Referee report on the Ph.D. dissertation entitled "*System supervision and abnormality detection through multivariate statistical methods*" by Bálint Levente Tarcsay

The importance of safety cannot be overestimated in chemical manufacturing systems, which must be ensured in an increasingly difficult regulatory (e.g., emission limits) and economic (fluctuating prices and costs, competition at a global level) environment. The advent of Industry 4.0 with the spread of extensive monitoring systems opens new horizons for process monitoring, control, and fault diagnostics, but it also puts an overwhelming load on the operators. This culminates in alarm management, as not all alarms indicate a critical issue that seeks immediate action, but missing a critical one may have fatal consequences. Identifying and suppressing the less important ones to allow sufficient attention to the eventual safety-critical ones becomes crucial in modern chemical systems. However, this is a challenging engineering task, and it is the topic of this Ph.D. dissertation, an important and timely problem. In this work, Bálint Levente Tarcsay provides model-based support for such tasks.

Bálint Levente Tarcsay presents a monographic dissertation spanning a total of 120 pages, organized into seven chapters. The quality of editing (typography), figures, and tables is high. The main section of the thesis is followed by an extensive bibliography and appendices that offer supplemental information for a more thorough understanding of the low-level details.

Chapter 1 is the introduction, with a total length of less than three pages. The section lists the goals/thesis points of the dissertation. At a high level, it introduces the problem and importance of fault detection and isolation. The work is also motivated by the abruptly increasing number of scientific publications over the last decades. It is fascinating that the number of patents per year is also plotted, which seems to be lagging behind the exploding number of publications. This itself is an excellent motivation to conduct application-oriented research. Finally, it describes the structure of the PhD dissertation. This chapter is like a high-level motivation for the work.

Chapter 2 is the literature survey on fault diagnosis. The section reviews the quantitative and qualitative fault detection and isolation techniques and process-history-based fault detection. The section concludes with the current bottlenecks and future trends in the field. The section is written at a high technical level, including the representative equations as well. The detailed description and illustration of DPCA help the readers better understand how the DPCA is derived from the PCA and, subsequently, to understand the thesis results.

Chapter 3 presents a DPCA-based fault detection strategy, which relies on identifying irregular data points, probably belonging to a fault by Q-statistic, then comparing the dynamics to a pre-

determined and saved fault library using the Fréchet-distance. A comparison of alternative methods supports the selection of this metric. This concept is presented through the case study of a typical three-tank benchmark system. Even though the benchmark problem is somewhat simple, and this particular case could have been solved with traditional engineering methods by mounting two additional flow rate sensors, the fault detection and isolation performance is stunning. It would have been interesting to deploy this method on a more complex system, e.g., the Tennessee-Eastman process. I would like to underscore that this criticism is for the benchmark problem, not the Ph.D. candidate's work.

In Chapter 4, risk assessment techniques were combined with fault detection. This is very useful and important, as the severity of different faults can differ significantly. In this chapter, Bayesian networks were developed for failure mode and effect analysis, allowing failure prioritization. The performance was presented in two simulated case studies: the three-tank benchmark problem and the dehydrogenation reactor, and good results were obtained in both cases.

Chapter 5 diverges somewhat from the other parts of the thesis as it introduces a distributed parameter system, to which a significant CFD, and CFD-inspired compartmental model development was presented in details. This alone is an important contribution and could have been a separate chapter. The CFD and compartmental models were validated experimentally, which is a valuable piece of the thesis. The model was used to generate a synthetic database (similar to the previous case studies), and the DPCA-based fault detection was deployed. In my view, applying the fault detection algorithm to the real system would have been outstanding.

Chapter 6 concludes the thesis, and Chapter 7 contains the thesis points (new scientific results), all supported by accepted publications. The thesis points are written concisely. This section also underscored the Candidate's excellent publication record and collaboration skills, as the thesis does not include some publications resulting from collaborative research.

Each case study is executed at a high technical standard, as described in Chapters 3-5. The results and individual conclusions are significant. The topic of the thesis is timely and holds significant practical potential. The literature study is thorough, which permitted the Candidate to formulate the scientific goals that were solved and described in the thesis. The methods represent the state-of-the-art, and they were applied systematically, enabling robust, reliable conclusions to be drawn. These conclusions and the limitations are described and discussed in a clear, meaningful way.

Despite the Candidate already sorted out the Reviewer's technical questions in a previous iteration, a few follow-up questions for the final defense may be sorted out:

1. The Bayesian networks were applied, and outstanding results were obtained in the investigated case studies. I was wondering if the Candidate could comment on the practical

aspects of building such Bayesian networks, including identifying the failure modes and getting the corresponding scores for a technically much more complex, integrated system with a significantly higher number of failure modes.

2. Suppose we accept that the number of patents correlates with the number of technical realizations. In that case, one can assume that the practical implementations in the field are lagging behind the research and development – which, in general, is not outstanding. What are the main obstacles that prevent the widespread implementation in the field? How does the research conducted by the Candidate contribute to the practical realizations?

3. Finally, I would like to ask the Candidate to briefly present the most challenging technical problem he encountered during his research and the key element in solving that problem.

To summarize, Bálint Levente Tarcsai proved he could conduct high-quality, independent scientific research with this dissertation work. I accept all the thesis points as new scientific results. Therefore, I recommend accepting the thesis and awarding the Ph.D. degree from the University of Pannonia.

Budapest,
10 September 2024

Botond Szilágyi
Associate Professor of Chemical Engineering
Budapest University of Technology and Economic