Manuscript No.: amt-2021-334

Title: Regularized inversion of aerosol hygroscopic growth factor probability density function: Application to humidity-controlled fast integrated mobility spectrometer measurements

We thank the anonymous referees for their valuable and constructive comments/suggestions on our manuscript. We have revised the manuscript accordingly and please find our point-to-point responses below.

Comments by Anonymous Referee #2:

General Comments:

The authors apply different inversion methods to invert the aerosol GR-PDF from the measured signals from synthetic HFIMS signals. They found that for the few test cases, Markowski-Towmey's method generally outperforms other methods. By doing this, they convincingly improved the data inversion of HFIMS data and promisingly HTDMA data, which were mainly based on predefined size distributions or least square methods. This well-written manuscript is easy to follow. I recommend it to be published in Atmospheric Measurement Techniques, However, a major revision is necessary to convincingly demonstrate that the data inversion of HFIMS (and HTDMA) is improved. I feel that the authors are too optimistic about the representativity of their limited synthetic data on real laboratory experiments and atmospheric measurements. Further, this manuscript will have a broader impact on the community if its outcomes (e.g., inversion codes) can be readily used for HTDMA measurements. My detailed comments are given below.

<u>Responses</u>: We thank the reviewer for the constructive suggestions and comments. Point-to-point responses to comments and questions are detailed below. Following the reviewer's suggestions, we examined the impact of additional Gaussian noises on the performance of inversion methods. We also applied the inversion algorithms to ambient measurement data and compared the performances. Additional analyses were also carried out to elucidate why Twomey's method is statistically better than the Tikhonov regularization methods. The new results and discussions are now included in the revised manuscript.

Detailed Comments:

1) More tests and/or discussions are needed to provide supports for the argument that Towmey's method outperforms other tested inversion methods. The three test cases are perhaps sufficient to show that Towmey's method is better than least square methods because the least square methods are notorious for solving ill-conditioned problems. However, the reason why Towmey's method is better than the Tikhonov regularization methods needs more clarification and/or data to support.

<u>Responses</u>: We thank the reviewer for this comment. To elucidate why Twomey's method is better than the Tikhonov regularization methods, we compare the results of inversions based on Twomey's method and 1st order Tikhonov regularization with the regularization parameter derived using three different approaches. The first approach (i.e., the L-curve approach, Hansen and O'Leary, 1993), derives the λ by seeking a trade-off between minimizing the residual term and minimizing the regularization term (i.e., roughness of the solution). The 2nd approach is based on the Hanke-Raus rule, which selects the λ

value that minimizes the λ-dependent residual term $\frac{1}{\lambda}\left\|\mathbf{Mc}^{\boldsymbol{Tik}}(\lambda)-\mathbf{R}\right\|_{\mathbf{2}}$ (Hanke and Raus, 1996; Sipkens et al., 2020). In the 3rd approach, the value of λ is optimized by comparing the inverted GF-PDF with the correct solution, i.e., minimizing the Euclidean distance between the inverted and the pre-defined GF-PDF. Therefore, inversion based on λ derived using the 3rd approach represents the best possible performance of the Tikhonov regularization. It is worth noting that the 3rd approach is not possible for real measurements, as the true GF-PDF is unknown. As shown in Fig. 1, the Tikhonov regularized solution strongly depends on the regularization parameter λ. The Tikhonov regularization with the optimized λ (derived using approach #3) provides the most accurate solution (i.e., lowest GF-PDF error) as expected, and outperforms Twomey's method. However, when λ derived using the L-curve approach or the Hanke-Raus rule is used, GF-PDF inverted using 1st order Tikhonov regularization generally has a larger error (i.e., γ^2) than that inverted using Twomey's method. The above comparisons indicate that while Tikhonov regularization can outperform Twomey's method in theory, the optimal regularization parameter λ cannot be obtained reliably using existing methods in practice, leading to inferior performance than Twomey's method. For example, the L-curve approach does not work well if the curvature of the L-curve is negative everywhere, and in such scenario, the leftmost point (i.e., with smaller λ) on the L-curve is taken as the corner (Hansen, 1994), leading to insufficient regularizations of the solution (Naseri et al., 2021). On the other hand, the Hanke-Raus rule often chooses a much larger λ compared with the optimal value, which results in over-smoothed solutions potentially with even larger errors. We also carried out similar comparisons of Twomey's method with 0th and 2nd order Tikhonov regularizations with λ values derived using the three different approaches, and the results are consistent. We have included the above comparison and discussion in the revised manuscript (line 335-370).

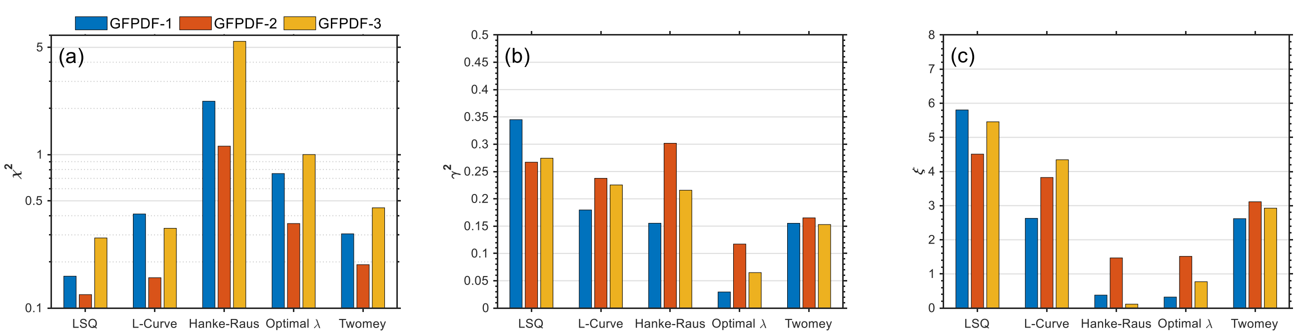

Figure 1. The reconstruction residual, $\boldsymbol{\chi^2}$ (a), the GF-PDF error, $\boldsymbol{\gamma^2}$ (b), and the smoothness, $\boldsymbol{\xi}$ (c) of GF-PDF inverted using LSQ, 1st order Tikhonov regularization with the regularization parameter derived from three different approaches (L-curve, Hanke-Raus rule, and optimized $\boldsymbol{\lambda}$), and Twomey's method. The colors correspond to the pre-defined GF-PDFs with one mode (blue), two modes (orange), and three modes (yellow). The results are averages based on inversions of 500 sets of synthetic HFIMS data for each of three pre-defined GF-PDFs.

2) The authors need to show the performance of Towmey's method with at least one dataset from either laboratory experiments or atmospheric measurements. Estimating the measurement uncertainties with only the counting uncertainties typically *underestimates the total uncertainties. Despite this, I am not concerned about the applicability of Towmey's to real datasets and its better performance of the than least square methods.*

<u>Responses</u>: Following the reviewer's suggestion, we apply the nonparametric inversion methods to ambient HFIMS measurements, and the results are compared in Fig. 2. The HFIMS responses reconstructed from GF-PDF inverted using unregularized LSQ, Tikhonov, and Twomey's methods generally match the measurement (black circle) well. The GF-PDF at 85% RH for ambient 35 nm particles consist of a smaller less-hygroscopic mode and a larger more-hygroscopic mode. As expected, the HFIMS response reconstructed from LSQ inverted GF-PDF has the minimum deviation from the actual measurement whereas the GF-PDF exhibits more oscillations near the tail of the second mode. These oscillations create a small third mode that is absent from the smoother GF-PDFs inverted using regularized methods (i.e., Tikhonov and Twomey's methods). GF-PDF inverted using Twomey's method and 0^{th} Tikhonov clearly distinguish the two growth factor modes. In comparison, the two modes become more overlapped in GF-PDF inverted using 1^{st} and 2^{nd} Tikhonov regularization, due to additional and possibly excessive regularization.

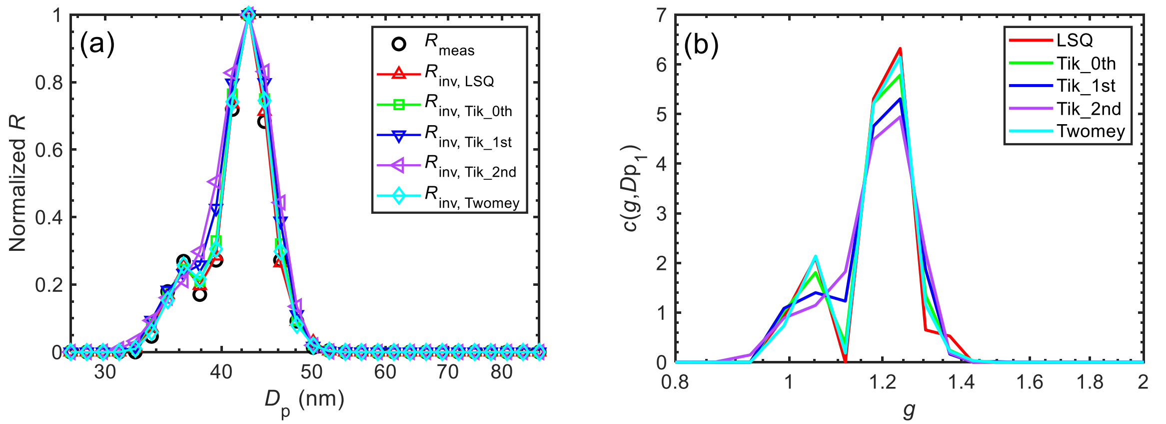

Figure 2. (**a**) Comparison between the HFIMS measured response (black circle) and the responses (marked lines) reconstructed from GF-PDF derived using different methods for 35 nm ambient aerosol at 85% RH. (**b**) Inverted GF-PDFs using different methods.

We also examined the statistics of the reconstruction residual and the smoothness of GF-PDF inverted from 3-day HFIMS measurements using the listed nonparametric methods. Among all nonparametric inversion methods, unregularized LSQ leads to the lowest reconstruction residual but the worst smoothness (Fig. 3). As regularizations are introduced in the Tikhonov algorithms, the inverted GF-PDFs become smoother at the expense of increased reconstruction residuals. The Tikhonov regularized solutions strongly depend on the regularization parameter λ. In this study, the value of λ has been derived using three approaches, including (1) the L-curve, (2) the Hanke-Raus rule, and (3) comparison of inverted GF-PDF with the true solution. Note the 3^{rd} approach (i.e., comparison of inverted GF-PDF with the true solution) is not possible for ambient measurements. Inversions of synthetic data show that the L-curve approach generally underestimates the regularization parameter (Fig. 5 in the manuscript), resulting in insufficiently regularized solutions (Naseri et al., 2021). For the 3-day ambient measurements, when λ is derived using the L-curve approach, the reconstruction residuals for the GF-PDF inverted using Tikhonov algorithms are very close to those of the unregularized LSQ, consistent with underestimated λ values (Fig. 3a and d). In contrast, Tikhonov regularizations with λ value determined using the Hanke-Raus rule tend to over-smooth solutions due to overestimated λ values, resulting in significantly increased errors in reconstructed HFIMS measurements (Fig. 3b and e). The 3-day ambient measurements are also inverted using Tikhonov algorithms with an empirical λ value of 0.03 (Fig. 3c and f), which corresponds to the mean value of optimized λ values (i.e., derived using the 3rd approach) for the synthetic HFIMS data. The inverted GF-PDF shows improved smoothness compared to the solution from the LSQ method, without introducing excessive reconstruction errors. While the empirical λ value appears to work quite well for the 3-day measurements, using this fixed regularization parameter may not be appropriate for other ambient measurements. For Twomey's method, both the reconstruction residual and the smoothness are between those based on the 0th order and 1st order Tikhonov regularizations with the empirical regularization parameter (λ = 0.03), suggesting an appropriate trade-off between the GF-PDF smoothness and the fidelity in reproducing the HFIMS measurements. Note that the statistics of the GF-PDF error cannot be derived as the actual GF-PDF of ambient aerosols are unknown. As a result, it is difficult to draw a definite conclusion regarding which method has the best performance in retrieving the GF-PDF based on the ambient measurements. The above results and discussion have been added in Section S5 of the revised supplemental information.

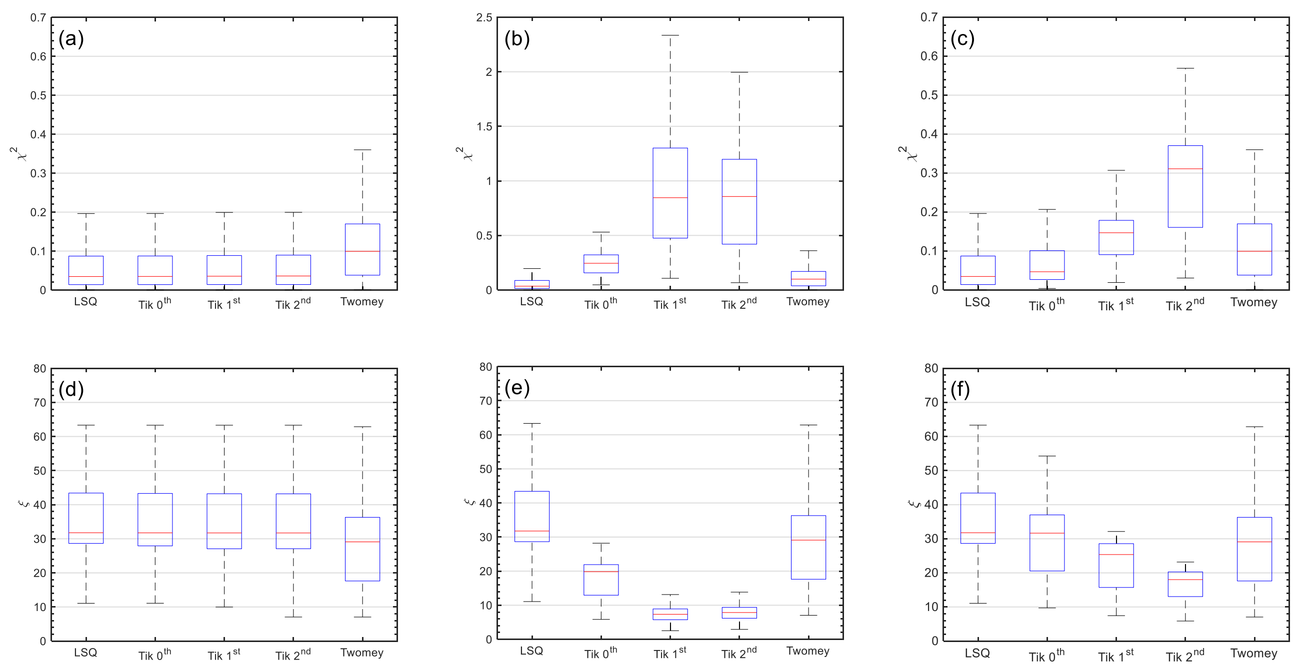

Figure 3. Comparison of reconstruction residual, χ^2 (**a, b, c**) and the degree of smoothing, ξ (**d, e, f**) of inverted GF-PDFs using different inversion methods (i.e., LSQ, Tikhonov of 0, 1, 2-th order, and Twomey's method), based on 3-day HFIMS measurements of ambient

aerosols of 35 nm at five different RH levels (20%, 40%, 60%, 75%, and 85%). The Tikhonov regularization parameters are derived using the L-curve approach (**a, d**), the Hanke-Raus rule (**b, e**), and an empirical value of 0.03 (**c, f**), respectively.

We agree that estimating the measurement uncertainties with only the counting uncertainties underestimates the total uncertainties, which also include the system noises (e.g., variations of the sample flow). In addition to noise due to counting statistics, we also included additional Gaussian noise (e.g., due to variations of sample flow rate) ranging from 1% to 10% in generating the synthetic HFIMS data and examined the impact of the additional noise on inverted GF-PDFs. Figure 4 shows that Gaussian noises up to 10% have negligible impact on the reconstruction residual (χ^2), the error (γ^2), and the smoothness (ξ) of GF-PDFs inverted using Twomey's method. Similarly, the impact is also negligible for GF-PDF inverted using unweighted LSQ and 0th, 1st, and 2nd order Tikhonov regularizations (not shown). The negligible impacts indicate that the noise of typical HFIMS measurements is dominated by counting statistics. The above results and discussion are detailed in a new section (Section 3.2) titled "Effect of measurement uncertainties" in the revised manuscript (line 253-299).

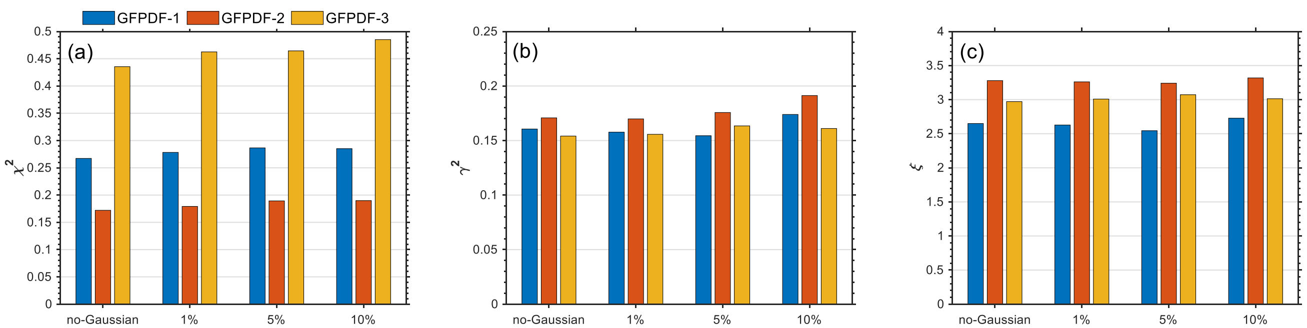

Figure 4. Comparison of reconstruction residual, χ^2 (**a**), the GF-PDF error, γ^2 (**b**), and the degree of smoothing, ξ (**c**) inverted GF-PDFs using Twomey's methods with additive Gaussian noises of different levels (i.e., none, 1%, 5%, and 10%). Colors correspond to the pre-defined GF-PDFs with one mode (blue), two modes (orange), and three modes (yellow). The results are averages based on inversions of 500 sets of synthetic HFIMS data for each of three pre-defined GF-PDFs.

3) Lines 30 - lines 115. The working principles of HFIMS are well summarized. However, they can also be removed or shortened to make space for more tests and discussion, as long as the inversion problem (e.g., Eq. 4) is clearly proposed.

Responses: We thank the reviewer for the comment. We did not include specific details of the HFIMS (i.e., the general principle and instrument setup), which has been presented in previous studies (Pinterich et al., 2017; Zhang et al., 2021). We summarized existing parametric inversion methods for both the HTDMA and HFIMS (i.e., ML and PL least-squares fitting). In Sect. 2.1, we present the mathematical derivation of both forward and inverse models, which are the key components of this study.

4) Line 116, Eq. 4. Please consider adding an error term (É) to Eq. 4 and other related equations to emphasize that the main challenge of data inversion is to deal will the uncertainties. The least-square methods are supposed to work pretty well if there is no error in the inversion problem as presented in Eq. 4.

Responses: We have clarified this in the revised manuscript (line 114-127).

5) Line 125. "The integration can be written as......". I recommend replacing "written as" with "approximated by". Discretizing a continuous distribution is also a step of inversion and there are inversion algorithms using improved discretization methods (e.g., Hagen and Alofs, doi.org/10.1080/02786828308958650).

Responses: We thank the reviewer for this constructive suggestion. "approximated by" is much more rigorous. We made the change in the revised manuscript (line 122).

6) As far as I am concerned, Towmey's method does not mathematically guarantee convergence of the inversion results. Optimized adjusting factor(s) are usually needed to guarantee that convergence without a great sacrifice of the computational expense. As a result, the convergence of Towmey's method for one dataset (e.g., synthetic data) does not guarantee its convergence for other datasets (e.g., laboratory experiments and atmospheric measurements). I recommend the authors address this very briefly in the main text. Considering broader applications of the inversion methods to HTDMA studies, I recommend the authors address this very briefly in the main text.

Responses: We thank the reviewer for this comment. To conduct a comprehensive test, we have carried out comprehensive tests of inversion algorithms using both synthetic HFIMS data based on representative GF-PDFs and ambient HFIMS measurements (i.e., as described in the revised SI), and the results indicate that Twomey's method performs well for both datasets. While the convergence of Twomey's method is achieved for the datasets used in this study, it is possible that Twomey's method may not achieve a converged solution on rare occasions, especially for measurements with very poor counting statistics. We have clarified this in the revised manuscript (line 178-179).

References

Hanke, M., and Raus, T.: A General Heuristic for Choosing the Regularization Parameter in Ill-Posed Problems, SIAM Journal on Scientific Computing, 17, 956-972, 10.1137/0917062, 1996.

Hansen, P. C., and O'Leary, D. P.: The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems, SIAM Journal on Scientific Computing, 14, 1487-1503, 10.1137/0914086, 1993.

Hansen, P. C.: REGULARIZATION TOOLS: A Matlab package for analysis and solution of discrete ill-posed problems, Numerical Algorithms, 6, 1-35, 10.1007/BF02149761, 1994.

Naseri, A., Sipkens, T. A., Rogak, S. N., and Olfert, J. S.: An improved inversion method for determining two-dimensional mass distributions of non-refractory materials on refractory black carbon, Aerosol Science and Technology, 55, 104-118, 10.1080/02786826.2020.1825615, 2021.

Pinterich, T., Spielman, S. R., Wang, Y., Hering, S. V., and Wang, J.: A humidity-controlled fast integrated mobility spectrometer (HFIMS) for rapid measurements of particle hygroscopic growth, Atmos. Meas. Tech., 10, 4915-4925, 10.5194/amt-10-4915-2017, 2017.

Sipkens, T. A., Olfert, J. S., and Rogak, S. N.: Inversion methods to determine two-dimensional aerosol mass-mobility distributions: A critical comparison of established methods, Journal of Aerosol Science, 140, 105484, https://doi.org/10.1016/j.jaerosci.2019.105484, 2020.

Zhang, J., Spielman, S., Wang, Y., Zheng, G., Gong, X., Hering, S., and Wang, J.: Rapid measurement of RH-dependent aerosol hygroscopic growth using a humidity-controlled fast integrated mobility spectrometer (HFIMS), Atmos. Meas. Tech., 14, 5625-5635, 10.5194/amt-14-5625-2021, 2021.