Response to Anonymous Referee #1 on "Identification of linear response functions from arbitrary perturbation experiments in the presence of noise

Part I. Method development and toy model demonstration"

Guilherme L. Torres Mendonça[1,2], Julia Pongratz[2,3], and Christian H. Reick[2]

[1]International Max Planck Research School on Earth System Modelling, Hamburg, Germany
[2]Max Planck Institute for Meteorology, Hamburg, Germany
[3]Ludwig-Maxmillians-Universität München, Munich, Germany

Correspondence: Guilherme L. Torres Mendonça (guilherme.mendonca@mpimet.mpg.de)

We would like to thank Anonymous Referee #1 for the fast response to our paper, which gives us the opportunity to discuss it more carefully. In view of her/his comments, we believe it is important to first clarify why the paper was framed in the present way and why we are convinced that our method must be introduced at the present level of detail – including a review of some aspects of regularization theory. We thus start with some general remarks clarifying these issues and then proceed to address the referee's comments point by point.

1 General remarks

First, although we do believe that with this paper we give a significant contribution to the ill-posed problems literature, we think it should be made clear that this paper was not framed having in mind the ill-posed problems community as the *main* audience. From the point of view of this community, the main novelty of the paper – as correctly pointed out by the referee – is our new approach to estimate the noise in the data to determine the regularization parameter. As we show in more detail in the point-by-point reply, this novelty is indeed emphasized several times throughout the paper.

Nevertheless, the focus of this paper is on the identification of linear response functions – which from the perspective of the ill-posed problems community can be seen as an application of our new approach to estimate the noise level in a regularization procedure to solve a particular ill-posed problem. Accordingly, the paper was framed mainly for a community of applied scientists (particularly in geosciences) interested in studying physical systems by means of linear response functions. From the point of view of these scientists, our new approach to estimate the noise level is only a single step in the RFI algorithm to identify response functions. The real novelty is the RFI method itself, which makes it possible for the first time to identify these functions taking noisy data from arbitrary perturbation experiments. Therefore, although we do emphasize several times that the main novelty of the method is our noise estimation approach, the main focus of the paper is on the resulting RFI method and how with this method one can identify linear response functions from perturbation experiment data. We believe that the fact that we focus on the resulting RFI method instead of only on the noise estimation procedure might have made it difficult for the referee to "see the forest for all the trees", since she/he evaluated only the methodology and the algorithm, and therefore could see the novelty only from the point of view of the contribution to the ill-posed problems literature.

The fact that this paper was framed mostly for a community interested in the response function approach also explains why we introduce our RFI method starting from a review of details that are well known to the ill-posed problems community. As discussed in the introduction of the paper, current methods in geosciences to identify linear response functions from data do not explicitly account for the ill-posedness of the identification problem. This indicates that the ill-posedness of this problem is not completely clear to at least a large part of the geosciences community interested in the response function approach. Hence, this ill-posedness is already an issue that must be explained to that audience. But most importantly, if one is not aware of the whole ill-posedness issue, one is probably not aware either of details of regularization theory, which is a means to treat that issue. Since for understanding our RFI method one must understand particular aspects of the large corpus of knowledge on regularization (like the discrete Picard condition that is necessary for understanding the noise estimation explained in section 3.4, and "Hansen's observation", needed to understand the additional noise level adjustment in section 3.5), to get our paper understood by our main audience we therefore find it essential to clearly explain those aspects. Further, because our method relies so much on those particular details, we believe that such review is beneficial even for people already working on ill-posed problems that may get interested in our method.

With these general remarks we now proceed to respond the reviewer's comments point by point.

2 Point-by-point reply

***AR#1*: Being a numerical analyst, I am not able to evaluate the application aspect of this manuscript – so I will focus on the methodology and the algorithm. A large part of the manuscript consists of a review of material that is already well described in the references given by the authors. Since this is not a review paper, I wonder why so much space is devoted to review?**

Authors: To answer this comment we refer to our general remarks above. In summary, as pointed out in the introduction of the paper, the main motivation for the design of this method is its practical application to study physical systems. Since scientists interested in applications are not necessarily aware of the nuances of regularization theory, we are convinced that an introduction to the aspects of the theory that are relevant for understanding our method is essential to get those scientists into the boat. In addition, because our method is so deeply rooted in certain details of regularization theory, we believe that explicitly stating those details is beneficial even for initiated readers. Still, we agree that it would make sense to make more clear what parts are review, and where the novelty begins.

***AR#1*: I find that due to this lengthy presentation and all the details, it is difficult "to see the forest for all the trees." Specifically, I find it hard to identify precisely what is the new contribution of this work.**

Authors: We really regret that the referee finds it difficult "to see the forest for all the trees". As stated in the general remarks above, we believe that in part this difficulty is explained by noting that our algorithmic improvements cannot be appreciated without considering also the application aspect of our study, namely the identification of linear response functions, which is our main focus and presents the most relevant novelty to our targeted audience – a method to identify linear response functions taking noisy data from arbitrary perturbation experiments. Still, obviously the main novel idea of this method – which *was* evaluated by the referee – to combine information from a SVD analysis of the perturbation experiment data and from an additional unperturbed data stream to more objectively estimate the noise level and thereby the regularization parameter could not be sufficiently conveyed. Nevertheless, we are surprised by this comment, because we repeatedly emphasize this novel idea, e.g. in the abstract (L5-7 and L11-13), in the introduction (p. 4, L103-105), in section 2 (p. 5, L144-146), when introducing section 3 (p. 6, L152-155, L158-162, L165-166), in section 3.4 (where the idea is in detail explained), in section 4 (where we numerically demonstrate how the noise estimation works), and in section 6 (p. 31-32, L744-748; p. 33, L759-778; p. 34, L779-790).

***AR#1*: The new algorithm is called the "RFI (Response Function Identification) method." This is a very generic name since the goal is, indeed, to solve the deconvolution problem in Eq. (1) – it says nothing about the particular approach taken, and any deconvolution method could go by that name.**

Authors: Only for response problems one can expect to have an additional unperturbed data stream from an independent data source to determine the regularization parameter, which is not the case for general deconvolution problems. In our view, for this reason the choice of the name makes sense instead of simply considering the RFI method as a general deconvolution method. Nevertheless, we are of course open to suggestions from the referee for a more appropriate name.

***AR#1*: The RFI method is summarized in Figure 1, which shows that this is nothing but "plain vanilla" Tikhonov regularization using the discrepancy principle for choosing the regularization parameter. The only novelty seems to be the choice of delta in the discrepancy principle. This could be described much, much shorter.**

Authors: The referee is correct in pointing out that the novelty of the method lies in the estimation of the noise level δ for determining the regularization parameter. We refer to our general remarks and to our first response as to why we believe it is important to introduce also aspects of the method that are not new from the point of view of the ill-posed problems community.

***AR#1*: I honestly do not understand the rationale behind the choice of delta. I can see that delta is the norm of a scaled noise vector, and the scaling depends on an index i_max that is "the last index i before the plateau sigma_i approx 0" [sigma_i being the singular values of the system matrix]. This means that i_max is the number of singular values that are not dominated by rounding errors (and perhaps approximation errors in the discretization). This has nothing to do with the noise in the data, which is the ingredient in the discrepancy principle. That is why I don't understand what is going on here.**

Authors: We fully agree with the reviewer's interpretation of i_{max} that this index gives the number of singular values that are not dominated by rounding and maybe discretization errors. Nevertheless, as we explain in the paper (section 3.4) and emphasize below, by determining i_{max} one can obtain important information about the noise in the data, namely the range of index values for which the SVD components of the data contain for sure only noise.

The rationale for our approach to estimate δ is in detail explained in sections 3.4 and 3.5. Specifically, the explanation for this estimation starts in p. 11, L312. How the approach works numerically is also carefully discussed in the examples of section 4, e.g. in Fig. 5, where we demonstrate how step 3 of the algorithm works (i.e. the scaling of the noise taken from the control experiment, introduced in section 3.4); in Fig. 6, where we show how step 6 works (i.e. the additional noise level adjustment in the presence of a monotonicity constraint, introduced in section 3.5); and in Fig. 8, where we discuss how our noise estimation procedure behaves in the presence of nonlinearities.

But to make our approach even clearer, in the following we visually explain its basic steps, once more using a numerical example. The data for the example were taken from the same toy model experiment used for the analysis of Fig. 5 in the paper. Our step-by-step explanation is shown in Figs. 1–6 below, in a sequence of Picard plots with a text description of each of the steps 1–3 of our RFI algorithm – the basic steps performed to estimate δ. For the particular case where the response function is monotonic, this estimate may be even further improved by the additional step 6 (described in section 3.5 of the paper).

More technically, the approach is the following:

- In the first step, we take a first-order noise estimate $\boldsymbol{\eta}_{ctrl}$ obtained from the additional unperturbed data stream $\boldsymbol{\Delta Y}_{ctrl}$ – the data taken from the *control experiment* (L328–330 of the paper; Fig. 2 in the present reply).

- In the second step, we define i_{max} as the last index before the plateau $\sigma_i \approx 0$ (L337 of the paper; Fig. 3 in the present reply).

- We then note the following: Since by definition for $i > i_{max}$ the singular values are $\sigma_i \approx 0$, by the Picard condition also the projection coefficients of the "clean" data $\boldsymbol{u}_i \bullet \mathbf{A}\boldsymbol{q}$ must have dropped to zero. As a result – as shown by Eq. (25) of the paper –, for $i > i_{max}$ the projection coefficients of the data $\boldsymbol{u}_i \bullet \boldsymbol{\Delta Y}$ are completely dominated by the noise contribution $\boldsymbol{u}_i \bullet \boldsymbol{\eta}$. Therefore, for $i > i_{max}$ the data components $\boldsymbol{u}_i \bullet \boldsymbol{\Delta Y}$ can be taken as an estimate of the respective noise components $\boldsymbol{u}_i \bullet \boldsymbol{\eta}$ (L312–327 of the paper; Fig. 4 in the present reply).

- In the third step, we (i) collect the SVD coefficients $\boldsymbol{u}_i \bullet \boldsymbol{\Delta Y}$ and $\boldsymbol{u}_i \bullet \boldsymbol{\eta}_{ctrl}$ for $i > i_{max}$ in two vectors, and compute their norms z and z_{ctrl} (L338–341 in the paper; Fig. 5 here); and then (ii) scale by z/z_{ctrl} the noise from the control experiment $\boldsymbol{\eta}_{ctrl}$ to obtain $\boldsymbol{\eta}'$ (L341–343 in the paper; Fig. 6 here).

- In this way, the magnitude of the SVD components for $i > i_{max}$ of $\boldsymbol{\eta}'$ matches that of $\boldsymbol{\Delta Y}$, and, because of Eq. (25), also that of $\boldsymbol{\eta}$. If the spectral distribution of $\boldsymbol{\eta}_{ctrl}$ is similar to that of $\boldsymbol{\eta}$ (spectral similarity assumption), then $\boldsymbol{\eta}'$ can be seen as an estimate for $\boldsymbol{\eta}$ (L344–L348 of the paper; also Fig. 6 here).

- With the resulting noise estimate $\boldsymbol{\eta}'$ we finally set the noise level $\delta := ||\boldsymbol{\eta}'||$, which is then used in the discrepancy method (L349-354 of the paper).

- If monotonicity of the response function should be accounted for, the resulting estimate of δ is further adjusted in an iterative way until the constraint is enforced (section 3.5 of the paper).

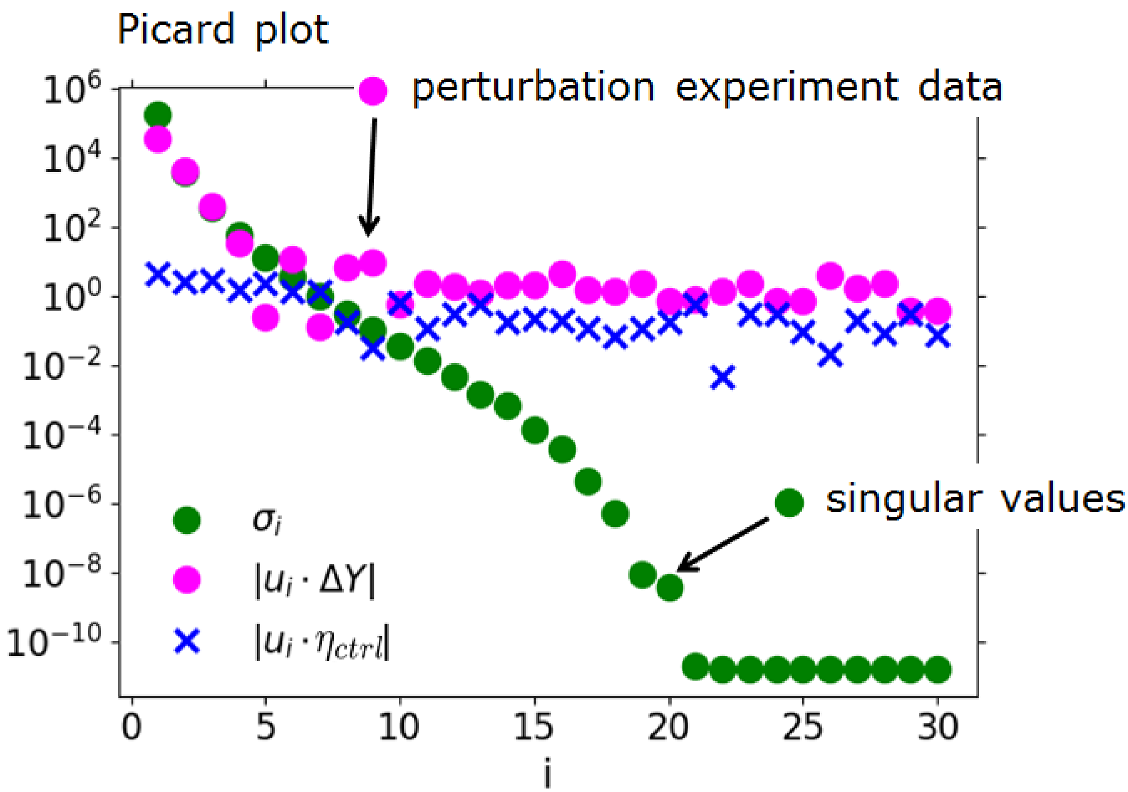

Figure 1. Picard plot for exemplary data taken from a simulation with our toy model (described in section 4.1 of the paper). The data vector $\boldsymbol{\Delta Y}$ from the perturbation experiment are in magenta. In green are the singular values of the matrix $\mathbf{A}$, and in blue the noise from the control experiment $\boldsymbol{\eta}_{ctrl}$.

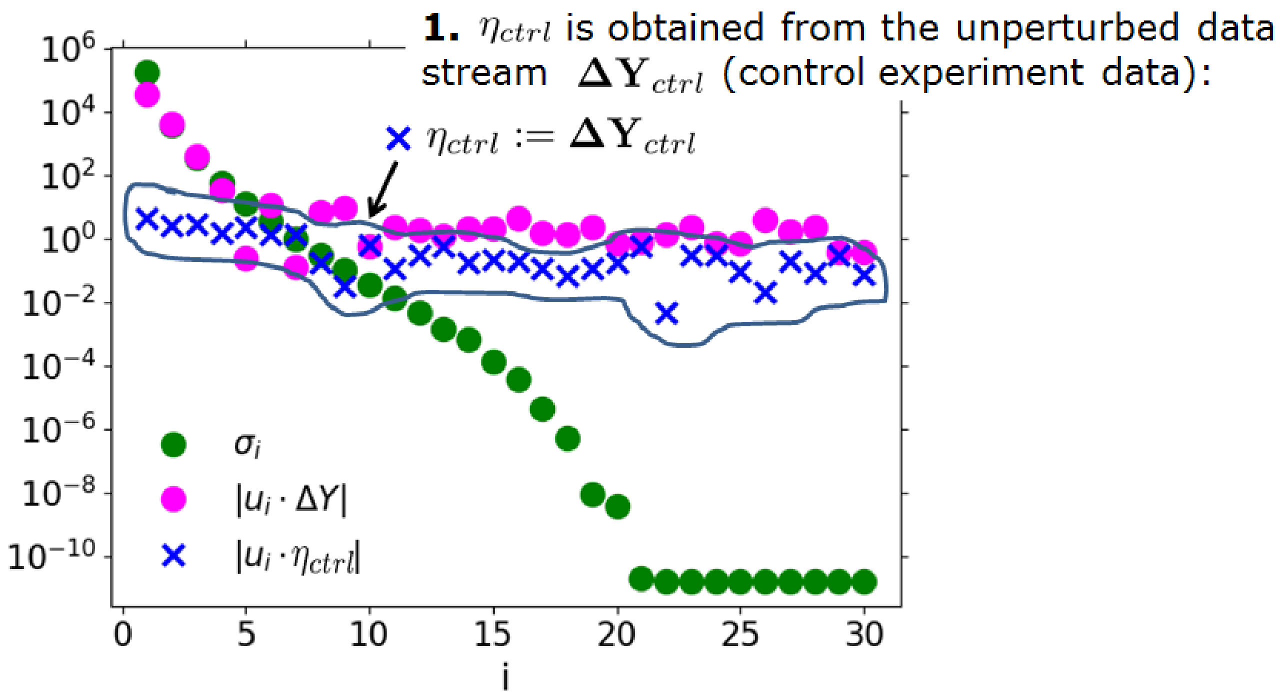

Figure 2. First step of the RFI algorithm: The noise from the control experiment η_{ctrl} is taken from control experiment data ΔY_{ctrl}.

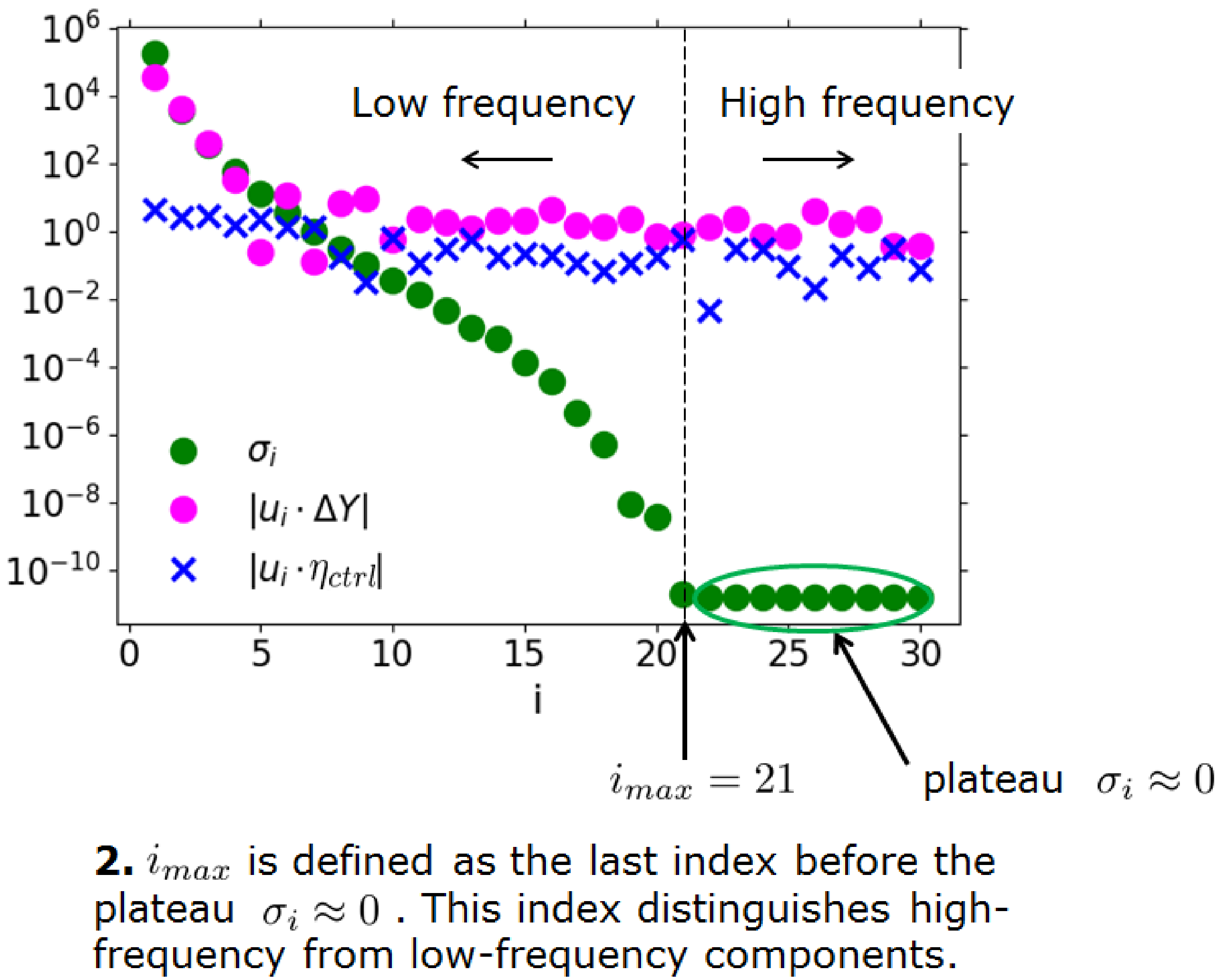

Figure 3. Second step of the RFI algorithm: The index i_{max} is determined as the last index i before the plateau $\sigma_i \approx 0$.

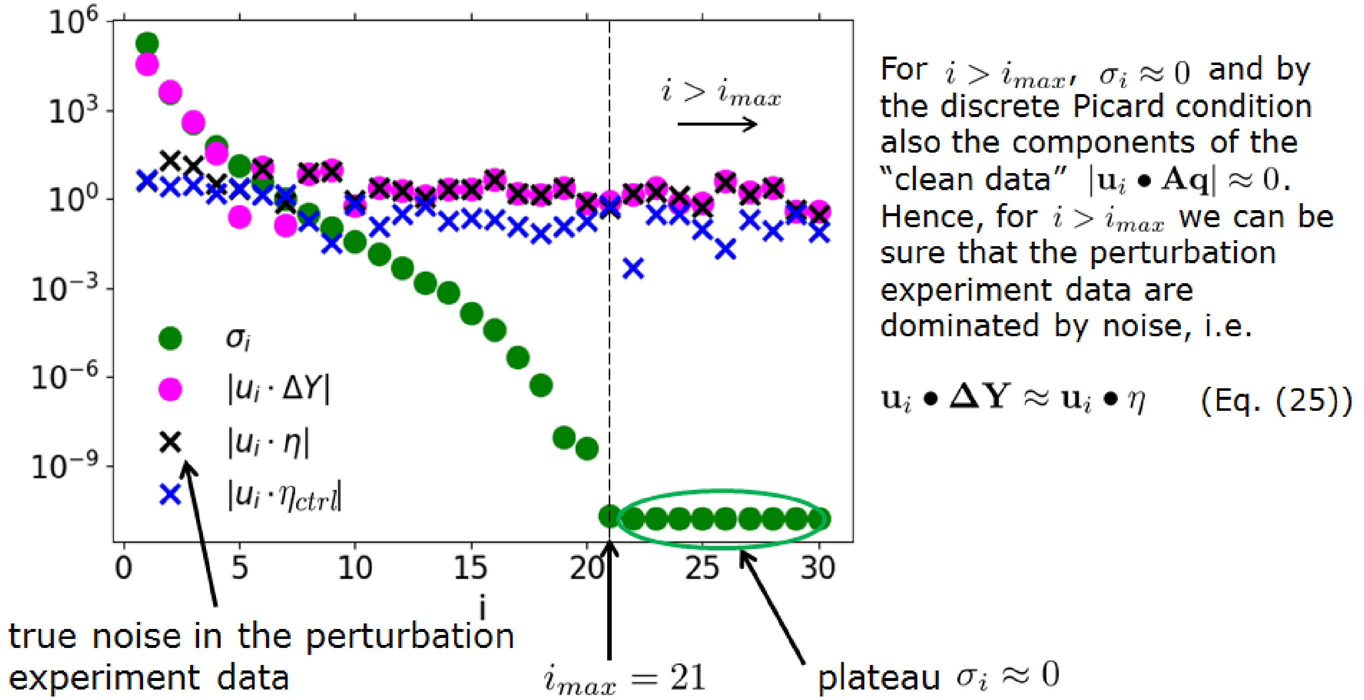

Figure 4. Explanation for the relevance of i_{max} to determine the noise in the data: For $i > i_{max}$, the data are dominated by noise. Therefore, for $i > i_{max}$, the data components $\boldsymbol{u}_i \bullet \boldsymbol{\Delta Y}$ can be understood an estimate of the noise components $\boldsymbol{u}_i \bullet \boldsymbol{\eta}$.

3. Define z as the norm of the high-frequency components of the perturbation experiment data, which by Eq. (25) are also the high-frequency components of the noise in the data. Define z_{ctrl} as the norm of the high-frequency noise components from the control experiment.

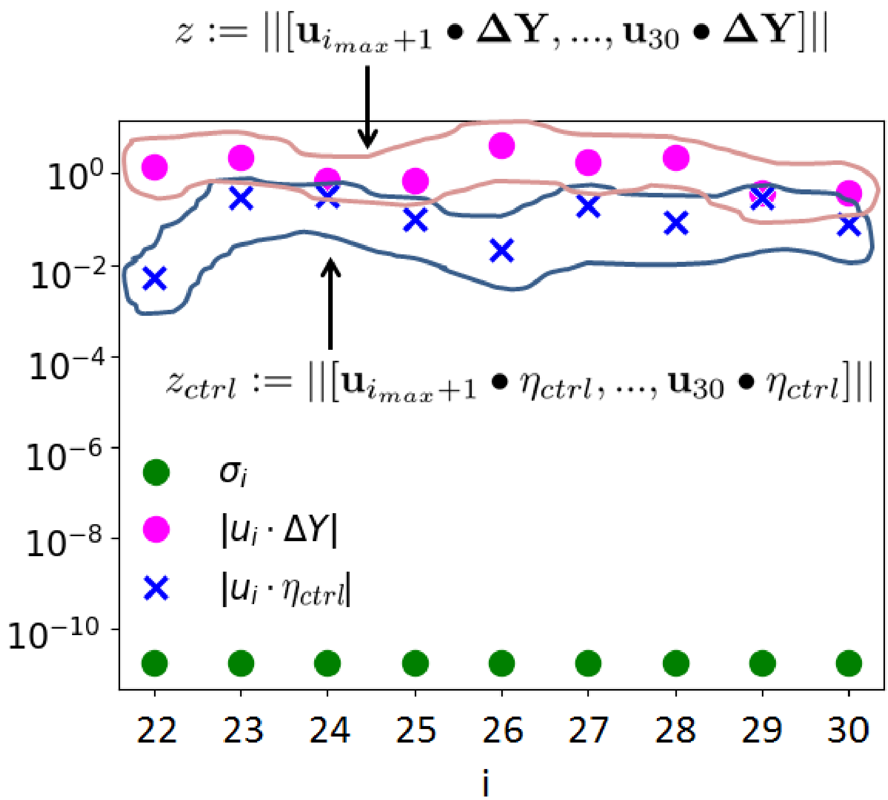

Figure 5. Third step of the RFI algorithm: Define z as the norm of the high-frequency components of the data $\boldsymbol{\Delta Y}$, which by Eq. (25) of our paper are also the high-frequency components of the noise $\boldsymbol{\eta}$. In addition, define z_{ctrl} as the norm of the high-frequency components of the noise from the control experiment $\boldsymbol{\eta}_{ctrl}$.

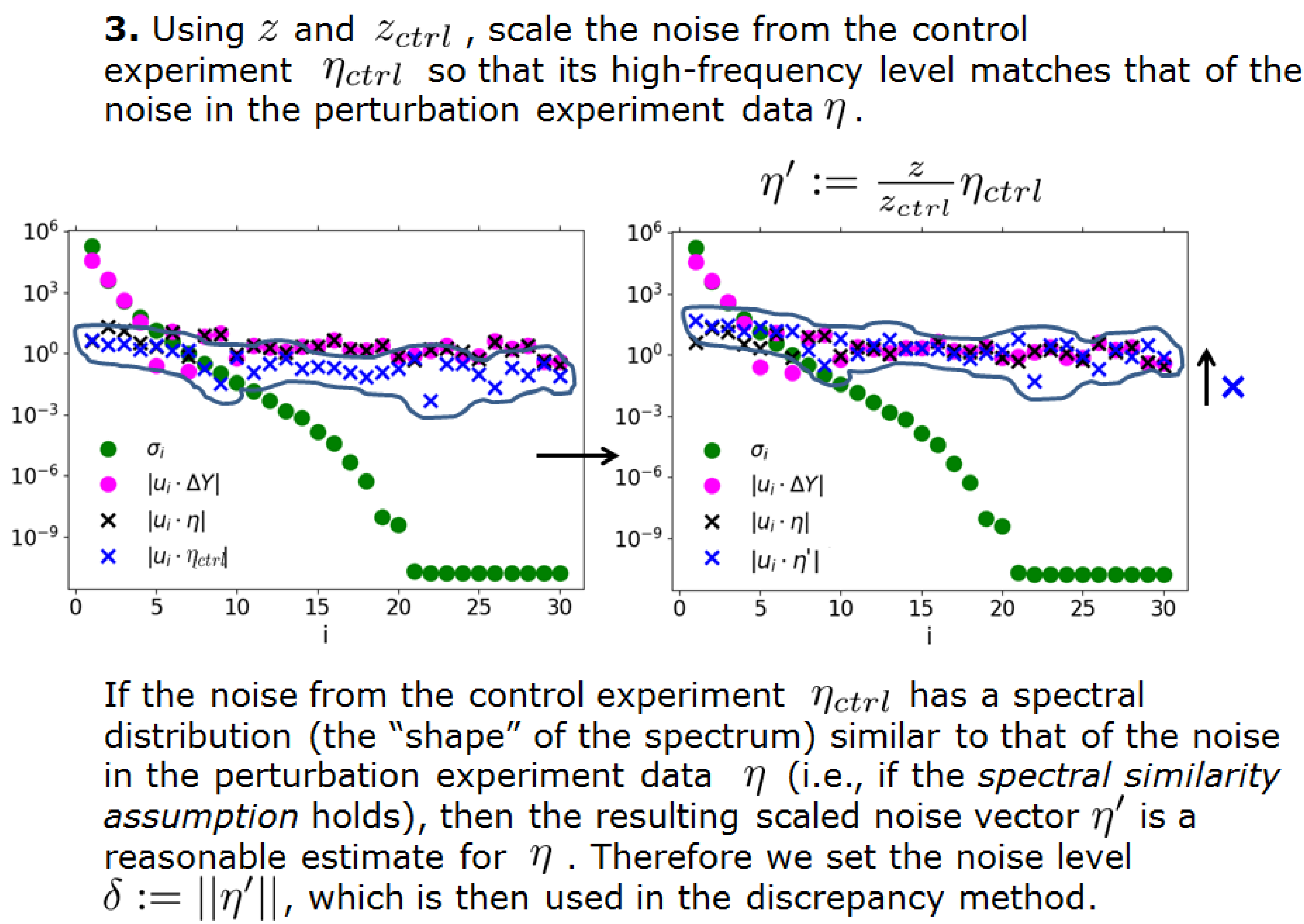

Figure 6. Third step of the RFI algorithm (continuing): scale by z/z_{ctrl} the noise from the control experiment $\boldsymbol{\eta}_{ctrl}$ to obtain $\boldsymbol{\eta}'$. If the spectral distribution of $\boldsymbol{\eta}_{ctrl}$ is similar to that of $\boldsymbol{\eta}$ (spectral similarity assumption), then $\boldsymbol{\eta}'$ can be seen as an estimate for $\boldsymbol{\eta}$.

***AR#1*: Due to the excessive amount of review material, the minimal amount of novelty, and the failure to motivate and explain how delta is computed, I recommend rejection of the manuscript. A much shorter and precise manuscript might be considered for publication.**

Authors: We still think that our study represents an important methodological advancement not only for the geosciences community interested in the response function approach – whose viewpoint is the main focus of the manuscript –, but also for scientists interested in ill-posed problems. As demonstrated by the large number of studies employing response functions in geosciences, especially in recent years – see introduction, in particular L35-41 –, these functions represent a powerful tool of increasing importance for the geosciences community. But as also discussed in the introduction, currently no method is available in the field to identify response functions taking data from any arbitrary type of perturbation experiment. Further, even in the case where a tailored perturbation experiment for identifying these functions is available, noise in the data usually hinders a reliable identification. This often makes it necessary to perform many experiments to obtain a better signal-to-noise ratio. These two difficulties (the need for a special perturbation experiment and for sufficiently "clean" data) thus severely restrict the applicability of the response function approach. Hence, by presenting a method to identify response functions from arbitrary perturbation experiments in the presence of noise, we believe that this paper gives a relevant contribution to the field that allows for a much wider applicability of the response function approach.

From the point of view of scientists interested in ill-posed problems, we are convinced that our method represents also an important advancement. The reason is that we present for the problem of response functions identification an approach to estimate the noise in the data on more objective grounds. As is known from the literature – and also discussed in our paper (see section 6, starting from L759) –, to obtain a regularized solution that converges to the "true" solution of the problem for decreasing noise level, regularization methods need to account for the noise level (Bakushinskii, 1984). But in practice this noise level is rarely known with accuracy: In fact, several studies investigate the typical situation where one has only a guess of the noise level at hand (e.g., Raus, 1992; Hämarik and Raus, 2006; Hämarik et al., 2011, 2012). With our approach, this noise level can in principle be more accurately estimated by using information from a SVD analysis of the data and from an additional unperturbed data stream: This is numerically demonstrated not only by the results of the application of our method to toy model simulations shown in the present paper, but also by the results of its application to a real problem in Part II. Further, although our method was designed for the identification of linear response functions, as discussed in section 6 (starting from L779) it may in principle find application in solving also other types of linear ill-posed problems.

But to make these advancements more obvious, it may indeed be useful to convey better to the reader what parts of our paper are reviewing standard knowledge in numerical analysis, and what is new.

With best regards,

Guilherme L. Torres Mendonça, Julia Pongratz and Christian H. Reick

References

Bakushinskii, A.: Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion, USSR Computational Mathematics and Mathematical Physics, 24, 181–182, 1984.

Hämarik, U. and Raus, T.: On the choice of the regularization parameter in ill-posed problems with approximately given noise level of data., Journal of Inverse & Ill-Posed Problems, 14, 2006.

Hämarik, U., Palm, R., and Raus, T.: Comparison of parameter choices in regularization algorithms in case of different information about noise level, Calcolo, 48, 47–59, 2011.

Hämarik, U., Palm, R., and Raus, T.: A family of rules for parameter choice in Tikhonov regularization of ill-posed problems with inexact noise level, Journal of Computational and Applied Mathematics, 236, 2146–2157, 2012.

Raus, T.: About regularization parameter choice in case of approximately given error bounds of data, Acta et Comment. Univ. Tartuensis, 937, 77–89, 1992.