

Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series “Springer Theses” brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student’s supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today’s younger generation of scientists.

Theses are accepted into the series by invited nomination only and must fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder.
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to scientists not expert in that particular field.

More information about this series at <http://www.springer.com/series/8790>

Nils Braun

Combinatorial Kalman Filter and High Level Trigger Reconstruction for the Belle II Experiment

Doctoral Thesis accepted by
the Karlsruhe Institute of Technology,
Karlsruhe, Germany

 Springer

Author

Dr. Nils Braun
ETP
Karlsruhe Institute of Technology
Karlsruhe, Germany

Supervisor

Prof. Florian Bernlochner
ETP
Karlsruhe Institute of Technology
Karlsruhe, Germany

ISSN 2190-5053

Springer Theses

ISBN 978-3-030-24996-0

<https://doi.org/10.1007/978-3-030-24997-7>

ISSN 2190-5061 (electronic)

ISBN 978-3-030-24997-7 (eBook)

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Supervisor's Foreword

We have entered an era in particle physics in which experimentation is led by computer programs and complex algorithms. Today's experiments have reached a complexity and intricacy such that no other way of analyzing the large streams of data—collected by instruments the size of houses and covering the area of soccer fields in some cases—can be imagined. It is fair to say that neither the discovery of the mechanism of charge-parity violation nor the confirmation of the existence of the Higgs boson, both feats that were rewarded with Nobel prizes, would have been possible without the use of state-of-the-art software and complex algorithms. And it is also reasonable to believe that such tools will play a crucial role in future discoveries. In this thesis, Dr. Nils Braun documents something one could consider the backbone of an experiment: During his Ph.D., he carried out work at Belle II and worked extensively on algorithms that reconstruct trajectories of charged particles traversing the detector. In addition, he made crucial contributions to the so-called online computing, which uses large computing farms to partially reconstruct collision events to decide which of them are of interest to physicists and should be stored to disk. At the time of writing, the Belle II experiment has just started to record electron and positron collisions produced by the SuperKEKB accelerator facility north of Tokyo in Japan. First preliminary physics results are expected this summer, and all of them will make use of the work carried out in this thesis.

Let me describe a typical collision at Belle II: Electrons and positrons are accelerated to nearly the speed of light and brought into colliding paths at the center of the Belle II detector. The detector has the size of a four-story building (with another building of similar size, shielded from the harsh radiation environment such collisions create, to read out the individual sub-detectors). During some collisions, a special type of quantum annihilation occurs, which produces two heavy Beauty quarks. These very unstable elementary particles decay after a very short lifetime of about 1.5 ps into a number of charged and neutral particles and these decay remnants allow us to study and understand the properties of our universe. Do new heavy fundamental particles exist? Is there a connection between the elusive Dark matter and everyday visible matter? In some sense, Belle II is a time machine that

allows us to recreate conditions shortly after the Big Bang, when Beauty quarks were abundant and their properties had an impact on the future evolution of the universe. But let us return to Belle II: The decay remnants of both Beauty quarks travel through the detector, and in case of charged particles, three dedicated track reconstruction detectors reconstruct their trajectories. At the very center of the detector and closest to the collision point, a pixelated small sensor with over 7 million individual pixels is located. The volume outside is instrumented by double-sided strip-detectors and a sizable gaseous drift chamber. Both of the outer detectors possess a fast readout, and one crucial aspect of the data-taking scheme at Belle II is that the innermost pixel detector cannot be fully read out for each interesting collision. The information of the two outer detectors thus has to be used to extrapolate to a small region of interest on the pixel detector, which in turn can be read out. If this region of interest is wrong, the recorded information is lost, and with it much of the necessary precision needed to carry out physics measurements at the forefront of our field.

This is where Dr. Nils Braun and his thesis enter the scene: he took up this challenge in his outstanding Ph.D. work. He demonstrated that a combinatorial Kalman filter can be used to carry out such an extrapolation with the necessary precision. His Kalman filter implementation now plays a key role in the sequence of track reconstruction algorithms currently being employed by the Belle II collaboration. His thesis offers a very pedagogical introduction into the topic and offers an excellent starting point for, e.g., track reconstruction algorithms for future experiments. In addition to these crucial contributions, Dr. Braun also played a major part in optimizing the execution time of the online reconstruction code that stores or discards collision events selected by a fast hardware-based trigger. This decision is taken by partially reconstructing the collision event using a large computing farm called the high level trigger. The decision has to be executed within a narrow time window (dictated by the incoming stream of collision data and the limited processing power to select collision events for permanent storage). Dr. Braun's contributions here not only helped to keep the reconstruction time within the allocated limits but also contributed to a new error-proof parallelization implementation, which will stabilize the high level trigger operations in the future. All of this is also outlined in his thesis, providing interesting metrics and supplemental material to optimize such systems.

Thus in summary: This thesis is a must-read, it offers comprehensive and detailed documentation of the Belle II charged particle reconstruction algorithms, and it summarizes crucial parts of the online computing system.

Nils, it was a true privilege to work with you and I wish you the very best for your future! No matter what you will end up doing, you will be outstanding at it.

Sincerely

Karlsruhe, Germany
July 2019

Prof. Florian Bernlochner

Acknowledgements

I always knew very well that there were many people out there I could rely on. I would like to express my sincere thanks to all of them, because without them this work would not have been possible.

My first thanks go to Prof. Dr. Michael Feindt for taking over the main lecture and Prof. Dr. Florian Bernlochner for taking over the co-lecture. Michael has given me a lot of freedom in my choice of topics since I wrote my Master’s thesis with him at our institute. I do not want to miss the insights into the industry of data science—also besides science—which I got through him. I would like to thank Florian very much for the always good and personal support. Not only the technical advice but also the help with the organizational tasks helped me a lot. His motivating nature and his many new ideas have brought me and our whole group forward.

Track reconstruction and software development are above all teamwork, and our former Convener for Tracking, Dr. Martin Heck, was not only a good leader of the tracking group but has also brought our local working group in Karlsruhe forward with ever new ideas and improvements. His wealth of knowledge and experience was an enrichment for me. I want to thank him for the many years of collaboration and support. My sincere thanks go to Dr. Thomas Hauth. He was my supervisor for large parts of the thesis and my first help in case of problems. I would like to thank him for his competent help, his support despite the great geographical distance, the new ideas we had in the many discussions, and his very open and friendly nature. I could not have wished for a better supervisor.

Since my time as a master student, I have felt very comfortable in our working group and this is mainly due to my colleagues. I would like to highlight my (former) office colleagues Christian, Moritz, and Markus, with whom I was not only able to discuss common topics but also common problems and had a lot of fun. I would also like to thank everyone who proofread my work. My big thanks also go to the admin team of the institute, who managed to maintain a very smooth operation despite many external problems.

Without the work of several open source developers of countless Python packages, the good questions and answers on many blogs and forums on the Internet (especially Stack Overflow) and the large and helpful community of software developers, this work would not have been possible. My thanks also go to all of them. I would also like to thank the many other developers in the Belle II collaboration, especially Martin Ritter, who answered many of my questions.

Finally, I would also like to thank my family, my friends, and my girlfriend Sophie. They have accompanied me throughout my studies and my doctoral thesis and have helped me with words and deeds.

Contents

1	Introduction	1
1.1	General Introduction into Belle II and This Work	3
	References	5
2	Experimental Setup	7
2.1	From Accelerated Electrons to Final State Particles	7
2.1.1	Accelerator Complex at KEK	8
2.1.2	Detector Design	9
2.1.3	Beam-Induced Background	15
2.2	From Digital Detector Data via the Trigger to the Storage Drives	17
2.2.1	Detector Readout and Hardware Trigger	17
2.2.2	Online Software Trigger	17
2.2.3	Storage and Offline Reconstruction	21
2.3	Belle Analysis Software Framework 2	21
2.3.1	Inter-Module, Inter-Process and Inter-Node Communication	22
2.3.2	Monte Carlo Simulation	22
2.4	Time Schedule and Performed Tests	24
	References	25
3	Foundations	27
3.1	Track Reconstruction	27
3.2	Track Finding Concepts at Belle II	29
3.2.1	Track Finding in the CDC	29
3.2.2	Track Finding in the VXD	32
3.2.3	MC Matching and Performance Indicators	36
3.3	Track Fitting	38
3.3.1	Tracking Parameters	39
3.3.2	Track Fit Using χ^2	39

3.3.3	Kalman Filter	41
3.3.4	Deterministic Annealing Filter	44
3.3.5	Extrapolation	45
3.4	Mathematical Foundations	45
3.4.1	Multivariate Analysis (MVA)	46
3.4.2	Treatment of Statistical Uncertainties	47
References		47
4	Fast Reconstruction for the High Level Trigger	51
4.1	Introduction	52
4.1.1	Event Composition	52
4.1.2	Level 1 Trigger	54
4.1.3	Requirements	55
4.2	Processing Time Studies	56
4.2.1	Speedup $\eta(p)$	57
4.2.2	Single-Core Reconstruction Time $T(1)$	69
4.2.3	Summary	72
4.3	Principles of <i>FastReco</i>	73
4.3.1	Continuous HLT Decision	74
4.3.2	Bhabha Veto	76
4.3.3	General Framework	80
4.4	Multiprocessing Using the ØMQ Library	82
4.4.1	Overview on ØMQ	84
4.4.2	Implementation for basf2	84
4.4.3	Performance	87
4.5	Summary	89
References		89
5	Event Timing	93
5.1	The Role of the Event Time	94
5.2	Event Timing Using the CDC Information	95
5.3	Flight Time Estimation	97
5.3.1	Performance of the Track Finding with Wrong T_0 Assumptions	97
5.3.2	Track Time Estimation	97
5.4	Event Time Extraction Using the Drift Length Information	100
5.4.1	Components of the Measured Time in the CDC	100
5.4.2	Performance on MC Simulations	102
5.5	Event Time Extraction Using χ^2 Information	103
5.5.1	Minimizing the χ^2 Function	104
5.5.2	Performance on MC Simulations	106
5.6	Event Time Extraction Using Hit Information	107

5.7	Combination of the Methods	109
5.7.1	Performance on MC Simulations	110
5.7.2	Comparison with Other Event Time Extraction Methods at Belle II	111
5.8	Summary	115
	References	115
6	Combinatorial Kalman Filter	117
6.1	Motivation	118
6.2	Principles	118
6.3	Application of the CKF at Belle II	120
6.4	CKF from CDC to SVD—An Example	122
6.4.1	Build Relations	123
6.4.2	Tree Search	123
6.4.3	Filters	124
6.4.4	Final Candidate Selection	125
6.5	CKF for SVD Hit Attachment	126
6.5.1	Characteristics of CDC Tracks	126
6.5.2	Legacy Implementation	128
6.5.3	Performance	131
6.5.4	Additional Runs of the CKF and VXD Track Finding	137
6.6	Merging of CDC and SVD Tracks	139
6.6.1	Overall Performance of the SVD CKF Tracking Chain	141
6.6.2	Comparison with VXDTF2	147
6.6.3	Performance on ROI Finding	148
6.7	CKF for PXD Hit Attachment	150
6.7.1	Influence on $B^- \rightarrow D^0 \pi^-$ Decays	155
6.7.2	Reduced PXD Setup	159
6.8	Performance in Phase 2	161
6.8.1	Characteristics of the Phase 2 Setup	162
6.8.2	Performance on Simulated Events with the Phase 2 Geometry	165
6.8.3	Performance on Phase 2 Data	166
6.9	Summary	171
	References	173
7	Conclusion	175