Report on the PhD thesis "Fragmentation through heavy and light-flavor measurements with the LHC ALICE experiment" of Zoltán Varga

One of the most surprising results at the LHC was the discovery of collective effects in proton-proton collisions with high charged particle multiplicity in the final state, such as long-range near-side correlations and anisotropic flow. To explain these observations, one either assumes that quark-gluon plasma (QGP) can be created in the collisions of small systems or looks for other physics effects that could induce collective behavior. The latter has serious implications also for the interpretation of heavy ion collision data and the observation of the QGP, a unique state of matter that existed in the very early universe.

So **the thesis addresses a very relevant question** in high energy particle and nuclear physics about the non-QGP origin of collectivity. It covers phenomenological studies using standard event generators for the fragmentation and hadronization process in proton-proton collisions, as well as a preliminary experimental investigation of the Koba-Nielsen-Olesen (KNO) scaling of the charged particle multiplicity of jets for various jet transverse momentum ranges.

The thesis is **written in English** which is a practical choice for a work produced for and within a large international collaboration. The **quality of the language is good,** easily readable, with a limited number of small mistakes.

After a short "Introduction" (Chapter 1) announcing the topics and the structure of the thesis, selected topics from "High-energy hadron collisions" (Chapter 2) are presented to set the stage for the PhD research. The experimental details are rather brief, e.g., the reconstruction of physics objects (e.g. tracks, vertices, electrons…) is not included. From Chapter 3 to 7 the various research topics and results are described which also appear as thesis points. The thesis **mostly concentrates on phenomenological studies of pp collisions**, even though the later chapters **include results using ALICE collision data**. The start of most chapters reads like a dense scientific article (and most seem to be copied from papers that the candidate is an author of, the self-quotation however is not marked by a reference). In a PhD thesis a bit more pedagogical explanation would be welcome.

The organization of the information has some shortcomings. A number of concepts and phenomena are mentioned with no explanation, some reappear later and gets explained, others never explained. The figures frequently appear rather far from their first mention. Some of their captions (or the discussion in the main text) would have benefited from more details. In general, the bibliography could have been improved, e.g. a link to the online documents added, publication data such as journal missing for several entries, along with some original references.

The results of the thesis have been **published in four articles** that appeared in refereed journals and **five conference publications** (most of which were also refereed), so they clearly **fulfill the formal requirements of a doctoral (PhD) degree.**

I accept all thesis points of the candidate as original research contributions. These cover the followings:

- Chapter 3 presents a simulation based study of jet properties and their multiplicity dependence in pp collisions and concludes that the observed jet shape / size originates from the fragmentation process (thesis #1).
- Chapter 4 looks into the scaling properties of light and heavy flavour jets in pp simulations, and then chapter 5 presents preliminary and not yet approved experimental studies in ALICE data of jet multiplicity distributions as a function of the jet pT to verify if KNO-type scaling holds (thesis #2).
- Chapter 6 discusses the ALICE measurement of azimuthal correlations of decay electrons originating from heavy flavour production to which the candidate contributed from the simulation side (thesis #3).
- Chapter 7 finally focuses on charm-baryon enhancement in simulation and its connection to the underlying event, for various baryon species (theses #4 and #5).

The presentation of chapter 5, which covers unpublished experimental studies still under development, is less mature than the rest of the dissertation, especially considering the figures. As an example, in Fig 5.18 showing the estimated systematic uncertainties, there is no mention of which subfigure belongs to which pT range. Some lines corresponding to the systematics are missing and the total systematic error does not seem to be consistent with the components (e.g in top left subfig - possibly due to a drawing feature when the y axis range is not properly chosen). Fluctuations of the uncertainty from bin to bin might suggest a large statistical component in their determination. However, I very much welcome this chapter as it shows **an interesting analysis of collision data by the candidate covering the essential experimental steps**. It has surely widened his experience and contributed to his development as an independent scientist at the intersection of theory and experiment.

My questions on the research and the theses are the following:

1) At the end of section 2.5 it is stated about small systems that manifest collective phenomena: “In recent years a consensus starts to emerge that the observed collective phenomena could be explained by semi-soft vacuum-QCD effects, such as multiparton-interactions [66] with color-reconnection [34] or minijets (semi-hard partons produced by incoming partons or bremsstrahlung) [67] production.” If so, could it be that the discovery of QGP is misguided and the observed phenomena in heavy ion collisions are also simply the result of such effects?
2) Why was the Bayesian unfolding chosen for the measurement of the jet multiplicity distribution in the study of KNO-like scaling (chapter 5)? Were other methods tested? How did you verify that the result is not sensitive to the assumed prior? The text is not very informative about this: “The initial MC distributions were reweighted and smoothed out as the new choice of prior for the unfolding process.“ Reweighted to what? Did you only remove statistical fluctuations from the prior by smoothing? Did you try unfolding with a flat prior, for example? Or using a different MC model? These could be used as meaningful closure tests, e.g. unfold a MC data with known truth using a different MC for the prior and the response matrix.
3) What is the typical uncertainty on the denominator and numerator values for the ratio shown in Fig 5.11? Are the up to ~10% fluctuations visible on the plot consistent with them? Why is the iteration uncertainty negligible (shouldn’t Fig. 5.18 green line show similar values as Fig 5.11)? For the binning systematic, how was the 5% variation chosen? Are the stat uncertainties displayed on Fig. 19? How are (will be) the uncertainties from Fig. 5.18 propagated to Fig. 5.19?

4) In the measurement of the azimuthal correlations of heavy-flavor decay electrons (chapter 6), why was the FONLL calculation chosen to determine the contributions of bb and cc production (their relative importance) for the Pythia 8 prediction? What are the advantages of FONLL when compared to other calculations? What is the theoretical uncertainty on the FONLL prediction? How different is the bb/(bb+cc) fraction in EPOS3? Is this difference a significant contribution when trying to understand the differences between the Pythia8 and the EPOS3 predictions? Can there be any significant effect from the decay tables used in the simulations? Were they the same for Pythia8 and EPOS3? In general, one of the main problems of the comparisons of theoretical models among themselves and with the data is that the theoretical predictions do not have their uncertainties marked and thus it is not possible to judge whether the differences are significant or not. What are the main sources and the typical sizes of theoretical uncertainties?

In summary, I accept all the results claimed in the thesis booklet, I **recommend the doctoral thesis for public defense**.

Budapest, 24 November 2024.

Dr. Gabriella Pásztor
ELTE TTK Institute of Physics