Dear Editor,

We thank the Referees for making many detailed and useful suggestions; we sincerely appreciate such a thorough review process. Referee 1 found our example "illustrative and intuitive" for which "the proposed method clearly works well", whereas Referee 2 considered our work "worth bringing to the attention of the readership of Cytometry A" and that "the illustration seems particularly enlightening".

We have addressed all the Referee comments, point by point, below. A Referee comment appear in *blue italics*, followed by our reply in black. In addition, we have made the data and MATLAB code used for this manuscript freely available (details below). We have also provided the MIFlowCyt details as supplements to this submission. We hope you will find our new submission much improved.

We look forward to your reply,

Kind Regards,

Amir Erez

Reply to Associate Editor

Associate Editor Comments:
In addition to reviewers with respect to meeting scientific standards for reporting, it seems this work currently does not meet the MIFlowCyt requirements (not provided, as required by Cytometry A), especially as noted with respect to data availability. Also missing are details of available implementations of their approach, per Cytometry A software and data sharing requirements.

Dear Associate Editor, we now make publicly available all data (including MIFlowCyt checklists) and analysis with the manuscript, organized as described below. We have added to the *Supplemental information* the description below.

- All MATLAB scripts used to generate the plots and analyze the data in
  https://github.com/AmirErez/BimodalLogspaceCytA/tree/master/Scripts

- All data used in the manuscript and accompanying MIFlowCyt checklists, which comes from two datasets:

  1. 20140920-OT1-dynamics - used for Fig. 3 in the manuscript.
  2. FeinermanScience2008 - previously published in Feinerman et al. Science Vol. 321, Issue 5892, pp. 1081-1084 (2008).

  These data with their respective MIFlowCyt checklists are freely available in

  https://github.com/AmirErez/BimodalLogspaceCytA/tree/master/Data

The main manuscript github page can be found in
https://github.com/AmirErez/BimodalLogspaceCytA

Reply to Referee 1

We thank the Referee for considering our example *"illustrative and intuitive"* for which *"the proposed method clearly works well"*, and that the real-world example *"seems to do its job"*. We also thank the Referee for the multiple constructive suggestions. We detail each change following the Referee's comments below:

1. It takes a reader while before it is clear what Q(I) and P(log I) mean in the introduction. They are formally introduced in the introduction, while it would be more coherent to do so in the Materials and Methods section. In addition, I would rewrite the introduction in a general non-formal way in order to make the problem statement clear to the reader.

We appreciate the Referee's suggestion and moved the formal relation between Q(I) and P(log I) to the Methods section. We have also rewritten the opening of the introduction to more gently introduce the reader to the problem.

2. While more advanced data-transformations are stated in the introduction (e.g. Logicle transformation), it might be a good idea to also make at least a comment about them in the discussion of the work and how the method would relate to these transformations, as they are more and more the standard when analyzing flow cytometry data.

We have added a comment about the more advanced transformations in the discussion. Please note that when analyzing the experimental data, in Fig. 3A, we show the data as presented by FlowJo using such a transformation. As is apparent in Fig. 3A, this does not solve the mismatch in the number of peaks.

3. The 'Methods' section needs considerable restructuring. Preferably the toy (or simulated) example is pushed to the first part of the results section, and the methods section contains the derivation and structure of proposed methods, a brief description of Hartigan's dip test and a description of the used datasets.

We have made particular effort to separate out the theoretical (methods) from the experimental (results). As such, we describe the used datasets where they are used, in the "Results - experiments" section. Further, in accordance with MIFlowCyt, the experimental details are summarized in the MIFlowCyt supplements and the data made freely available. We have added a comment to this effect at the beginning of the *Methods* section and details of the data availability in the *Supporting information*.

4. There is no description of the data that is used in this study whatso-

We have addressed this, providing description of the data in accordance with MIFlowCyt. We have made the data and scripts which create the figures publicly available. We provide details in the *Reply to Associate Editor*. Briefly, all data, scripts, and MIFlowCyt details can be found in https://github.com/AmirErez/BimodalLogspaceCytA

5. The link to phase transitions needs more explanation or to be removed (line 53-54).

Removed.

6. The second method, how biological information can be used to properly estimate the mismatch between logarithmic and linear binning is in need of a proper presentation, as the idea seems an interesting one, but is not clearly laid out.

To clarify this method, we have added a paragraph, which begins with "We briefly sketch what follows". (Last paragraph in page 12). The purpose of this section is how, by considering both ERK1 and ppERK levels, one can ameliorate the mismatch between the two representations, resulting in more reliable gates.

7. A recent publication in Cytometry A by Johnsson et al. (2017) show that dip and bandwidth tests perform statistically consistent for unimodality, so it would be preferable to make a comment about that and to incorporate some of the literature in this manuscript. In addition, they propose that for the dip test you should have at least 10,000 cells while results in the manuscript suggest to have at least 100,000 cells. They analyzed the FlowCAP1-dataset, which might be a good addition to this work, as it is well known in the cytometry community.
See: What is a Unimodal Cell Population? Using Statistical Tests as Criteria for Unimodality in Automated Gating and Quality Control. Johnsson et al. (2017), Cytometry part A.

We thank the Referee for pointing out this interesting and recent manuscript which complements well the main thrust of our manuscript. As per Referee 2's suggestions, we have trimmed the detailed discussion of dip statistics (for which there are many, as Referee 2 points out), and focused on the mismatch

instead. Accordingly, a deeper analysis of Johnsson et al. falls outside the focus of our work. Indeed, as Referee 2 points out, drawing conclusions about whether 10,000 or 100,000 cells are sufficient for a certain statistical test, depends on many factors and is not necessary for the message in our article. We cite Johnsson et al. in the manuscript in the appropriate place, together with related work.

8. Some parts of the manuscript, such as the intuitive argument on page 8, should be placed in the discussion.

Moved. We thank the Referee and agree that it is better placed in the discussion.

9. Some labels are missing in figures. Please add those in figure 2 and 6. Also, a number of captions do not address figures in full, see for example the caption under figure 1, in which the left column is not addressed.

We thank the Referee for pointing these out. We have done all changes as requested

- In Fig. 1: Caption modified.

- In Fig.2: $y \longrightarrow$ label added to the horizontal axis; the vertical axis shows 3 different functions, color-coded in the legend.

- In Fig.5: (formerly Fig. 6) we added the appropriate labels.

10. Please have a look at the manuscript organization of Cytometry part A: http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1552-4930/homepage/ForAuthors.html.

We have made the appropriate changes in manuscript organization.

Minor comments:
11. There are quite a number of spelling mistakes and grammatical errors in the manuscript.

We thank the editor for catching these mistakes, and hope that these errors did not obscure his or her first reading of the manuscript. We have taken great care in revising this version and are happy to work with the editor and / or representatives from Cytometry A to correct any remaining errors.

12. Please adhere to the journal's format of abbreviations, such as flow cytometry (FCM).

Fixed

13. Sometimes abbreviations are not defined such as 'RHS' and 'LHS' on page 5 L22 (presumably right hand side and left hand side).

Fixed

Changed.

The Referee refers to Fig. 4, (formerly Fig. 5). We gathered the Referee meant the plots of the marginal distributions $P(I_{ERK1})$ and $P(I_{ppERK})$ drawn in black and red (respectively) to the right of and on top of their joint heatmap. We have modified them, and agree with the Referee that the result is more aesthetic.

Removed.

Fixed.

Reply to Referee 2

We thank the Referee for considering our work *"worth bringing to the attention of the readership of Cytometry A"* and that *"the illustration seems particularly enlightening"*. We also thank the Referee for the multiple constructive suggestions and agree with the prediction that this more terse manuscript conveys the message more readily. We detail each change following the Referee's comments below:

The notion and terminology employed in the manuscript is inconsistent with standard usage in probability and statistics, which is unfortunate as it is likely to limit/inhibit readers. I appreciate that changing this might be time consuming, though I personally think it would be worthwhile. What is labeled as Q(I) and called a distribution in the present paper is instead a probability density function (pdf). Let us say there is a single continuous random variable, I, the intensity of a fluorescence then the pdf would usually labeled $f_I(x)$, with associated cumulative distribution function (cdf) being $F_I(x) = int^x_{-\infty} f_I(y) dy$. What is called P(log I) would then the pdf of the random variable log(I), which would normally be written as $f_{log(I)}(x)$.

We thank the Referee for the suggestion regarding precise notation of the pdfs in the manuscript, and that the Referee acknowledges that such a change of notation is a time consuming. Certainly, for some of the readership of Cytometry A the suggested change of notation $Q(I) \to f_I(x)$ and $P(\log I) \to f_{\log I}(x)$ would improve readability. We fear, however, that a possibly comparable proportion of the readers might mistake f_I and $f_{\log I}$ for having the same *functional form*, just of a different argument, which is manifestly not the case in our example. Indeed, that is precisely the reason we chose P and Q to make this distinction explicit to the reader with a less formal mathematical background. More so, P, instead of Q, was granted to $f_{\log I}$ on merit that the reader is most likely acquainted with a logarithmic representation of the data, and is more habituated to a pdf denoted P. This choice of P for the pdf of $\log I$ is consistent with Cytometry A's list of accepted abbreviations which chooses P for probability.

We hope the Referee understands and accepts the thought process which led to our choice of notation.

That the relationship between the two is

$$f_I(x) = f_{log(I)}(log(x))/x \tag{1}$$

is due to the "change of variable" formula, which can be found in most undergraduate probability texts.

We agree with this statement and indeed we have cited such a book by Nitis Mukhopadhyay in ref 25. The purpose of including this well-known relation is to introduce the reader less versed in probability theory with an essential component for the understanding of the manuscript.

Immediately after Eq. 1 (i.e. current eq. 1), is presented, I think one could pose the question: Q1: when is $f_I(x)$ uni-modal while $e^x f_I(e^x) = f_{log(I)}(x)$ is multi-modal and vice versa? That is the essential element of what follows, and the example provided makes the matter clear.

We have added the question to the text as the Referee suggested.

Later in the paper where data is used to create empirical estimates of f_I and $f_{log(I)}$, exactly how they are made should be stated by way of formulae and more standard notation employed. E.g. $\hat{f}_I^{(n)}$ where n is the sample size. Note that Eq. 1 will not hold for most forms of estimation, which is why the distinction needs to be made.

We have added explicit details of how empirical densities are estimated in the relevant parts of the manuscript.

2) While some of the the argument in the paragraph starting "The difference..." is reasonable, it is not in its entirety and makes matters sound more frightening than they are. For many cell population enrichment purposes, there are gates (see all of hematopoiesis) that are based on logarithmic scales that are demonstrably segregating populations as established by cell sorting and downstream experiments. Consequently, this broad brushstroke dismissal of the meaning of log-scale separation seems at odds with established results.

We have changed the paragraph to address the Referee's concerns.

3) The methods section appears to mix up two things: a) when is Q1 (above) answered in the affirmative; b) when can you reliably assess this when the pdfs in question are not directly related as above, but are empirical estimates? The mixture of the two seems unnecessarily confused. In the present section, you are looking at the modeled pdf where there is no uncertainty so why would one need to confirm the computation with a (known to be unreliable) statistical test on simulated data?
5) What is the relevance of the section on Hartigan's dip test? Tests for bimodality are notoriously challenging / conservative and all are established to be of limited power. There are plenty of references to that effect, so that is not a primary finding of the present paper. Presumably the purpose here is to point that well known challenge to this readership, which I think would be better stated explicitly, particularly as there is more than one dip test out there.

We thank the Referee for pointing out this lack of focus in the methods section. The Referee is correct in understanding that some of our purpose here is to show that determining the number of peaks of experimental data is a difficult task. A second purpose is to refer to recent literature which employs Hartigan's test for a similar experimental situation. This is the relevance and particular choice of Hartigan as an example. Taking into account the Referee's comment, we have shortened the discussion of Hartigan's method in the main text, moving its bulk to the supplementary section for reference. We hope that in this way we clarify the focus of the methods section as per the Referee's suggestion.

4) In that section, the authors assume that $log(I)$ is a Gaussian mixture model. Tedious, unenlightening computations are then displayed much to this reader's displeasure, where only the outcome is necessary to be relayed. This is made more egregious due to the use of S (normally reserved for a survival function, i.e. one minus the cdf, in all stats papers) and F (normally reserved for a cdf). Essentially much of pages 4 and 5 could be deleted on the assumption that a sufficiently interested reader probably knows how to take derivatives and/or work mathematica, and the chase can be cut to.

We accept the Referee's comment that the derivations do not need to be spelled out in full in the main body of the manuscript. Accordingly, we have moved some to the *Supporting information* section, in our attempt to accommodate Cytometry A's readership which is of heterogeneous mathematical background. We are pleased to accommodate the Referee's suggestions that the letters S and F be left alone, and instead have changed $F \rightarrow A$ and $S_{1,3} \rightarrow B_{1,3}$, which we hope the Referee will find less objectionable.

5) appears together with (3) above.

6) The labels of Fig 3 are not self-consistent with the rest of the paper. What does events in $Q(I)$ mean given heretofore $Q(I)$ has been the probability density of the fluorescence at I? One presumes at this stage you are considering estimates of f_I. That being the case, a description of the empirical estimator would seem appropriate. In particular as one might note that (*) will no longer hold for these estimated pdfs (unless the pdf of one of I or log(I) is estimated and the transform used to estimate the other, which would be non-standard).

We have clarified the labels of Fig. 3 and have added a description of how the data are drawn from $Q(I)$.

7) The notation used for a joint pdf, $P_2(...)$, is again not standard, and one would expect $f_{(log(I),log(J)}(x,y)$. Moreover, given the emphasis in the paper on the impact of change of variables on univariate random variables,

We hope the Referee understands that the same logic that guided our choice of P, Q has led us to choose P_2, namely: (i) it is consistent with $P(\log I)$ in that P is used to describe the pdf of the logarithm of intensity; (ii) P_2, instead of P, is used to make explicit the fact that these functions are different, not just of a different argument; (iii) P is consistent with this journal's accepted abbreviations which specify P for probability.

We have added a short comment about the multivariate change of variables, which lies outside the focus of the work. In our analysis, we focus on the marginal distributions which are univariate. We fear a deeper discussion of this will distract the readers from our main message.

8) From that section on, what was used in the present notation as a value a pdf is evaluated at, e.q. $Q(I)$, becomes a measured value. Moreover the conditioning etc. is all written in a non-formal, non-standard way. Such abuse of notation is alleviated with use of standard notation, which would be advised.

We believe that we have used standard notation. We refer the Referee to the book by Koller and Friedman (ref. 32 in the main text). Prof. Koller's book is widely acclaimed in the computational biology field, a highly cited standard reference (according to Google Scholar, 4438 citations since its publication in year 2009). The notation in the conditioning part of the manuscript matches that in Prof. Koller's book. The field of computational biology combines several disciplines with different notation conventions.

A) One can't help but wonder if Q1 could be addressed in generality outside a parametric example.
B) I expect that the authors will not wish to pay this too much heed, and I would not blame them for doing so, but mathematically speaking (rather than practically so) Q1 is not actually a sensible one. To be mathematically precise one would need to pose its equivalent for cdfs. The reason is that pdfs are only defined up to sets of Lebesgue measure 0 and so any unimodal pdf can be replaced with an arbitrarily multimodal one while still giving rise to the same cdf (e.g. spike a normal pdf with the value 100 at all rationales, and you will still have a normal cdf). In other words, a random variable does not have a single pdf and it is actually the case that one is defining equivalence classes through the cdf and picking a representative.

We thank the Referee for both these comments. We also wonder (i) whether the question can be addressed in generality; (ii) is there/how to find a class of representations that always preserve the number of peaks while maintaining the advantages of the logarithmic transform. These are,

however, mathematical questions rather than immediately practical ones; this manuscript attempts to communicate with the cytometry community in a practical, readable manner. Such depth of mathematical analysis required to address the above is outside the scope of this journal. Once the review process is finished, we encourage the Referee to contact us if they so wish, to discuss these questions in more mathematical terms.

C) Cytometry readings are made for many purposes. While the vast majority that I am aware of are on logarithmic scales, and indeed it is befuddling that they look beautifully log-normal, some are linear scales. For example forward and side scatter (FSC and SSC).

We have added a comment regarding this to the discussion. Naturally, if no logarithmic transform is employed, then the possibility for mismatch disappears.

D) The sentence "... similarly to the way Landau theory defines the critical point in second order phase transitions [28]" seems to be an irrelevant remark for the present journal.

Removed.