The paper *Mathematical foundation of Capon's method for planetary magnetic field analysis* provides the underlying formalism for applying Capon's method to planetary magnetic fields and illustrates it with simulated data relevant for BepiColombo mission. While this is a valuable contribution to the field, a number of points need to be addressed before publication.

1. *Extension of Capon's method to planetary magnetic fields.* As indicated in the first and last paras of Section 3, L73–75 and L202–206, the paper generalizes the Capon method, previously used for the analysis of wave data. This is a major result that could be emphasized better, perhaps in a separate discussion section. This section could include a closer analysis of the case in the paper as compared to the wave case, by referring to Motschmann et al. (1996).

The discussion section could also detail the key principle(s) underlying the method, like maximum likelihood / minimum variance, along the line of Narita (2019). The divergence free feature of the magnetic field could be discussed on top, as in Motschmann et al. (1996).

Related:

- L166-168: I am not sure I understand the text here, even though I could essentially follow Eqs. 30 to 41. Eventually, the underlying model makes the main contribution to the data (e.g., in the test application of Section 5). I guess this 'minimal contribution' has rather the meaning of Motschmann et al. (1996), where the filter w absorbs all the energy not associated with k (here not associated with the parametrized field) and leaves the part related to k undistorted (here the part related to the parametrized field). Same issue at L256.
- L200-201: Is this expression derived by Narita (2019)? Or could be derived by further processing of the maximum likelihood estimator? (e.g., in the suggested discussion section?)

2. *Illustration & Validation.* In view of upcoming BepiColombo data, the authors chose to illustrate the method with simulated observations of Mercury magnetic field. While this is certainly helpful to prepare BepiColombo, I wonder if it is also the best test bed for the method. Earth magnetic field is known much better, at various altitudes – such that the weight of the external field and its influence on the results could be analyzed too. Including an example at the Earth, or at least a brief discussion of this validation possibility, would be more than welcomed.

Regarding the test exercise of Section 5, Table 1 shows that the largest errors are associated with g_2^1 and, to some extent, with h_2^1. Is this by chance, or related to some systematics?

3. *Technicalities.* Considering the target audience of the journal, different to a good extent from the signal processing community, the mathematical language of the paper may prevent the optimal transmission of the message. Additional explanations may help, inserted in the text or collected in an Appendix – when detailing the math would perturb the flow too much:

- L69-71: Please clarify this sentence, possibly including an example.
- L93-97: The intuitive introduction of the filter matrix w via Eq. 8 s a bit confusing, since eventually the non-parametrized (external) part of B does not show up in the g_C formula, Eq. 41.
- Eqs. 9 and 10 fall pretty much out of the blue. The use of M and P becomes clear later, but some clarification would be good already at this point.
- L106-110: Please detail why the determinant vanishes (even though it may look straight), how does statistical average prevent this, how is statistical average achieved.
- L127: Please indicate also the second order moments.
- Eq. 21: Please explain why 2<Hg> o <v> and not <Hg> o <v> + <v> o <Hg> (given that, in general, the external product does not commute).
- Eq. 27 and L154: Please explain why this is not enough to uniquely determine w.
- L155-158: Feels confuse. As long as the filter matrix truncates the non-parametrized part, it is not clear why its contribution to the data matters, neither how 'this yields the following procedure'.
- L191: Why is tr P a convex function?

- Eqs. 42 and 43: Please detail what is meant by ‘input’ and ‘output’. Regarding Eq. 43, is there an equation analogous to Eq. 23, to clarify the meaning of ‘signal’ and ‘noise’ also for output?
- Eq. 44: Please explain why this ratio is dominated by 1/trace.
- L266-267: Please provide a brief demonstration.
- L267-269: Please explain briefly what is this about.
- L278: How is the ‘compromise’ quantified?
- L278-282: This is quite opaque for those not familiar with signal processing and in particular with these techniques.

4. *Others*

- L17-18: What is non-ideal orbits?
- L18: simulated Mercury magnetic field data
- L54: ‘closing the void’ => ‘covering the range’ ?
- Eq. 54 is identical to Eq. 41.
- L325: in => at
- L352: In principle, one could analyze also the external field, if some model is adopted.