Supervisor's Report on the MSc Thesis

Student: Bc. Georgij Ponimatkin
Thesis Title: Identifying Heavy-Flavor Jets Using Vectors
                 of Locally Aggregated Descriptors

Mr. Georgij Ponimatkin's MSc thesis is devoted to the possibility of identifying jets originating from the decay of heavy quarks using modern methods based on machine learning in order to achieve greater accuracy compared to the standard methods used in particle physics. In his thesis, Georgij focused specifically on the development of the set based tagging JetVLAD model and subsequently evaluated its performance on simulated proton-proton collision data at 200 GeV available at the RHIC accelerator at Brookhaven National Laboratory. In his work, Georgij clearly demonstrated that the JetVLAD model achieves very good classification performance in the studied range of jet transverse momenta from 5 to 40 GeV/c in p+p collisions at RHIC, both for efficiency ($> 80\%$) and signal purity ($> 80\%$) as well as for the rejection factor (> 200). The model is also resilient to pileup effects and thermal background, which should allow its application on real data.

The thesis is written in English and is divided into four chapters and a conclusion. In Chapter 1, Georgij introduces basic concepts from the field of jet physics with a focus on jets containing a heavy quark and a summary of current physical measurements in this area. Chapter 2 and Chapter 3 contain an introduction to computer vision, machine learning and an overview of current results in this area. Georgij's own research work is described in Chapter 4, which deals with the JetVLAD model and its specific application to simulated proton-proton collision data for the acceptance of a typical midrapidity detector, such as the STAR experiment at RHIC or in near future the sPHENIX experiment. Hyperparameter optimization and demonstration of the JetVLAD model performance is performed both on p+p data simulated using Monte-Carlo generator PYTHIA, which is commonly used in particle physics, and on p+p data, to which a thermal background is added. The background is supposed to mimic (in a simplified way) the situation prevailing in heavy ion collisions, where the study of hard probes, such as heavy quark jets, is potentially influenced by the presence of a background from soft particles forming the nuclear medium (bulk). At the end of the diploma thesis, a discussion of the obtained results is presented.

Georgij showed a great deal of innovative thinking, diligence and worked during his thesis project enormously hard. Work on the development of the JetVLAD model and its application and optimization took place in close cooperation with Dr. J. Šivic (CTU Prague), Dr. R. Kunnawalkam Elayavalli (Wayne State University/Yale University) and Prof. J. Putschke (Wayne State University), but the main credit for the achieved results and the writing of the article "Identifying Heavy-Flavor Jets Using Vectors of Locally Aggregated Descriptors", which was published in JINST 16 (2021) 03, P03017, belongs in full right to G. Ponimatkin. The achieved results allowed Georgij to present a talk at the prestigious international conference Deep Inelastic Scattering (DIS 2021) in the USA and also at the 20th Conference of Czech and Slovak Physicists in Prague. This can be considered another great achievement and is by no means common for a MSc. student. Last but not least, I would like to

mention that in addition to a large amount of work done on his own diploma thesis, Georgij has always been kind and gave advice to other students working in the Ultrarelativistic Heavy Ion Group at the Nuclear Physics Institute of the CAS.

In conclusion, I am convinced that Georgij Ponimatkin has met the objectives in the thesis assignment and I recommend the thesis to be admitted for the defence and propose to evaluate it with a grade **A** (**excellent**).

In Prague, June 8, 2021

RNDr. Jana Bielčíková, Ph.D.
Nuclear Physics Institute of the CAS/
FNSPE CTU Prague