

Jury members' comments on the private PhD defense of Gustavo Quintana-Carapia

General comments

- Discuss more clearly the aspects of scalability of the methods and their limitations in terms of model order, size of data sets (e.g., related to the inversion of LS matrices).

Lyudmila Mihaylova

- You develop results for uncertainty quantification; what type of uncertainties can your approach deal with and which levels of noise can you consider? Elaborate on SNR levels; I'm also very interested in scalability, so how big the matrices can be.
- What recommendation up to scale/order can method work with. What are the limitations? It would be good to have more precise discussion about these aspects in dissertation: what do you mean by large scalability: 100, 10^4 , ...? A more detailed discussion is desirable.

Stephane Chretien

- About trying to find a good transformation from LS to EIV model. I am wondering if you can think of other more accurate starting points instead of LS. Closer to the problem you want to study, but still preserving computational lightness? The closer you start from objective, better the Taylor expansion would be.. Something to investigate as well...
- About robustness: is it an issue in practical problems with heavy tailed noise, outliers,... Perhaps the use of median of means plugged into your approach might produce interesting alternatives for outliers or heavy tailed data. But this is an intricate thing to deal with. Might be interesting to ponder this in future.
- On deep learning: could this be an alternative?

Guillaume Mercere

- Chapter 2, Section 2.2.1. Can you give me some classic solutions/algorithms instead of LS in general? (Say I'm not familiar with EIV or TLS?) Usually with LS problems: multiply with K^\top and compute pseudoinverse. Do you know another solution where they don't use K on the left-hand-side but something else? To get rid of noise by using a matrix on the left-hand-side, is related to EIV. There are some recursive solutions for this. It could be interesting to compare with such IV solutions.
- In Chapter 3: at beginning you say to use linear systems theory to model a sensor. Can you discuss this assumption: is it true in practice? Can you guarantee it behaves like an LTI system?

- Equation 3.1: an LTI system, described with SS representation. You only put output noise on this representation. People sometimes also add process noise. Why? How to modify your approach if there is process noise? Is it still white in this case?
- Switch to 4.2, simulation results: first part dedicated to K , randomly generated; not sure if still possible, but in all your results it would be nice to have somewhere explicitly $K = \dots$ (even if randomly generated). You use only one realization also for matrices A , B , C , D in the text, so it can be added explicitly.
- To show comparison between structured and unstructured case, use unstructured results for structured case; then use unstructured results for structured case and cover all the possibilities. Now you compare curves related to two different simulations.

Interesting question, I want to try it this evening

- On the impact of selected order on results. You were talking about a related bias/variance trade off. So, do you know in the literature some solutions to avoid overfitting? Can you cite something? Can you use something in your case to test for a good order of system?
- Do you know of cross-validation: split data into two parts or several parts; some for estimation, some for validation? Have you tried this? Can this be used here to select good orders? Better results? Influence of system order on bias/variance; if I only use one dataset it is complicated and usually you can increase complexity to mimic a data set, but to avoid overfitting (too high order) use a test/validation data set. with this i should see what is the best value... Have you tried this?
- Concerning the ML approach in simulations, you say that it takes 30 seconds to complete. What is done exactly in these 30 seconds; how many samples, realizations,... I am a bit surprised you require 30 sec for 2 parameters.

Nikos Deligiannis

- When solving algorithms. Do you opt for closed-form solutions, or do you use optimization based algorithms? How to solve LS problems? Do you solve them using pseudo-inverse or do you do gradient descent?
- Today there is an overwhelming amount of data sensed and communicated; How would you apply your method in a scenario where dimensionality increases significantly? This relates to the questions of previous jury members; what is the implication of using numerical versus closed form solutions?
- What if the $K^T K$ matrix is not invertible? What if you have fewer samples than unknowns; what can you do? Solutions that assume structure like in compressed sensing, regularization,... can you use them, what would be the impact of that?
- You mention that you improve complexity with your approach, but I missed a complexity analysis. Did you do this? Anything in processing time, flops,...?
- You mentioned deep learning could be applied; could you show how? An issue is that neural nets are trained in high SNR data, but applied in engineering with lower SNR. This clogs performance; any ideas? Is this a limitation?

- You mention that you consider Kalman filters; have you seen approaches where they write Deep Learning in form of Kalman Filter; Could this be applicable?

Philippe Dreesen

- Your main assumption is that a sensor outputs a steady state value that is proportional to the to-be-measured quantity. This means that you assume that the sensor is a linear system. What if this assumption is not met, and the sensor has a saturation for high values? Can you still use your methods?
- Related to earlier questions about compressed sensing. There are some recent results on the connections between sparsity and low-rank Hankel matrix completion and approximation. Say that you have an issue with missing data, e.g., from sensor outages. Would it be possible to plug in a procedure that does structured matrix completion and approximation into your methods?

Rik Pintelon

- Returning to a question by Philippe Dreesen on missing data. If you can deal with missing data, you can deal with saturation in sensors: you just remove the saturations. What is the rank of the Hankel matrix? Is it full rank or not? The question is if matrix completion can help? It works in low-rank approximations; not if the matrix is of full rank.

Detailed comments provided by jury members

Guillaume Mercere

- It is important for me to know why these new developments are generated. That is the reason why I suggest adding sections or paragraphs dedicated to
 - the reasons why, in the industry, we need such a new technique. As I said during the defense, do you have a list of sensors for which waiting for the steady regime for getting the measurement is not possible? Why do you need a fast algorithm? What are the reasons why industry people need fast algorithms? In the end, it seems that, for you, it is more important to have a fast biased result than an unbiased estimate. Why? What do you mean by fast?
- When you generate different simulation examples, explain why the new ones are interesting or, more importantly, what they bring w.r.t. the former simulations results. This is the case for instance when I compare Section 4.2.2 and Section 5.1.

Nikos Deligiannis

- Comments related to the structure of the thesis:
 - In the introduction chapter it is best to describe a set of applications and settings that motivate your research and proposed solutions. This can make your thesis more accessible. Furthermore, you can add a subsection that describes in a high-level manner

the remainder and structure of the thesis (describing briefly the contributions of each chapter). You can also mention your contributions there but in a less technical way since the problem statement only comes in the Chapter preliminaries.

- Further comments:
 - Position your work with regard to the state of the art as well as regarding solutions based on low-rank matrix completion and compressed sensing.