reviewers question is obviously still valid.

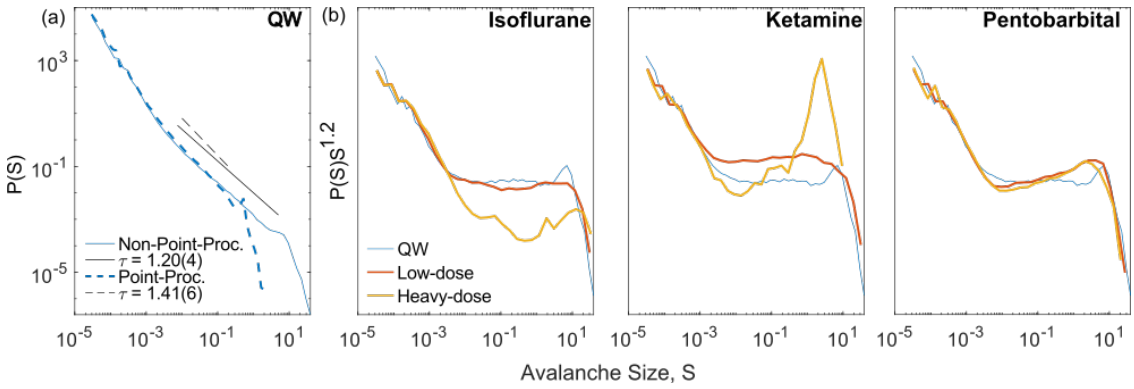

Figure R2: (a) Quiet-wakefulness avalanches using adjacent pixels rather than the expanded neighbourhood used in the main paper. Distributions are shown for point-processed (dashed) and non-point-processed signals (solid), with their associated estimated exponents. (b) The non-point-processed distribution from panel a) is compared against the equivalent distribution across different drug cases (not including the awake-like cases). Each curve is rescaled to the $\tau = 1.2$ obtained in QW to show differences more clearly.

Exponents less or equal to 1 as we observe in some limited cases are fundamentally different from those larger than 1. This is because the former range of values is theoretically problematic since they violate normalizability for power-laws that are not truncated for $S_{max} \to \infty$. At the same time, the standard scaling relation between τ, α, and γ no longer holds since its derivation requires $\tau, \alpha > 1$ (see also our Reply to Reviewer 3's comment I). Because our MLE estimator includes a finite S_{max}, it does allow for $\tau \leq 1$. These cases then indicate that the system size and the finite S_{max} have become vital to the analysis. Yet, it is important to point out that out of the five recordings for which we obtain an estimate of $\tau \leq 1$ (out of a total of 50 recordings) the statistical uncertainties allow for the actual exponent to be bigger than 1 in three cases already at the one-sigma level and in the other two cases at the two-sigma level. Thus, it is possible that $\tau > 1$ in all cases, as expected for the standard notion of criticality which requires that power-law behavior holds in the asymptotic limit.

Exponents in the range $1 < \tau \lesssim 1.2$ are significantly smaller than what is expected for a mean-field branching process, for which $\tau = 3/2$. This indicates that in our case the underlying network structure of the cortex is quite distinct from an all-to-all coupling, consistent with our analysis of the partial correlation function. In fact, these smaller values of τ are closer to what is expected for a branching process on a two-dimensional lattice, namely $\tau \approx 1.26$. This is now discussed in our revised manuscript. Note that a more local coupling at the mesoscale resolution at which we observe the dynamics of the cortex does not contradict the observations with different imaging modalities at smaller scales with very limited fields of view (see, for example, [Fontenele et al., 2019]), for which an all-to-all coupling aka a mean-field branching process might be more appropriate, though sub-sampling issues could also affect the estimation of the exponents in those cases [Levina et al., 2022, Carvalho et al., 2021].

Refs: [1] Martin, D. A. et al. Box scaling as a proxy of finite size correlations. Sci. Rep. 11, 1–9 (2021). [2] Mariani, B., Nicoletti, G., Bisio, M. et al. Disentangling the critical signatures of neural activity. Sci Rep 12, 10770 (2022). https://doi.org/10.1038/s41598-022-13686-0 [3]E Tagliazucchi, P Balenzuela, D Fraiman DR Chialvo. Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis-Frontiers in Physiology. . . , 2012 - frontiersin.org

Reviewer 2:

The study by Curic et al. addresses changes in cortical dynamics during anesthesia in transgenic mice using 2-photon imaging. Spontaneous activity was recorded over a large area of the brain and was analyzed for avalanche dynamics using three different types of anesthetics at 2 levels each – light and surgical plane. The authors focus on the change in avalanche size distributions with respect to the wake, quiet resting state and anesthesia. They find that size distributions change more when reaching the surgical plane in particularly with ketamine. The authors conclude that diverse changes in size distributions indicate significant individual variability and diverse underlying mechanisms.

Comment I

Major critique: 1. The novelty of this study is unclear. Many studies have demonstrated significant and variable changes in cortical dynamics with increasing dosage of anesthetics using a variety of measures. The current study using avalanche size distribution statistics is in line with those general reports. Even the more distinct effect reported here for ketamine is in line with the common observation of burst suppression at surgical plane. It is unclear how the use of avalanche measures contributes significantly to our understanding of the effects of anesthesia on the brain.

We agree that many of the observations regarding the diverse consequences of different anesthetics have been previously identified in the context of understanding the mechanisms of general anesthetics and the neural dynamics associated with reduced consciousness. However to our knowledge this is a novel effort to apply these observations in order to understand avalanche dynamics and the critical brain hypothesis specifically. We believe this is novel and notable as the disparate alterations to neural dynamics previously observed with varying doses and types of anesthetics does not necessarily predict multiple pathways away from criticality. Unique neural dynamics may all be subserved by a common deviation from the critical state. Our results show conclusively that indeed there are multiple pathways that can be induced away from the default critical state associated with quiet wakefulness. We believe this is an important finding furthering our understanding of brain dynamics and the robustness of its intrinsic optimality making it of interest for a broad scientific community, and we have revised our manuscript throughout to emphasize the main contribution of our paper more clearly. We have also made revisions to better contextualize these results with regard to the substantial prior work on neural dynamics and anesthetics.

Comment II

2. The second problem with this study is the claim of different mechanisms underlying the effect of anesthetics on the awake, quiet state. The study would improve significantly if the authors indeed could identify or isolate at least one of those mechanisms. Unfortunately, currently those claims remain rather non-specific, are derived mainly from some of the known receptor targets (which on a global scale are rather difficult to pin down mechanistically), and are suggested to include potential differences in the brain regions involved.

This is an important limitation of our study. The biological mechanisms of anesthetics are notoriously multifactorial, and the work we present here cannot relate the different alterations to the critical state we have discovered to any specific signalling pathway acted upon by the anesthetics we utilized. We have clarified throughout the manuscript that we are simply leveraging the fact that ketamine, isoflurane, and pentobarbital induce an anesthetic state through different signaling pathways to investigate whether this results in different alterations to the critical state. In particular, our previous use of the word "mechanism" may have been unclear. We have done our best to remove the word in any context where it might be misunderstood (in particular in the abstract and introduction).

Comment III

3. There are several technical aspects not considered by the authors that weaken their study further. While the two-photon imaging field has somewhat gravitated to using the raw Df/f-signal, its complexity though makes interpretations of the underlying neuronal dynamics difficult. For example, absent spike deconvolution, the df/f decay is dominated by the decay in the fluoroprobe used reducing temporal resolution. Similarly, estimating f_0 from the average signal is highly unusual and can lead to false reconstructions of the underlying spike trains if non-stationarities are not controlled for. This latter point is particularly important given that the effects of some anesthetics are known to wear off quickly making it difficult to properly control the depths of anesthesia for the duration of the recording.

We agree that these are important points and we have now addressed all of them, including with the help of additional experimental recordings. We would first just clarify that our recording are utilizing single-photon, wide-field imaging, rather than two-photon imaging.

First, we reprocessed four of the QW wakefulness recordings using f_0 calculated from only the first 1000 frames (20 seconds), as opposed to over duration of the recording. As Fig. R3 shows the avalanche statistics are essentially independent of the way f_0 is chosen. As we temporally bandpass filter the pixel-wise fluctuations, this removes any signal decay over the recording duration, making estimates of f0 robust across the recording length. We have updated the methods section to make this explicit.

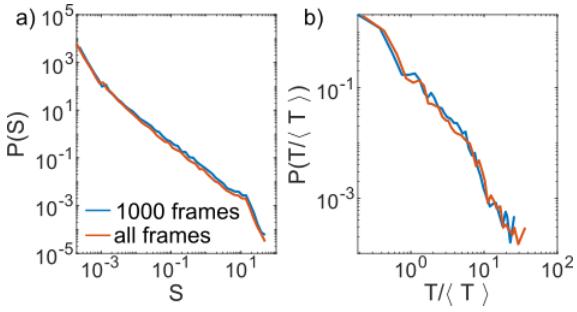

Figure R3: Comparison of avalanche size (a) and duration (b) distributions calculated from a subset of four QW recordings using F_0 calculated from the first 1000 frames, as well as over the whole recording.

Next, we applied a deconvolution method to remove the effects of the slow calcium decay using the code provided in [Pnevmatikakis et al., 2016] with default parameters. We should note that deconvolution methods are not as well established for single-photon wide-field optical imaging techniques as they are for two-photon methods. Despite this, the reconstruction appears rather faithful as Fig. R4a shows for a representative pixel as example of this method. Due to the computational time required for this method, we restricted our analysis to a single hemisphere of QW recordings (previous analysis showed no substantial differences between hemispheres). As each pixel collects bulk fluorescence from a $\sim 40 \mu m^2$ column of cortex, the pixel-wise signal represents the average calcium fluorescence of hundreds of neurons. Consequently the deconvolved signal estimates the multi-unit activity (MUA) of this population of neurons, and we extracted calcium transients or population 'spikes' from this MUA, which are shown in Fig. R4(a,c). We then repeated the avalanche analysis in the same manner as in the manuscript. As Fig. R4d shows, the size distribution of spike-derived avalanches is largely indistinguishable from the original distribution. Most importantly, the decay exponent (which is the relevant aspect for criticality) is statistically indistinguishable between the two cases. Avalanche durations also follow similar statistics as the original (Fig. R4e), albeit with a shift in the mean due to the calcium decay. We have now added a discussion of all these aspects together with Fig. R4 to section **S6** in the supplementary material and highlight in the main paper that regardless of which approach is used to define avalanches, our overall conclusions remain the same.

Lastly, to address the concern that "effects of some anesthetics are known to wear off quickly making it difficult to properly control the depths of anesthesia" we acquired additional recordings for both surgical plane ketamine and isoflurane (this was not done for pentobarbital due to difficulty in requisitioning). These recordings were 15 minutes, twice as long as those used in the manuscript, and were divided into an early and late epoch. Fig. R5 compares the (concatenated) avalanche statistics of the early and late epochs against the avalanche statistics for surgical plane isoflurane (left) and ketamine (right) presented in the paper (QW reference is also shown). Both epochs of isoflurane were consistent with one another and the results presented in the manuscript. Ketamine had some some variability but qualitatively matched the previous results in the manuscript — namely, a departure from criticality marked by the presence of large waves which manifest as excessive large avalanches. We have added a corresponding statement regarding the control of the depths of anesthesia to the manuscript and added Fig. R5 together with the above discussion to the supplemental material.

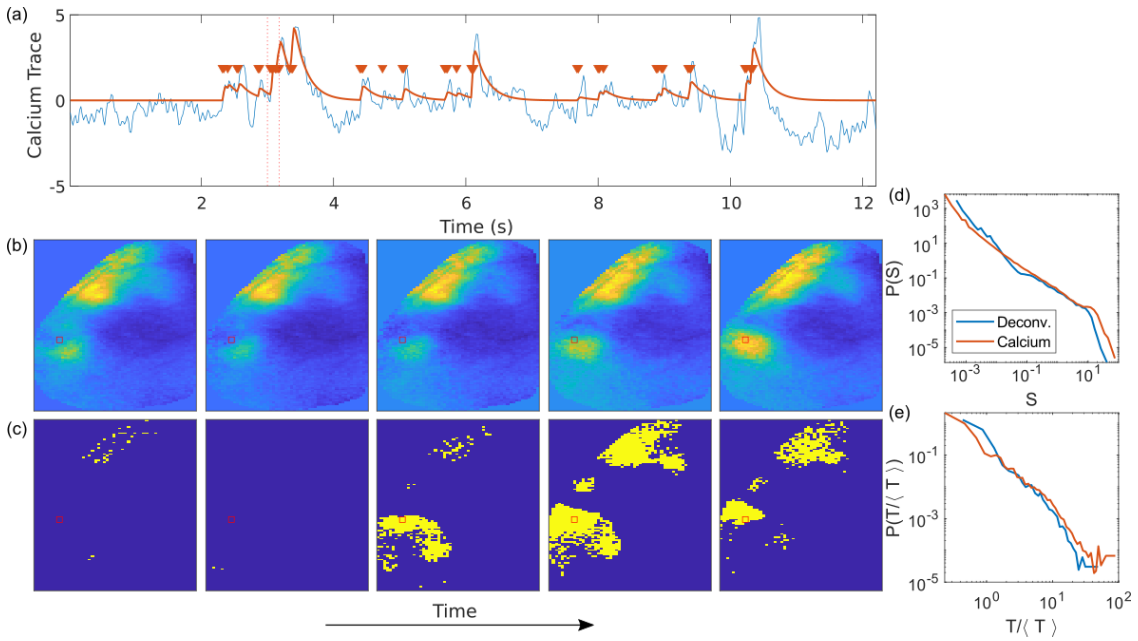

Figure R4: a) An example showing the original calcium signal (blue) along with the deconvolved signal (orange) for a single pixel. Triangles show times at which calcium transients were identified. Dotted-vertical lines show the segment where panels in (b) and (c) are taken from. b) Sample calcium transience signal over the FOV used in the original analysis (only a single hemisphere), with time going from left to right. Red square indicates the pixel corresponding to the trace in a). c) The spikes obtained from the deconvolved calcium signal for the same panels as in b). d) Avalanche sizes distributions obtained from the spikes shown in c) compared with the original calcium signal (concatenated over all QW recordings). e) Same as d) but for avalanche durations, re-scaled by their respective averages.

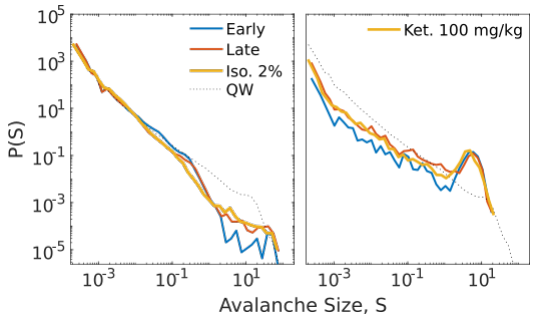

Figure R5: Avalanche statistics for the early and late epochs for surgical plane isoflurane (left) and ketamine (right). Both the yellow and dot-dashed curve in each panel are the same as in manuscript and serve as reference.

Reviewer 3

Curic and coworkers explore the changes in critical behavior of the mouse cortex upon different anesthetic interventions, and show that light anesthesia largely preserves scale-free statistics, while surgical one changes the statistics depending on the anesthetic agent. The results are very interesting since they open new avenues in understanding the relation between chemical perturbations on a neuronal circuit and its dynamical state. The article is clear, very well written and organized. The experiments and analytical tools are convincing, and the results are of interest for a broad scientific community, from physics of complex systems to medicine. However, before publication, I would like the authors to clarify some aspects of their work, as elaborated below.

Comment I

Major concern: 1. The present study is somehow related to the previous work of the group (Yaghoubi et al., Sci Rep 2018), in which they studied criticality in primary neuronal cultures. There, the authors explored different critical exponents and their relation through scaling laws. A relation between critical exponents

derived from the scaling laws provides validation to the analyses, but it was missing in the present study. Can the authors comment on this aspect? For instance, can they measure the duration of avalanches and extract other exponents?

We agree with the reviewer and have done this analysis. Originally our focus was to compare avalanche sizes as these are a) more reliable due to their larger range, and b) more well studied in the literature while avalanche durations are often not discussed (e.g., [Scott et al., 2014]). We agree though that avalanche durations are also important and should be included.

We found that if we concatenated avalanches across all in-group recordings, statistically plausible power-law regions with an exponent α for the durations were difficult to find over extended regions, likely due to variability across the recordings (we will discuss per-recording results in the next paragraph). This variability decreases the fitting range for both size and duration distributions, but in the former the initial range is wide enough so that even when concatenating there is still a substantial fitting range left. Nevertheless, we did find that the α obtained did appear to satisfy the scaling relation, which is calculated as

$$\gamma = \frac{\tau - 1}{\alpha - 1}$$

(τ is the avalanche size exponent and γ is the exponent relating avalanche sizes to durations), as can be seen in Figs. R6(a - c). Here the fitting region for α was chosen using the region obtained from $\Sigma = [S_{min}, S_{max}]$, the fitting range of the avalanche sizes (Fig. R6(a)), which was then used to calculate $[\min_{S\in\Sigma}(\langle T\rangle(S)), \max_{S\in\Sigma}(\langle T\rangle(S))]$, the region obtained from the scaling relation fit with exponent γ (Fig. R6(c)).

On a per-recording basis avalanche duration distributions did have statistically plausible regions of power-law. These ranges were typically just over a decade, but avalanche durations only span about two orders of magnitude to begin with. Using the per-recording estimated avalanche size and duration exponents we tested to see if the scaling relation held. We found that in QW, the relation held for 10 out of 12 QW recordings within uncertainties of one σ. Of the two recordings where the scaling relation only held within three σ, we found that it was because the size exponent τ was very close to 1 where the scaling relation formula becomes problematic and systematic uncertainties are large. In the recordings for which the scaling relation held within the statistical uncertainties of one σ, the size exponent was much closer to 1.2. We are convinced that all these observations together with the results across the concatenated recordings discussed in the previous paragraph provide validation of our analysis of criticality in QW.

Does the scaling relation hold for anesthetics? In the majority of high-dose recordings, the avalanche distributions do not follow a power-law to begin with. For the other ones, on a per-recording basis, the scaling relation was found to be valid for all six 1%, and all awake-like 2% isoflurane recordings within one σ. Whereas the average avalanche size exponent τ was about 1.2 in these cases, we typically encountered significantly smaller values of τ for the other conditions (see Table 1 in the manuscript). As mentioned above, the scaling relation formula becomes problematic for values close to but bigger than 1 and systematic uncertainties are large. For $\tau \leq 1$, it fails altogether (see reviewer 1, comment III), a situation that we encountered a few times for low-dose pentobarbital but very rarely in other cases (see Table 1). Excluding those recordings, the scaling relation held within two σ in five out of the seven remaining ketamine 10 mg/kg recordings. In low-dose pentobarbital, the scaling relation held within one σ in all three recordings, and in one out of three high-dose recordings. In the other two high-dose recordings, it held within two and three σ, respectively.

In summary, the scaling relation typically held within the statistical uncertainties consistent with a critical state. We have updated both the main text and supplementary material to include the analysis of the avalanche durations and the scaling relation summarized above. Consequently, we now report exponents as averages across the per recording estimations in the manuscript. To show the consistency, exponents obtained via concatenation are now provided in the supplementary material.

Comment II

Minor: 2. Experimental data is obtained through fluorescence calcium imaging on a large window of the mouse cortex, of about 9x9 mm2. Each pixel on the images may therefore contain the information of several neurons, many of them acting as "hidden layers" that affect the behavior of the accessed pixels but without direct monitoring. Can the authors explain if they have taken this background activity into account?

This is unfortunately an inherent limitation of wide-field imaging where we are dealing with a coarse grained picture. Each pixel represents the collective calcium trace of hundreds of individual neurons. Indeed, the calcium transients we show in Fig. R4a are not due to any individual neurons but rather represent large population activation. Ongoing advances in mesoscale fluorescence microscopy in awake animals continually improves the fidelity of these recordings, but resolving single neuronal contributions is not possible with our current methodology.

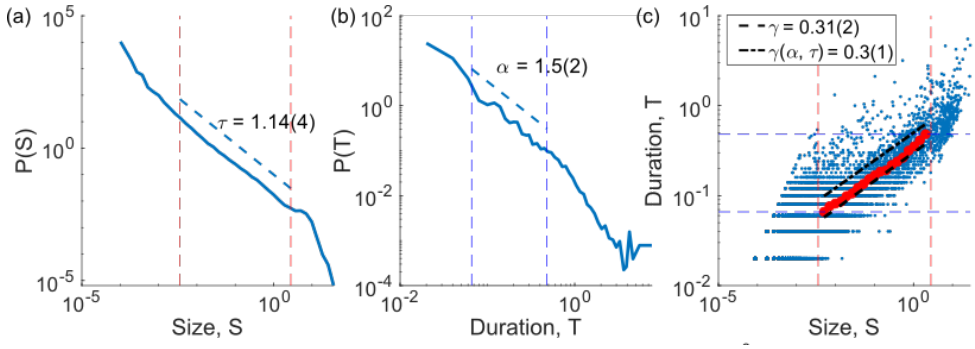

Figure R6: a) Avalanche size distributions for quiet wakefulness concatenated across all recordings. The dashed red lines denote the estimation region. b) Same as a) but for avalanche durations. c) Avalanche durations versus size, with red dots showing the mean avalanche duration for a given size. The red vertical lines are the same as those in a). The estimation region for avalanche duration is obtained from the smallest and largest average duration for the sizes within the red lines. These are the same blue lines in b). The fitted exponent γ is shown as a dashed black line, with the exponent estimated from the scaling relation, $\gamma(\alpha, \tau)$, also shown.

Comment III

3. The authors use the genetically encoded calcium indicator jrGECO1a to monitor activity. Other sensors may differ in kinetics and brightness upon calcium binding and may therefore change the shape of the fluorescence traces. Can the authors comment on possible artifacts associated to the selected fluorescence probe?

The kinetics of protein biosensors for calcium have a large impact on the recorded fluorescence traces, especially the decay kinetics. We utilize jRGECO as this sensor has a favorable combination of brightness (high SNR) and decay time [Dana et al., 2016]. Within red-shifted calcium biosensors, which are favorable in widefield imaging for their reduced hemodynamics contamination, jRGECO represents the best available widely used sensor. More broadly, there are further limitations in estimating neuronal activity, both in the use of a calcium reporter and the camera frame rate (50 Hz). Calcium flucutations are accurate reporters of neuronal depolarization but are themselves slow and lagged relative to voltage fluctuations, and the known firing rates of various neuronal classes (greater than 100Hz, in some cases) imply that a superfast fluorescent voltage biosensor in combination with frame rates in the 200 Hz range may be necessary for artifact free monitoring of neuronal activity. We have included a shortened summary of the above regarding our choice of calcium indicator in the results section of the manuscript.

Comment IV

4. How do the authors understand the fact that a disruption of the excitation-inhibition balance in light anesthetics lead to the roughly the same power law statistics? It is very surprising to me.

The reviewers question is an important one, but we can only speculate. While excitation-inhibition (EI) balance can indeed be disrupted by anesthetics, it is not obvious that this disruption would occur globally and uniformly. If perhaps changes in local EI are compensated elsewhere, then potentially avalanches obtained globally maintain criticality.

To test this idea we utilized an approach by the group of Dr. Klaus Linkenkaer-Hansen which recently published a method to estimate the EI balance from electrophysiological signals [Bruining et al., 2020]. This method has not been established in wide-field optical imaging of calcium, however. As such our analysis here should be regarded as preliminary. Nevertheless, applying this method pixel-wise, we generated averaged EI balance maps shown in Fig. R7. In quiet wakefulness the average of EI values across the cortical surface for the baseline is $EI = 0.90(6)$, close to 1 predicted by the critical point hypothesis. Moreover, the values are distributed relatively homogeneously across the cortical surface (we also show the Gini coefficient as per the reviewers next comment). While light anesthetics do not change the average EI substantially, the values themselves are distributed less homogeneously across the cortical surface. Higher doses of anesthetic can greatly disrupt either the average (as in the case of Iso. 2%) or the localization (e.g., Ket. 100 mg/kg). We also find awake-like recordings (labeled AL) show large differences when compared to their corresponding predominant cases. The analysis including Fig. R7 is now present in the supplementary material and we mention it in the Discussion section of the main paper.

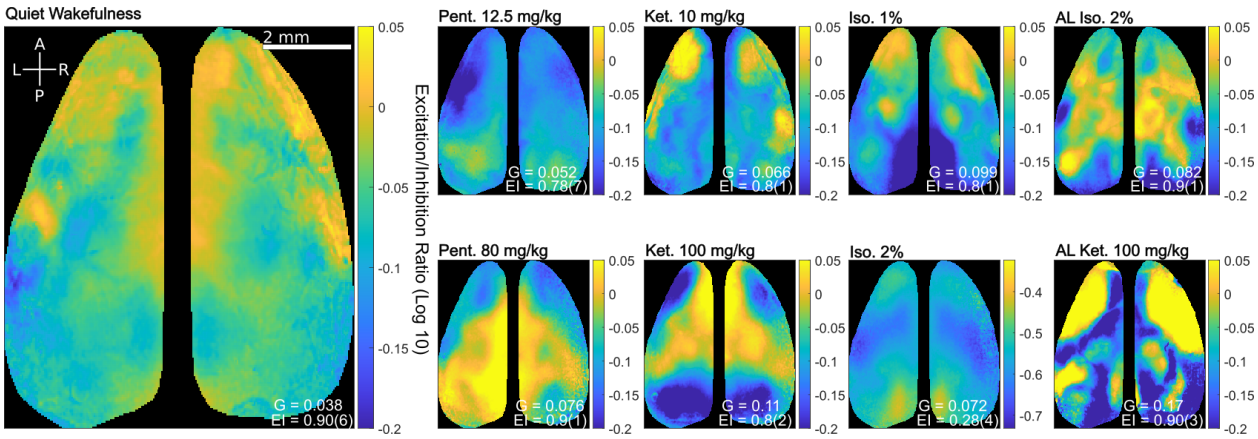

Figure R7: EI balance maps averaged over all animals belonging to a particular category. EI is the average (with standard deviation) over the field of view and G is the Gini coefficient. Color-bars show the (logarithmic) value of each color, with a value of $EI < 1$ means inhibition dominated while $EI > 1$ means excitation dominated.

Comment V

5. I found very interesting the analysis of avalanche initiation. In some cases, initiation is extremely focalized, as for ketamine 100 mg/kg. Maybe the authors could quantify the focalization of initiation by computing the Gini coefficient, as in e.g. Montalà-Flaquer et al., "Rich dynamics and functional organization on topographically designed neuronal networks in vitro", iScience 2022. Additionally, do the authors have a model to understand why this extreme focalization occurs?

We agree with the reviewer that this could be beneficial for interpreting the results and have updated Fig. 5 of the main text with the values (attached also here as Fig. R8). We calculated the Gini coefficient as

$$G = (2n^2\bar{x})^{-1} \sum_{ij}^{n} |x_i - x_j|, \tag{R1}$$

where x_i is the value of the $i-$th pixel, and $\bar{x}$ is the average across all pixels. In the case of ketamine and isoflurane (in particular at high doses) the respective increase and decrease of the Gini coefficient (relative to baseline) matches closely with what one would expect from the avalanche statistics. Under ketamine avalanches were observed to have wave-like patterns initiated preferentially in the anterior regions, thus leading to an increase in G. In isoflurane the calcium dynamics were largely small scale fluctuations uniformly distributed across the cortical surface, decreasing G. Low-dose pentobarbital on the other hand does show a decrease in G, consistent perhaps with the observation of an increased propensity for large avalanches. However, this is not really observed in the higher dose, despite similar avalanche statistics.

We have now expanded the spatial analysis section in the main paper to include the Gini coefficient as discussed above.

Regarding a potential model to understand the focalization — unfortunately no as this would likely rely on detailed models of both the mesoscopic network connectivity, as well as the mechanisms of the anesthetic (which can vary at different dosages as well). There is existing literature on these anesthetics, though the exact mechanisms are not entirely understood. For example, as we mention in the manuscript: *Previous work regarding the effect of sub-anesthetic ketamine has shown the drug acts on interneurons [Behrens et al., 2007, Gerhard et al., 2020] with a global effect on cortical glutamate signalling [McGirr et al., 2017].* This suggests EI balance has some role to play (see also our reply to the comment directly above), though that is not that surprising. However, this mechanism is not thought to hold under surgical-plane ketamine and each anesthetic would need a similar bespoke analysis.

Other Changes

- As per nature communications software transparancy guidelines we have included the clustering algorithm in the supplementary material. The code itself will be available in the associated repository.
- Fig. 3c accidentally had inconsistent number of bins between distributions. This has been fixed.

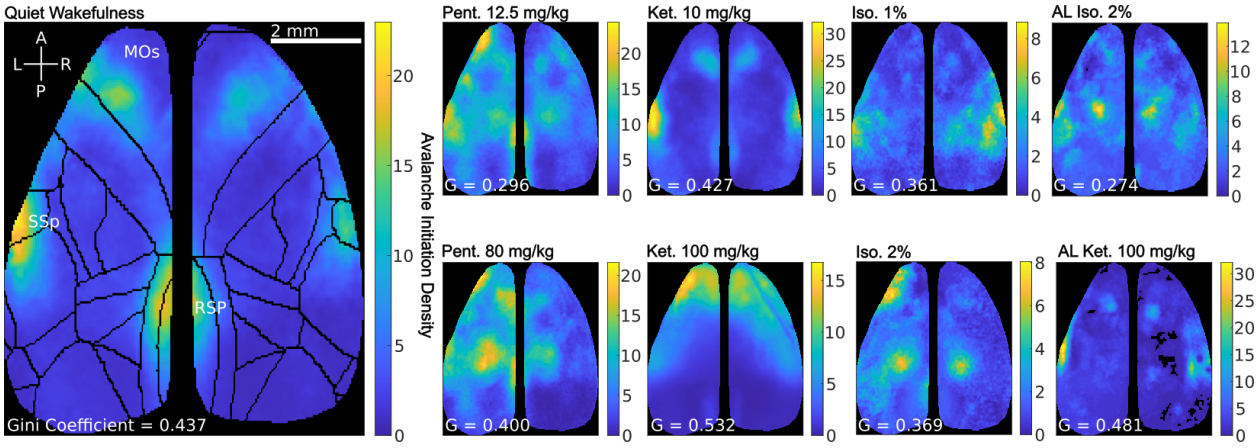

Figure R8: Avalanche initiation maps. Color-bars indicate the number of times a pixel was classified as an initiation site, averaged over all in-group maps, weighted by the number of observed avalanches, with the Gini coefficient G indicated. Awake-like recordings were kept separate as per the previous analysis. The somatomotor (MOs), somatosensory (SSp) and retrosplenial (RSP) areas are also labeled as regions of interest for reference in the first panel. The panel for quiet-wakefulness has the Allen Institute atlas outlined for reference.

References

[Behrens et al., 2007] Behrens, M. M., Ali, S. S., Dao, D. N., Lucero, J., Shekhtman, G., Quick, K. L., and Dugan, L. L. (2007). Ketamine-induced loss of phenotype of fast-spiking interneurons is mediated by nadph-oxidase. *Science*, 318(5856):1645–1647.

[Bruining et al., 2020] Bruining, H., Hardstone, R., Juarez-Martinez, E. L., Sprengers, J., Avramiea, A.-E., Simpraga, S., Houtman, S. J., Poil, S.-S., Dallares, E., Palva, S., et al. (2020). Measurement of excitation-inhibition ratio in autism spectrum disorder using critical brain dynamics. *Scientific reports*, 10(1):9195.

[Carvalho et al., 2021] Carvalho, T. T., Fontenele, A. J., Girardi-Schappo, M., Feliciano, T., Aguiar, L. A., Silva, T. P., De Vasconcelos, N. A., Carelli, P. V., and Copelli, M. (2021). Subsampled directed-percolation models explain scaling relations experimentally observed in the brain. *Frontiers in neural circuits*, page 83.

[Dana et al., 2016] Dana, H., Mohar, B., Sun, Y., Narayan, S., Gordus, A., Hasseman, J. P., Tsegaye, G., Holt, G. T., Hu, A., Walpita, D., Patel, R., Macklin, J. J., Bargmann, C. I., Ahrens, M. B., Schreiter, E. R., Jayaraman, V., Looger, L. L., Svoboda, K., and Kim, D. S. (2016). Sensitive red protein calcium indicators for imaging neural activity. *eLife*, 5:e12727.

[Fontenele et al., 2019] Fontenele, A. J., de Vasconcelos, N. A. P., Feliciano, T., Aguiar, L. A. A., Soares-Cunha, C., Coimbra, B., Dalla Porta, L., Ribeiro, S., Rodrigues, A. J. a., Sousa, N., Carelli, P. V., and Copelli, M. (2019). Criticality between cortical states. *Phys. Rev. Lett.*, 122:208101.

[Funk et al., 2016] Funk, C. M., Honjoh, S., Rodriguez, A. V., Cirelli, C., and Tononi, G. (2016). Local slow waves in superficial layers of primary cortical areas during rem sleep. *Current Biology*, 26(3):396–403.

[Gerhard et al., 2020] Gerhard, D. M., Pothula, S., Liu, R.-J., Wu, M., Li, X.-Y., Girgenti, M. J., Taylor, S. R., Duman, C. H., Delpire, E., Picciotto, M., et al. (2020). Gaba interneurons are the cellular trigger for ketamine's rapid antidepressant actions. *The Journal of clinical investigation*, 130(3):1336–1349.

[Levina et al., 2022] Levina, A., Priesemann, V., and Zierenberg, J. (2022). Tackling the subsampling problem to infer collective properties from limited data. *Nature Reviews Physics*, pages 1–15.

[McGirr et al., 2017] McGirr, A., Berlim, M. T., Bond, D. J., Chan, P. Y., Yatham, L. N., and Lam, R. W. (2017). Adjunctive ketamine in electroconvulsive therapy: updated systematic review and meta-analysis. *The British Journal of Psychiatry*, 210(6):403–407.

[McShane et al., 2012] McShane, B. B., Galante, R. J., Biber, M., Jensen, S. T., Wyner, A. J., and Pack, A. I. (2012). Assessing rem sleep in mice using video data. *Sleep*, 35(3):433–442.

[Pnevmatikakis et al., 2016] Pnevmatikakis, E. A., Soudry, D., Gao, Y., Machado, T. A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W., Ahrens, M., Bruno, R., Jessell, T. M., Peterka, D. S., Yuste, R., and Paninski, L. (2016). Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron*, 89(2):285–299.

[Scott et al., 2014] Scott, G., Fagerholm, E. D., Mutoh, H., Leech, R., Sharp, D. J., Shew, W. L., and Knöpfel, T. (2014). Voltage imaging of waking mouse cortex reveals emergence of critical neuronal dynamics. *Journal of Neuroscience*, 34(50):16611–16620.

[Soltani et al., 2019] Soltani, S., Chauvette, S., Bukhtiyarova, O., Lina, J.-M., Dubé, J., Seigneur, J., Carrier, J., and Timofeev, I. (2019). Sleep–wake cycle in young and older mice. *Frontiers in systems neuroscience*, page 51.

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

The authors considered my concerns. I would like to persuade the authors to include, as Supp. Info, Figure R1 (included in the response to referees). It is my impression that, although the correlation length is not (in the actual implementation) exhibiting noticeable changes, the information is very relevant for the readers interested in further inquiring the topic.

I have no further concerns or suggestions.

Reviewer #2 (Remarks to the Author):

In their revised version, the authors provide additional analysis as requested by Refs. 1&3 and additional data in their response (Fig. R5) to my critique which show similarity in neuronal dynamics measures during the first and 2nd 15 min of anesthesia.

The revised version does not address at all the core problem of this work. I congratulate the authors for showing a thorough understanding of the various 'flavors' in avalanche analysis and potential pitfalls when estimating/quantifying heavy-tail distributions. However, this is not the primary problem with their work and, unfortunately, the additional analysis requested by Refs. 1&3 does miss this issue completely.

As pointed out in my initial critique, the main issue is the lack of measuring their independent variable, the anesthesia level of the animal. In their methods, the authors describe that "For deep-anesthesia experiments, the surgical plane was verified with both tail and foot pinch prior to mesoscale calcium dynamic acquisitions." This approach is insufficient for the work and conclusions drawn in the present manuscript. Particularly, when a main conclusion is that 'wake-like' neural dynamics is found at 'high-level' of anesthetic dosage such as 100 mg ketamine/kg body weight or 2% isoflurane.

There are numerous major issues with their approach and conclusion:
1. The authors do not provide any quantification of the anesthesia level. Instead, they equate their dosage of the drug (injection in the case of ketamine/pentobarbital, volatile anesthetics in the case of isoflurane) with the level of anesthesia in the animal. This is obviously wrong. The level of anesthesia is highly variable from animal to animal and depends on a plethora of physiological parameters (metabolism, heart rate, etc.). Worse, failure to properly administer the drug can not be excluded from this work. Variability in injection site or isoflurane/PB flow rates/mask fitting are typical issues with these approaches. Tail and foot pinch tests are highly qualitative, insufficient to demonstrate a continuous level of anesthesia, and prone to misinterpretation.

Their main finding that, in a few animals, neuronal dynamics is wake-like when given a higher dosage of anesthetics, could simply be due to improper administration of the anesthetics and/or improper assessment of the level of anesthesia or, in general, the state of the animal. The additional data provided does not prove that a proper level of anesthesia was obtained for the small subset of animals appearing 'wake-like' in their neural dynamics.

The work simply does not demonstrate the scientific standard required to carry out these

experiments. Continuous measures of relevant physiological parameters and uptake of the anesthetics using an intracranial micro dialysis probe, pre- and post-anesthetics periods for each animal would be the minimum of experimental documentation required. If foot and tail pinches are being used, responses to these triggers need to be backed up by corresponding physiological measures.

2. An n=1 for ketamine and n = 2/3? for isoflurane is simply too small of a sample size for any statistically significant conclusion. In the worst scenario, the current data suggests a failure of drug administration and/or proper assessment of the physiological state of the animal. In the best case, this small subset could be used as preliminary data to design a more thorough investigation of this issue. Please note that the latter, if thoroughly conducted, has the potential to lead to a revision of animal care and use policies (no one wants to experiment on an animal that is improperly anesthetized for the research proposed). Thus, this is not a small issue but would be very interesting to explore.

3. The authors' findings strongly depend on classifying a 1% level of isoflurane as 'light' and a 2% level of isoflurane as 'deep' anesthesia. Given that dose-response curves typically require a logarithmically spaced exploration of the independent variable, the ad hoc classification assumed by the authors is simply not sufficient to support the conclusions drawn for isoflurane.
Pts. 1 – 3 show that the authors' claim in the Abstract 'surgical plane anesthesia induces multiple dynamical modes, some of which maintain critical avalanche dynamics, while others do not.' This claim is unsubstantiated.

4. The author did not check the state of the animal when given a low level of anesthetics (Lines 553 – 556). Accordingly, classifying this condition as 'light anesthesia' confuses the state of the animal with the dosage of the anesthetic (see also my first point). Accordingly, the conclusion that avalanche dynamics are not sensitive to a light level of anesthesia is incorrect. 'Light anesthesia' needs to be replaced with 'low dosage of anesthetics'.
5. Lack of statistical tests to support claims for low dosage of anesthetics. The first main result, that 'light' anesthesia does not seem to significantly affect certain avalanche measures, is not supported by statistical tests (Fig. 2d). The authors state in the Abstract 'We show that while light anesthesia largely preserves scale-free statistics, …'. The corresponding main text (lines 217 – 222; Fig. 2) uses uncommon verbiage such as 'statistically consistent' or avoids the issue of significance completely. Note that the DKS measure is highly biased for finding differences between distributions, but the test should be on the distribution of DKS measures. Proper non-parametric tests need to be added with corrections for repeat measures to support their claims. Please also note the limits of this claim pointed out in pts. 3 & 4 related to isoflurane and the potential significant effect of 10mg/kg of ketamine.

Reviewer #3 (Remarks to the Author):

The authors addressed all my concerns satisfactorily. I recommend publication of their work.

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

The authors considered my concerns. I would like to persuade the authors to include, as Supp. Info, Figure R1 (included in the response to referees). It is my impression that, although the correlation length is not (in the actual implementation) exhibiting noticeable changes, the information is very relevant for the readers interested in further inquiring the topic.

I have no further concerns or suggestions.

Reply: We agree and have now included the figure and the corresponding discussion in the supplemental materials, see new section S9.

Reviewer #2 (Remarks to the Author):

In their revised version, the authors provide additional analysis as requested by Refs. 1&3 and additional data in their response (Fig. R5) to my critique which show similarity in neuronal dynamics measures during the first and 2nd 15 min of anesthesia.

The revised version does not address at all the core problem of this work. I congratulate the authors for showing a thorough understanding of the various 'flavors' in avalanche analysis and potential pitfalls when estimating/quantifying heavy-tail distributions. However, this is not the primary problem with their work and, unfortunately, the additional analysis requested by Refs. 1&3 does miss this issue completely.

As pointed out in my initial critique, the main issue is the lack of measuring their independent variable, the anesthesia level of the animal. In their methods, the authors describe that "For deep-anesthesia experiments, the surgical plane was verified with both tail and foot pinch prior to mesoscale calcium dynamic acquisitions." This approach is insufficient for the work and conclusions drawn in the present manuscript. Particularly, when a main conclusion is that 'wake-like' neural dynamics is found at 'high-level' of anesthetic dosage such as 100 mg ketamine/kg body weight or 2% isoflurane.

Reply: We thank the reviewer for more clearly expressing what they believe is their fundamental critique. This was not evident in the original round of reviews where this fundamental critique was expressed in the final sentence of major critique #3 after an extended

critique of the calculation of dF/F and arguing for the implementation of deconvolution (which we addressed in our previous revision).

Now knowing what the reviewer believes to be the fundamental flaw with our study, we can provide data to reassure the reviewer that **our primary conclusion that there are multiple pathways away from criticality as revealed by anesthetic agents** is valid. These data are discussed below and shown in the revised supplemental material.

There are numerous major issues with their approach and conclusion:
1. The authors do not provide any quantification of the anesthesia level. Instead, they equate their dosage of the drug (injection in the case of ketamine/pentobarbital, volatile anesthetics in the case of isoflurane) with the level of anesthesia in the animal. This is obviously wrong. The level of anesthesia is highly variable from animal to animal and depends on a plethora of physiological parameters (metabolism, heart rate, etc.). Worse, failure to properly administer the drug can not be excluded from this work. Variability in injection site or isoflurane/PB flow rates/mask fitting are typical issues with these approaches. Tail and foot pinch tests are highly qualitative, insufficient to demonstrate a continuous level of anesthesia, and prone to misinterpretation.

Reply: The reviewer is correct that we did not present physiological data to support the level of anesthesia. In the revised manuscript (see new Fig. 5) and supplemental material (see new section S4), we provide continuous physiological measurement of heart rate to address this critique. The measurements show minimal variability and hence, stationarity in this physiological indicator of anesthetic depth in animals with high-dose anesthesia over the course of each recording.

Their main finding that, in a few animals, neuronal dynamics is wake-like when given a higher dosage of anesthetics, could simply be due to improper administration of the anesthetics and/or improper assessment of the level of anesthesia or, in general, the state of the animal. The additional data provided does not prove that a proper level of anesthesia was obtained for the small subset of animals appearing 'wake-like' in their neural dynamics.

Reply: We respectfully draw the reviewer's attention to the title of our manuscript and to the primary findings, whereby we show that anesthetic agents can induce multiple pathways **away** from default critical dynamics associated with quiet wakefulness. The observation that in some cases of high-dose anesthetics critical dynamics is preserved and resembles that of quiet wakefulness does not affect our main conclusion at all. The animals that displayed awake-like dynamics had similar physiological parameters as others (see our Reply to 1.) and were verified for surgical depth indicating that improper administration of the anesthetics and/or improper assessment of the level of anesthesia is unlikely. In order to not distract from our main conclusion and given the small number of animals that displayed awake-like dynamics, we have now deemphasized the discussion of the latter (including moving Fig. 4 to the supplemental materials, now Fig. S6) and removed any speculation regarding their relevance to anesthesia awareness from the revised manuscript.

The work simply does not demonstrate the scientific standard required to carry out these experiments. Continuous measures of relevant physiological parameters and uptake of the anesthetics using an intracranial micro dialysis probe, pre- and post-anesthetics periods for each animal would be the minimum of experimental documentation required. If foot and tail pinches are being used, responses to these triggers need to be backed up by corresponding physiological measures.

Reply: The reviewer's suggestions are interesting. In this revision we provide continuous physiological measurements (see our Reply to 1.) to augment the foot and tail pinches as requested. With respect to microdialysis quantification of anesthetics, we must highlight that this would compromise the mesoscale field of view and the edema would impact critical dynamics and therefore are not compatible with the question being posed by our study.

2. An n=1 for ketamine and n = 2/3? for isoflurane is simply too small of a sample size for any statistically significant conclusion. In the worst scenario, the current data suggests a failure of drug administration and/or proper assessment of the physiological state of the animal. In the best case, this small subset could be used as preliminary data to design a more thorough investigation of this issue. Please note that the latter, if thoroughly conducted, has the potential to lead to a revision of animal care and use policies (no one wants to experiment on an animal that is improperly anesthetized for the research proposed). Thus, this is not a small issue but would be very interesting to explore.

Reply: We agree that the small number of animals displaying awake-like dynamics makes it difficult to draw firm conclusions. Because of this and the fact that they do not affect our main conclusions, we have deemphasized the discussion of these observations as discussed above. We would like to repeat though that animals that displayed awake-like dynamics had very similar physiological parameters as others (see our Reply to 1.) and were verified for surgical depth. This indicates that improper administration of the anesthetics and/or improper assessment of the level of anesthesia is unlikely.

3. The authors' findings strongly depend on classifying a 1% level of isoflurane as 'light' and a 2% level of isoflurane as 'deep' anesthesia. Given that dose-response curves typically require a logarithmically spaced exploration of the independent variable, the ad hoc classification assumed by the authors is simply not sufficient to support the conclusions drawn for isoflurane.

Reply: We thank the reviewer for this feedback and therefore have revised our terminology and use 'low-dose' and 'high-dose' to reflect this critique.

Pts. 1 – 3 show that the authors' claim in the Abstract 'surgical plane anesthesia induces multiple dynamical modes, some of which maintain critical avalanche dynamics, while others do not.' This claim is unsubstantiated.

Reply: To avoid further misunderstandings, we have revised the statement in the Abstract to "surgical plane anesthesia induces multiple dynamical modes, most of which do not maintain critical avalanche dynamics".

4. The author did not check the state of the animal when given a low level of anesthetics (Lines 553 – 556). Accordingly, classifying this condition as 'light anesthesia' confuses the state of the animal with the dosage of the anesthetic (see also my first point). Accordingly, the conclusion that avalanche dynamics are not sensitive to a light level of anesthesia is incorrect. 'Light anesthesia' needs to be replaced with 'low dosage of anesthetics'.

Reply: To avoid any confusion about the state of the animal and the dosage, we have followed the advice by reviewer #2 and now consistently refer to "low dosage of anesthetics". With respect to the avalanche dynamics, please see our reply to comment 5 directly below.

5. Lack of statistical tests to support claims for low dosage of anesthetics. The first main result, that 'light' anesthesia does not seem to significantly affect certain avalanche measures, is not supported by statistical tests (Fig. 2d). The authors state in the Abstract 'We show that while light anesthesia largely preserves scale-free statistics, …'. The corresponding main text (lines 217 – 222; Fig. 2) uses uncommon verbiage such as 'statistically consistent' or avoids the issue of significance completely. Note that the DKS measure is highly biased for finding differences between distributions, but the test should be on the distribution of DKS measures. Proper non-parametric tests need to be added with corrections for repeat measures to support their claims. Please also note the limits of this claim pointed out in pts. 3 & 4 related to isoflurane and the potential significant effect of 10mg/kg of ketamine.

Reply: To quantify the statistical significance for the low dose anesthetics and beyond, we have now performed a two-sample Wilcoxon Rank Sum test (ranksum, MATLAB) with Bonferroni-Holm correction between QW and the cases for which power-law could be established (all low-dose cases, and high-dose pentobarbital). The findings for both the avalanche size exponents and DKS values confirm our previous statements. Specifically, for the critical exponents tau we found that all low-dose cases were not statistically different from QW at the 95% significance level, while high-dose pentobarbital was statistically different ($p = 0.004$). The Rank Sum test between the distributions of DKS values suggested only the low-dose pentobarbital is statistically different ($p = 0.002$). However, we urge caution when interpreting these statistical results for the DKS values due to the relatively low number of samples in the pentobarbital cases. We updated our manuscript to include the statistical analysis.

Reviewer #3 (Remarks to the Author):

The authors addressed all my concerns satisfactorily. I recommend publication of their work.

Reply: Thank you.

REVIEWERS' COMMENTS

Reviewer #2 (Remarks to the Author):

1. In the new figure 5, the authors now provide heart rate estimates from the raw calcium recordings as a physiological parameter monitoring. As heart rate typical drops with level of anesthesia, this data provides a proxy which supports their claim of stationary recording conditions and/or maintained surgical plane of anesthesia.
2. Results and discussion from the few awake-like animal findings under anesthesia have now been de-emphasized and are now mainly presented in the Suppl Info. Conclusion and statements in the Abstract have been appropriately updated to reflect the uncertainty surrounding these occurrences.
3. Wording has been updated appropriately to remove subjective interpretations of anesthesia levels.
4. Information on statistical significance has now been provided.

Minor:
Line 666: ‘alternative hypothesis’

Replies to Reviewers

We have included verbatim the comments of the reviewer (only one this round), as well as a point by point reply to each comment. Our replies are indicated in *italics*.

Davor C.

Reviewer #2 (Remarks to the Author):

1. In the new figure 5, the authors now provide heart rate estimates from the raw calcium recordings as a physiological parameter monitoring. As heart rate typical drops with level of anesthesia, this data provides a proxy which supports their claim of stationary recording conditions and/or maintained surgical plane of anesthesia.

2. Results and discussion from the few awake-like animal findings under anesthesia have now been de-emphasized and are now mainly presented in the Suppl Info. Conclusion and statements in the Abstract have been appropriately updated to reflect the uncertainty surrounding these occurrences.

3. Wording has been updated appropriately to remove subjective interpretations of anesthesia levels.

4. Information on statistical significance has now been provided.

We thank the reviewer for their comments.

Minor:
Line 666: ‘alternative hypothesis’

We have fixed this, thank you.