UNIVERSITY OF WEST BOHEMIA

doc. Ing. Ondřej Straka, Ph.D.
Department of Cybernetics
University of West Bohemia
Univerzitni 8, 301 00 Plzeň
Czech Republic

Thesis title: **State estimation and fault detection with reduced error sensitivity to parameters**

Thesis author: **Ing. Jaroslav Tabaček**

State estimation of stochastic dynamic systems is an essential technique for tracking and navigation, signal processing, system identification, optimal control and fault and change detection, and many other technical and non-technical areas. The state is observed indirectly using noisy measurements provided by sensors. The system behavior is represented by a dynamic model with limited accuracy. The model parameters are often uncertain. The problem of state estimation with such models is addressed by robust state estimation. The thesis of Jaroslav Tabaček focuses on such a problem of robust state estimation and its application in fault diagnosis. The tasks are treated for both monolithic models and a network of interconnected models.

The thesis is 95 pages long and is structured as follows: Chapter 1 provides short surveys of robust state estimation methods, fault diagnosis, and corresponding distributed algorithms. Chapter 2 motivates the thesis by presenting the advantages of the approach taken by the desensitized Kalman filter (DKF), which has yet to attract the interest of the community. The thesis thus aims to address several drawbacks of the filter and demonstrate the benefits of its application to fault diagnosis for monolithic and distributed systems.

Chapter 3 then introduces a generalization of the original DKF to the case of correlated process and measurement noises. The generalization is introduced for the original DKF with filter gain given by an implicit relation (DKF for linear systems with noise correlation – Algorithm 3.1) and also for the DKF with an explicit specification of the filter gain (Special DKF for systems with noise correlation – Algorithm 3.2). These algorithms are derived under the assumption of zero gain sensitivity. This assumption is relaxed in the derivation of the exact desensitized Kalman filter (XDKF), which minimizes a convex combination of the mean-squared error criterion and the sensitivity second non-central moment. Its optimal form is described by Algorithm 3.3. In addition to this optimal form, the author proposes several suboptimal forms of the XDKF. The suboptimal form assuming zero gain sensitivity special DKF is shown to be an equivalent to the special DKF. Other suboptimal forms focus on the accumulation of updates if individual second moments are not required, steady-state solution, or separating data- and time-update steps. The last subsections of the chapter deal with the stability issue of the algorithms and with a problem formulation involving normalized objectives. Such problem formulation facilitates a more suitable optimization problem parametrization.

Chapter 4 then shifts the attention from linear to nonlinear models and proposes an analogous algorithm for models linearized by means of Taylor series. Chapter 5 applies the proposed algorithms to fault diagnosis problems. The faults are represented by models that differ from the nominal model describing the fault-free behavior of the system. The switching among the models is modeled by a homogeneous Markov chain with prespecified transition probabilities. The proposed solution is built on the interacting multiple models (IMM), which is a well-known technique to address the problem of an exponentially increasing number of hypotheses. The state estimation for individual models is then solved by the proposed XDKF algorithms.

Chapter 6 focuses on the application of the XDKF algorithm to the problem of distributed state estimation, where the system is represented by a network of linear models with non-overlapping states. In the network, the states of models connected to a particular model act as inputs in the model dynamics. In such a setting, the filters associated with neighbor models usually exchange the state estimates and corresponding estimate error covariance. To minimize the communication, the author assumes that only state estimates are exchanged, and their inaccuracy is taken into account by using the XDKF. Finally, Chapter 7 focuses on combining the fault diagnosis from Chapter 5 with the distributed architecture to address fault diagnosis of large-scale systems.

The performance of the proposed algorithms is illustrated using numerical examples. The fault diagnosis algorithms are illustrated using an application for a building heating system, which is very interesting. Chapter 8 then concludes the thesis, highlights the contributions and provides challenges for future work.

The thesis is well-structured, with each part of the contribution having its own chapter containing a short introduction, conclusion, and a simulation example demonstrating the techniques and methods. It is also carefully written and checked with only several inaccuracies, which will be given further, and a few typos. Those aspects make the thesis highly accessible.

General comments to the thesis: The thesis contributes to a systematic development of robust state estimation algorithms and their application in fault diagnosis. The topic of the paper is highly topical and the proposed algorithms complement the range of different robust estimation methods with a new approach with attractive properties. Compared to the state-of-the-art desensitized algorithms, the proposed algorithms relax the zero gain sensitivity assumption, which might seem unsubstantiated. Besides linear models, state estimation algorithms are also designed for nonlinear models, which are ubiquitous in all application areas. I particularly appreciate the utilization of reduced sensitivity to parameter changes in distributed state estimation so that it reduces requirements for communication among individual nodes. Improved fault diagnosis quality of algorithms then naturally follows from the utilization of the proposed desensitized filters in the detection algorithms when the system model parameters are uncertain. The results achieved in the thesis are, in my opinion, highly valuable and of great potential for applications. The work has not avoided some errors and inaccuracies, a list of which is given below.

Comments to the publication activity:

Jaroslav Tabaček published the research results in two journal papers (Journal of Franklin Institute) and four papers presented at international conferences. Based on his scientific contributions in the thesis and published articles, it can be argued that he pushed forward the state estimation field and demonstrated his abilities of systematic work.

Below is a list of issues that can be considered inaccurate or require a more detailed specification:

- Page 1, 1st paragraph: It shall be emphasized that most of the claims hold only for linear models.
- The condition $\text{cov}(\mathbf{v}_k, \mathbf{x}_k) = 0$ can be replaced by the whiteness of $\mathbf{v}_k$ (already assumed) and $\text{cov}(\mathbf{v}_k, \mathbf{x}_0) = 0$.
- The condition $\text{cov}(\mathbf{e}_k, \mathbf{x}_k) = 0$ can be replaced by the whiteness of $\mathbf{e}_k$ (already assumed), $\text{cov}(\mathbf{e}_k, \mathbf{x}_0) = 0$ and $\text{cov}(\mathbf{e}_k, \mathbf{v}_\ell) = 0 \; \forall \ell < k$. This is important as $\text{cov}(\mathbf{e}_k, \mathbf{v}_\ell) = \mathbf{S}$ for $\ell = k$.
- The reason that average variance PKF from 100 Monte Carlo simulations is higher than the average for other filters is rather caused by the following facts:
  - The variance provided by the PKF corresponds to the mean squared error (MSE) which is the smallest MSE among all unbiased estimators given the model matches the system.
  - The variance provided by a non-optimal filter (either due to an approximation or due to a model mismatch) may be smaller (optimistic estimate) or larger (pessimistic estimate) than the MSE. Thus, the comparison of the average variances computed by optimal and non-optimal filters offers little information.
- The RMSE specified by (3.38) is rather square root of time-averaged squared error. There is no index of Monte Carlo simulation over which the sample mean should be calculated.
- In the derivation of iterative relations for computation of the optimal gain (3.77) the author assumes zero partial derivative of K_{θ_p}. If such an assumption would be used already in (3.73), no iterations are needed. This would correspond to constraining the gain to functions for which the partial derivative of their total derivative is zero. Justification of such constraint should be provided.

University of West Bohemia
Univerzitni 8, 301 00 Plzeň

Email: straka30@kky.zcu.cz
Phone: +420 377 632 553

UNIVERSITY OF WEST BOHEMIA

- It is stated on page 23 that Assumption 3.2. eliminates the stochastic terms $\mathbf{e}_k$ and $\tilde{\mathbf{x}}_{k|k-1}$ from the sensitivity update and the sensitivity is no longer a stochastic variable. It is, however, still a function of the state estimate $\hat{\mathbf{x}}_{k|k-1}$, which is, as a function of the noise measurements, still a stochastic variable.
- The difference between total RMSE (e.g., Fig. 3.7) and cumulative RMSE (e.g., Fig. 4.1) is unclear.
- The time index of models is missing in the specification of the Markovian jump linear system (M_k should be used instead of m). Its specification significantly affects the interacting multiple model algorithm (IMM). The same applies to the local Markov jump linear system (7.1).
- Why does the state x have mode index as superscript? Does it mean that each mode has different states? If so, what is the relation between states of different modes?
- The arg max operator rather than max should be used in (7.29).

Minor issues, typos:

- Page 4: "It is assumed that measurements are mutually independent and uncorrelated with the process noise." It should rather be that "...the measurement noises are mutually independent...".

Questions for the exam:

- Is there a justification for zero gain sensitivity Assumption 3.1 while it is the criterion J that should be insensitive to parameter estimate?
- Would it make sense to define the sensitivity of the state prediction error (3.50) as a sensitivity to the parameter estimate error $\tilde{\theta}$ rather a sensitivity to the parameter estimate $\hat{\theta}$?
- Can you compare zero gain sensitivity assumptions? Assumption 3.1 states that partial derivative of the gain w.r.t. the parameter estimate is zero while Assumption 3.2 states that total derivative of the gain is zero.
- Relation $\sum_p \gamma_p \leq \sum_p \bar{\gamma}_p$ must hold for stable weights. Does this relation hold for weights $\nu_{\gamma,p}$ defined by (3.119)?

Final recommendation:

The thesis meets the requirements of independent creative scientific work and contains original and published results of the author's scientific work. I **do recommend** the thesis for presentation/defense with the aim of receiving a Ph.D. degree in the field of Control Engineering and Robotics.

Plzeň, April 25, 2023
doc. Ing. Ondřej Straka, Ph.D.

University of West Bohemia
Univerzitni 8, 301 00 Plzeň

Email: straka30@kky.zcu.cz
Phone: +420 377 632 553