REVIEWER'S OPINION OF MASTER'S THESIS

Thesis name:	Knowledge transfer for linear systems with bounded noise
Author's name:	Bc. Eva Lainová
Type of thesis:	Master's thesis
Faculty/Institute:	Faculty of Nuclear Sciences and Physical Engineering
Department:	Department of Mathematics
Thesis reviewer:	Ing. Milan Papež, Ph.D.
Reviewer's department:	Artificial Intelligence Center
Address:	Charles Square 13, Prague 2
Email:	papezmil@fel.cvut.cz

The main objective of the thesis is to study and experimentally validate a fully probabilistic design (FPD)-based approach for knowledge transfer between Bayesian filters, each estimating latent variables of a linear state-space model with uniform noise disturbances. Transfer learning is a very attractive and vibrant research subject—especially in the community of machine learning and artificial intelligence. However, its deployment in the field of statistical signal processing has begun only recently, making the objectives of the thesis very relevant. The difficulty of the assignment is adequate for the Master's degree programme: "Applied Mathematical Stochastic Methods".

Chapters 1, 2, 3 and 4 are the introduction into the thesis. The author's contributions are presented in Chapters 5 and 6. More concretely, Chapter 1 provides a brief description of linear state-space models, various approaches for constructing support of the uniform distribution and definition of FPD. Chapter 2 introduces (i) an isolated Bayesian filter for a linear state-space model with uniform noise, discussing the approximation necessary to keep the time-update step closed under the uniform distribution; and (ii) the instantiation of FPD-based knowledge transfer for a pair of uniformly-modeled filters, mentioning the connection between the intersection of the (source and target) filters' supports and different transfer regimes. To establish a baseline for the transfer learning, Chapters 3 and 4 describe centralized and decentralized information fusion methods, respectively. Chapter 5 experimentally validates the aforementioned methods on a second-order dynamic system, using synthetic data, and a third-order navigation system, using both synthetic and real data. Chapter 6 discusses the results.

The document of 46 pages (including 18 figures, 2 tables and 3 algorithm summaries) has a very good graphical level. There are no critical errors in the language use. However, the text is sometimes repetitive (e.g, 4th paragraph on page 20) and certain formulations are misleading or difficult to understand (e.g., 8th paragraph on page 26, 2nd paragraph on page 30). There is a substantial absence of articles. The text is satisfactorily structured into six chapters. There are many wrong references to mathematical objects (e.g., 5th paragraph on page 19) and figures are sometimes referred to as Figure or picture (e.g., 2nd paragraph on page 19). The document contains 30 bibliographical items, where some of these—namely, the conference papers—have an inconsistent style. Some parts of the text, such as (2.35), are not directly used in the thesis or are introduced repeatedly (e.g., 2nd paragraph on page 29). Some abbreviations, e.g., FPD-IF, are not defined. Instead of using mathematical symbols, some quantities are described by directly equating words to values, i.e., "accuracy=3.4" (4th paragraph on page 45). There are no units of measurement in the figures and tables.

I find the thesis to completely fulfill the assignment.

The author correctly cited bibliographical sources to distinguish his own results and thoughts.

The technical level of the thesis is good. Nonetheless, I am concerned with the author's interpretation of various physical and mathematical aspects of the thesis:

- In Case 2, after (2.33), it is mentioned that the intersection is an empty set. However, in the sentence preceding (2.35), the author claims that this intersection is a non-empty set.
- The variational distribution in (2.29) is incorrect.
- The optimization objective in (3.4) contains undefined symbols. There is practically no explanation behind the meaning of this optimization problem.
- There is no mention of the discretization method used to obtain (5.12).
- The author defines (5.14) as the total norm-squared error. However, after selecting the entries of the state vector, (5.14) becomes just the total squared error.
- The quantities ω_1, ω_2 and ν are defined as noise bounds, yet the description of the results in Chapter 5 indirectly refers to these quantities as "precision", "accuracy" or "uncertainty". There is an insufficient level of detail on how ω_1, ω_2 and ν were obtained in the real-data experiments.
- There is no attempt to tune or optimize ω_1, ω_2 and ν in the real-data experiments.

I have two main concerns with the results presented in Sections 5 and 6:

- The results in Figure 5.1 are not convincing. The values of ω_1, ω_2 and ν are not very realistic for sensors that are commonly used in navigation systems. The KT filter is practically uninfluenced by the source filter, as also shown in Figure 5.5 where the performance of the KT filter is nearly constant. The experiment was performed over 500 repetitions. However, there is no error bar which could prove significance of the small difference between the methods (the std. deviation for the IF filter—such as the one presented for the other filters in Figure 5.5—is missing).
- Indeed, we do not know the true state in real-data applications. In such cases, the best we can do is to measure the filter's performance via the innovation sequence, i.e., the sequence of differences between the observation y(t) and the prediction ŷ(t) given by a specific filter (model). Any reasonable choice of a cumulative distance between y(t) and ŷ(t)—such as the total squared error (5.14) adopted in Table 5.1—provides a fair comparison among various filters. Therefore, I do not agree with the author's explanation for the poor performance of the KT filter, as discussed in Section 6 and Conclusion.

  Given that the process and observation noise variables are individually i.i.d. and mutually independent (as assumed by the author in Sections 1 and 2), the innovation sequence is stationary and ergodic. The author fulfills these assumptions in the synthetic-data experiments, and, therefore, the results presented in Figures 5.1 and 5.5 are methodologically valid (yet they are not convincing due to the improper settings, as mentioned above). These assumptions are certainly not fulfilled in the real-data experiments. Therefore, it is necessary to perform further analysis of the innovation sequence. For example, evaluating the autocorrelation would allow for assessing the suitability of the model (5.9) to the real data collected by the author, revealing that the main reason for the poor performance of the KT filter (and the other non-isolated filters) lies in the model mismatch (e.g., there is no model of the bias of the sensors in (5.9)), which can jeopardize the KT filter's ability to reject imprecise source knowledge.

The author demonstrated the ability to work on an interesting and original research problem. Therefore, I recommend the thesis for defending the Master's degree. Nonetheless, due to the aforementioned shortcomings, I suggest grading the thesis with **C - good**.

Questions:

- Can you present an equivalent of Figure 5.1 using more realistic settings of ω_1, ω_2 and ν? For example, you can set values close to those you used in the real-data experiments, allowing the change in ω_1 and ω_2 by, say, two orders of magnitude (to reduce the simulation time).
- In Section 6 you mention that it is unclear why the centralized fusion is more successful in estimating the acceleration than the other approaches. Could you briefly describe what is the observability and reconstructability of a dynamic system, and whether it can help you in explaining this issue?

Date: 20th May 2022

Signature: