March 9, 2018

CMS Computing and Storage upgrade for Heavy Ion Collisions Supplement to DE-FOA-0001664

Ivan Amos Cali[1], Yen-Jie Lee[1], Boleslaw Wyslouch[1]

[1]Massachusetts Institute of Technology, Laboratory for Nuclear Science, Cambridge, Massachusetts 02139

Research in Nuclear Physics: CMS Level 1 Trigger Rate Upgrade for Heavy Ion Collisions

Institution:	Massachusetts Institute of Technology Office of Sponsored Programs
Street Address:	Room NE18-901 77 Massachusetts Avenue Cambridge, MA 02139
Principal Investigator:	Prof. Boleslaw Wyslouch
Phone:	(617) 253-7800
E-mail:	wyslouch@mit.edu
Administrative Contact:	Michele Hudak
Phone:	(617) 324-5382
E-mail:	hudak@mit.edu
Task Supervisor:	Prof. Boleslaw Wyslouch

Proposed Project Period:	June 1, 2017 October 31, 2018

Funding Opportunity Announcement Number:	DE-FOA-0001664
DOE/Office of Science Program Office:	Office of Nuclear Physics
DOE/SC Program Office Technical Contact:	Dr. James Sowinski
DOE Grant Number:	DE-SC0011088
Topic Area:	Heavy Ion Nuclear Physics

Contents

1 Introduction

Relativistic heavy ion collisions allow experimental studies of the quark-gluon plasma (QGP), the equilibrated high-temperature state of deconfined quarks and gluons. At the Large Hadron Collider (LHC), the abundant production of hard probes such as vector bosons, heavy quarks, jets, and quarkonia produced in the highest energy heavy ion collisions has opened a new era in the characterization of QGP, providing new information on initial state properties, parton energy loss and color screening. The capabilities of the Compact Muon Solenoid (CMS) detector in the detection of high momentum photons, electrons, muons, charged particles, and jets have proven to be a unique match for the opportunities at the LHC. A key aspect of this success has been the convergence of detector needs for precision measurements in p-p discovery physics in a high luminosity, high pileup environment and for QGP studies in high multiplicity Pb-Pb collisions. This convergence has allowed CMS to adapt a large number of new analysis techniques to studies of heavy ion collisions, such as particle-flow jet reconstruction, studies of jet substructures, lifetime fits for secondary J/ψ studies, b- and c-tagging of jets, missing p_{T} measurements of W production and energy flow in dijet events.

The full power of these and other techniques will be exploited in future heavy ion studies beginning in 2018 and Run III, when an increase in the collision energy to $\sqrt{s_{NN}} = 5.02$ TeV and an eventual increase in Pb-Pb collision rate to as high as $\sim 30-50$ kHz will enhance the production rate of hard probes by more than an order of magnitude. In this high energy and high luminosity era, CMS will undertake precision studies of complex observables such as γ-, Z^0- and $\mathrm{W}^{\pm}$-jet correlations, heavy flavor jet quenching, multi-jet correlations and the azimuthal anisotropy of high p_{T} jets and quarkonia. Recently, extensive studies of fully reconstructed heavy flavor mesons such as D^0 and B^+ in heavy ion collisions also show that CMS is ideally suited for the studies of heavy flavor production and heavy quark energy loss in the quark-gluon plasma.

The CMS trigger and data acquisition system is the key to achieving the ultimate high p_{T} physics reach in heavy ion collisions. Uniquely, the CMS trigger system only has two main components, the hardware-based Level-1 (L1) trigger and the High-Level Trigger (HLT), which is implemented as "offline" algorithms running on a large computer farm with access to the full event information. The conceptual aspects of applying this system to heavy ion collisions, which are characterized by a much lower rate but a much higher multiplicity than those encountered in p-p collisions, were first developed by the US CMS HI group (see, e.g. [1]). The HLT trigger work formed the basis for our previous proposals to the DOE Office of Nuclear Physics. As we demonstrated in those proposals, triggering on high p_{T} probes for Pb-Pb collision rates up to several kHz (i.e. the design value for Pb-Pb) is possible using event rejection solely or mostly at the HLT level. In the 2011 Pb-Pb run, this allowed a reduction of the event rate to storage by more than an order of magnitude compared to the collision rate, without prescaling the most interesting high p_{T} observables. In the 2015 Pb-Pb run, the heavy flavor meson trigger developed by the US CMS HI group from a previous DOE proposal was deployed for the first time. Those triggers increased the high p_{T} D^0 meson statistics for example in p-p (Pb-Pb) collisions by a factor of 600 (20), opening a new era for the precision heavy flavor physics in heavy ion collisions.

In 2017, a new 4-layer pixel detector was installed in CMS that greatly improved the impact parameter resolution of the charged tracks. This is a unique opportunity to combine the capability of CMS for the studies of high p_{T} probes with the heavy flavor program down to $p_{\mathrm{T}} \sim 0$, complementary to the physics goal of the major ALICE upgrade foreseen for Run III.

It has now become clear that the LHC will be able to significantly exceed the Pb-Pb design luminosity in future runs, possibly reaching up to 40-50 kHz already during the 2018 Run. The L1 rate upgrade work formed the basis for our previous proposals to the DOE Office of Nuclear Physics. The project was approved by the CMS management and upgrades to the CMS were performed to obtain the desired L1 rate of 30 kHz. The overall data to disk bandwidth is not of 3 GB/s averaged during the day that corresponds to 255 TB/day of RAW data. CMS is now able to collect un unbiased single track, heavy flavor meson, jet triggers samples that are critical for the CMS physics program and at the same time record a large statistics Minimum-Bias triggered sample that could be used for studies of low p_{T} heavy flavor mesons. For the 2018 Pb-Pb data-taking it is expected a delivered luminosity of up to 1.8 nb^{-1}. Six bilions Minimum Bias events will be collected during the 2018 Pb-Pb run.

The declared physics goal will add a significant stress to the computing and data sotrage system. The minimum bias dataset alone require 4 PB of storage already having taken into consideration a reduced raw data format. Using 25 k-cores, the reconstruction of the sample cannot be done faster than 54 days. In this proposal, we will present the computing and storage needs for the 2018 Pb-Pb run. Computing resources should be available not only for the data reconstruction but also for the consequent data analysis and montecarlo production. Section 2 gives insight on the physics motivation that leads to an extensive data collection, including the 6 bilions Minimum Bias Sample. Section 3 will describe the computing and storage needs for the collection of the 2018 Pb-Pb data. Finally, the proposed schedule, management plan, and the budget are described in Sections 4 and 5.

2 Physics motivation

In this Section we discuss the physics performance that could be obtained collecting an high statistics Minimum Bias sample during the 2018 high luminosity Pb-Pb run using three examples of measurements with rare probes. The performance of the upgraded system, which uses background subtraction for L1 event rejection and provides access to the full expected luminosity of 3 nb^{-1}, is compared to that of the current system. These examples illustrate the large improvements in statistical precision afforded by the improved ability to select jet events and high p_{T} particles during high luminosity Pb-Pb collisions with L1 trigger rate upgrade.

Below we discuss a selection of measurements that exploit the future high-luminosity data sets to provide important insights into the parton energy loss mechanism and heavy flavor particle production:

1. The flavor dependence of jet quenching is an important test ground for various parton energy loss models. Compared to light quarks, heavy quarks are expected to suffer from smaller radiative energy loss when passing through the medium because gluon radiation is suppressed at angles smaller than the ratio of the quark mass M to its energy E [2]. This effect (and its disappearance at high p_{T}) can be studied using charged particles, heavy flavor mesons and b-jets.

2. Measurements of the charged particle and heavy flavor meson azimuthal anisotropy (v_2) at very high p_{T} give access to the in-medium path-length dependence of energy loss. The azimuthal anisotropy as a function of particle p_{T} also yields important information about the parton energy dependence of the energy loss. At low p_{T}, hadronization of heavy quarks such as parton-to-jet fragmentation and the recombination of heavy quarks and light flavors has a significant impact on the magnitude of the heavy flavor meson v_2.

3. Studies of jet-heavy flavor meson correlation could give important insights into the parton energy loss mechanism and for the studies of medium induced radiation and medium response to energetic heavy quarks.

These and other measurements rely on a sufficiently selective jet and heavy flavor meson trigger to provide access to the full delivered luminosity. At the same time, a large minimum-bias sample is needed in order to access very low p_{T} heavy flavor mesons and to perform data-driven tracking and jet energy corrections.

2.1 Nuclear Modification Factors of Heavy Flavor Mesons

One of the proposed observables that reveal the flavor dependence of in-medium parton energy loss is the reduction of heavy flavor meson yield. This can be studied by measurements of nuclear modification factors (R_{AA}), defined as the ratio of the yield in nucleus-nucleus collisions to that observed in pp collisions, scaled by the number of binary nucleon-nucleon collisions. Theoretical calculations of R_{AA} for fully reconstructed D and B mesons are shown in Figure 1. At low p_{T}, the production rate of heavy flavor mesons in Pb-Pb is sensitive to the elastic energy loss of the heavy quark, the gluon shadowing effect in the nuclear parton distribution function, and the rate of recombination of the heavy quark and light flavors at the hadronization stage. At high p_{T}, the size of the suppression is sensitive to the heavy quark radiative energy loss in the pQCD-based models. Precision measurement of the R_{AA} from intermediate p_{T} ($\sim$ 10 GeV) to very high p_{T} (200-400 GeV) in bins of event centrality could provide insights about the pathlength and momentum dependence of the heavy quark energy loss and potentially distinguish between models based on AdS/CFT and pQCD.

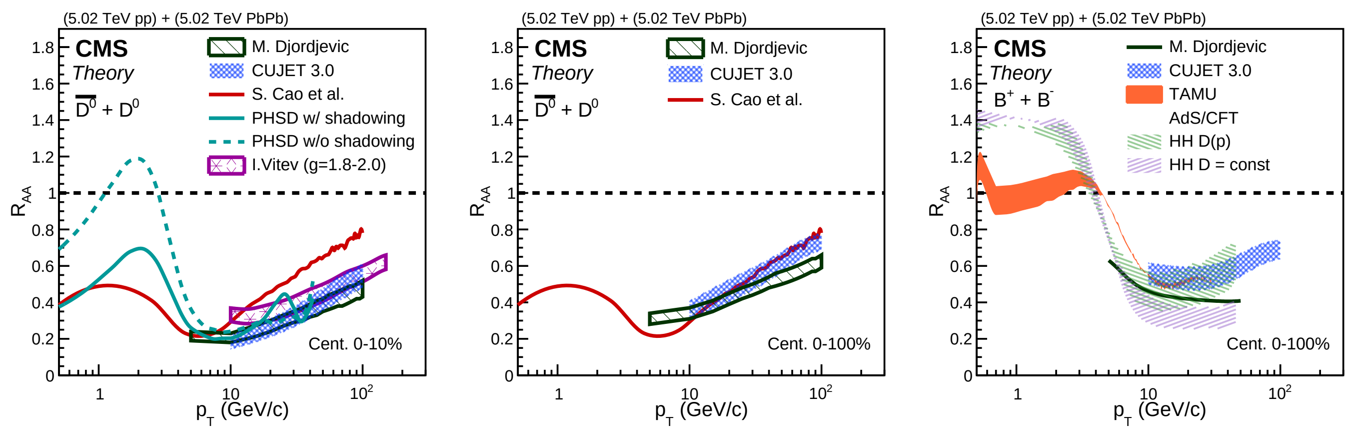

Figure 1: Theoretical calculations of nuclear modification factors of charged particles, D^0 and B^+ mesons.

It is also predicted that a significant fraction of the low p_{T} heavy quarks could hadronize via recombination with other quarks from the medium. This motivates measurements of the R_{AA} from D^0, D^+, (B^+), D_s^+ (B_s) and Λ_c (Λ_b) which could provide crucial information about the interactions of the heavy quarks and the produced medium. The D_s and B_s measurements are especially sensitive to the strangeness content of the produced strongly interacting medium. Figure 2 shows the performance of $D^+ \to \phi\pi^+$, $D_s^+ \to \phi\pi^+$, $\Lambda_c \to pK^-\pi^+$ and $B^+ \to \bar{D}^0\pi^+$ reconstruction in heavy ion collisions with the CMS detector, projected to the expected statistics to be collected with the L1 trigger rate upgrade. Significant signals of D_s^+ and Λ_c could be observed, enabling the measurement of D_s and Λ_c R_{AA} with the CMS detector for the first time.

Figure 3 shows the expected performance with the data recorded in 2015 and the projected performance in 2018 and beyond. With the high statistics jet and heavy flavor triggered sample, the precision of the R_{AA}

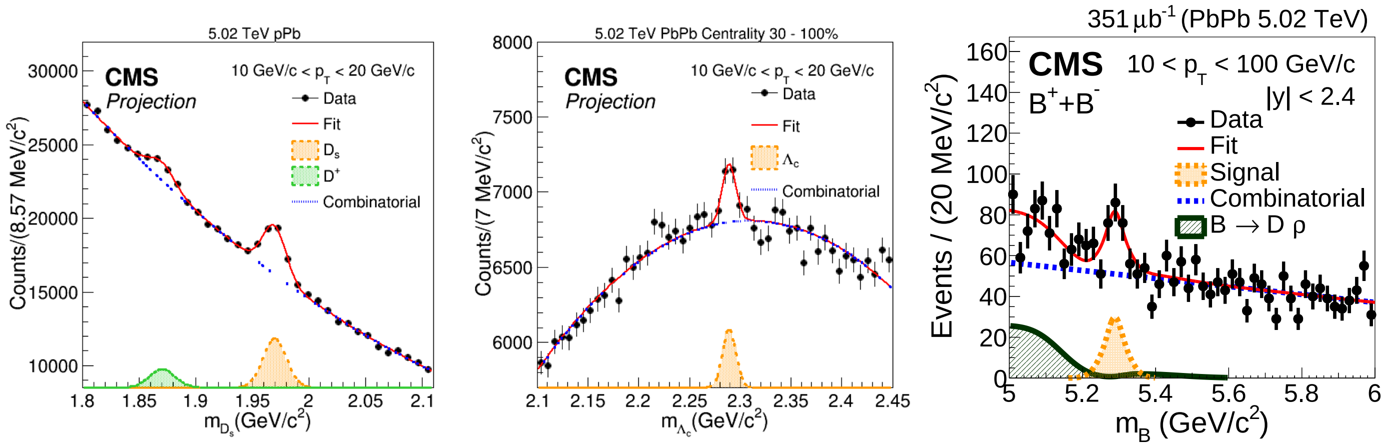

Figure 2: Invariant mass distributions of fully reconstructed heavy flavor mesons

measurements at high p_{T} could be greatly improved. At the same time, with the L1 trigger rate upgrade proposed in this proposal, the much larger minimum-bias sample enables CMS to perform these studies down to the very low p_{T} region. The expected precision of D^0, B^+ and D_s R_{AA} measurements from low p_{T} to high p_{T} could provide a strong constraint on theoretical models, and the difference in the suppression magnitude between those mesons could be observed for the first time.

2.2 Azimuthal anisotropy of Heavy Flavor Mesons

Azimuthal anisotropy of hadrons provides information about their production with respect to the reaction plane (Ψ_n). At high p_{T} , the larger in-medium path-length of the mother partons emitted in the direction of the reaction plane leads to a stronger suppression of the yield due to jet quenching. Therefore, measurements of the v_n coefficients from Fourier expansion of the particle distributions $dN/d\psi$ are sensitive to path length dependence of the parton energy loss. At low p_{T} , a large v_2 (elliptic flow) signal is considered as evidence for collective hydrodynamical expansion of the medium. Measurements of heavy flavor meson v_n could provide important information about the thermalization of the heavy quarks in the medium. Precision measurements of v_n as a function of heavy flavor meson p_{T} could teach us how the azimuthal anisotropy of the light flavor partons contribute to the observed anisotropy through the recombination of heavy quarks and light quarks. As shown in Figure 4, the predicted elliptic flow (v_2) signal covers a large range of values due to the difference in the treatments of in-medium parton transport and parton energy loss. Figure 5 shows the expected performance of elliptic flow (v_2) measurements with the data recorded in 2015 and the projected performance in 2018 and beyond. The new precise data will be able to constrain various components of the theoretical models to determine the heavy quark diffusion coefficient in QGP, and to reveal flavor dependence of the parton energy loss.

2.3 D^0-Jet and D^0-hadron correlations

CMS has published jet-hadron and jet shape and dijet missing transverse momentum measurements which showed a significant modification of the angular structure of the jets. A significant amount of energy is trans-

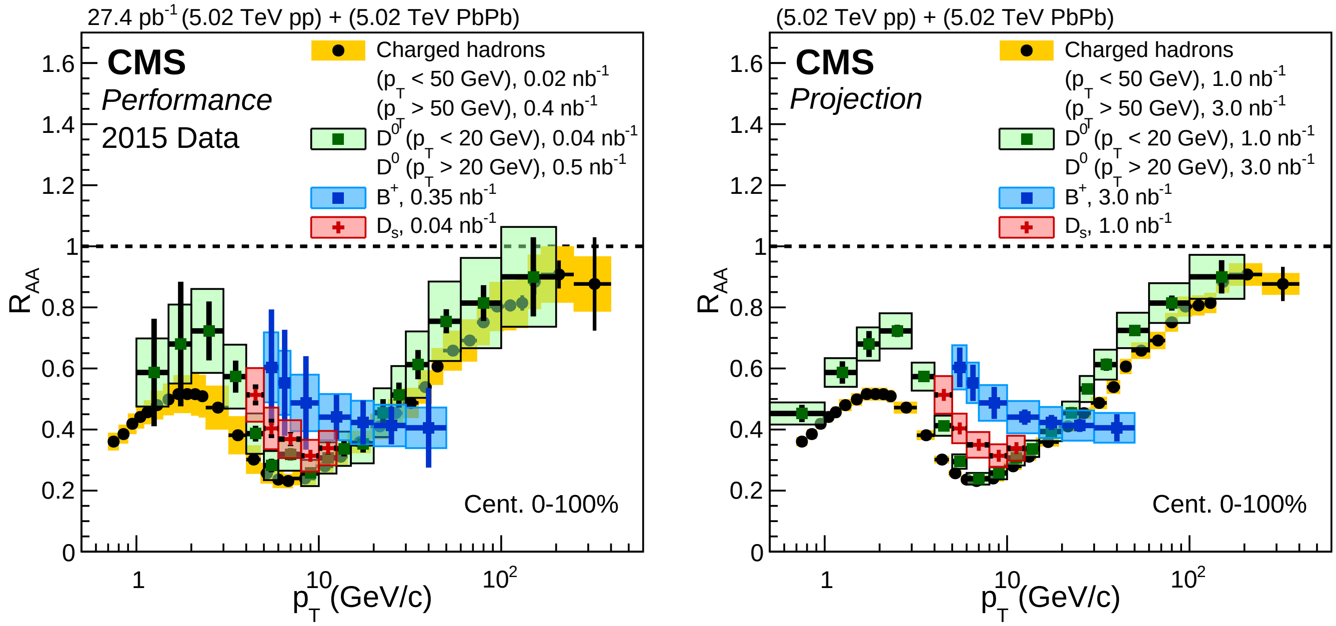

Figure 3: Nuclear modification factors of charged particles, D^0, D_s and B^+ mesons with 2015 data (left panel) and the statistics expected with L1 trigger rate upgrade (right panel).

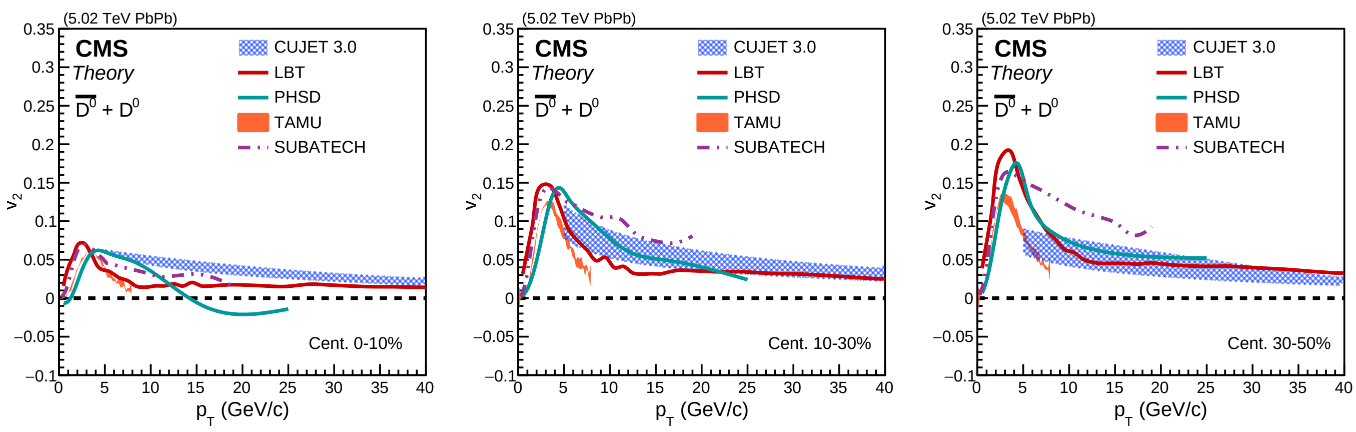

Figure 4: Theoretical calculations of heavy flavor meson v_2 in 0-10% (left) 10-30% (middle) and 30-50% (right) PbPb collisions at 5.02 TeV.

ferred to very large angle with respect to the jet axis. Studies of D^0-jet and D^0-hadron angular correlation are sensitive to the energy loss mechanism. Due to the large mass of the c-quark, the proposed measurement could also potentially distinguish models that explain the out-of-cone radiation as signals of medium response

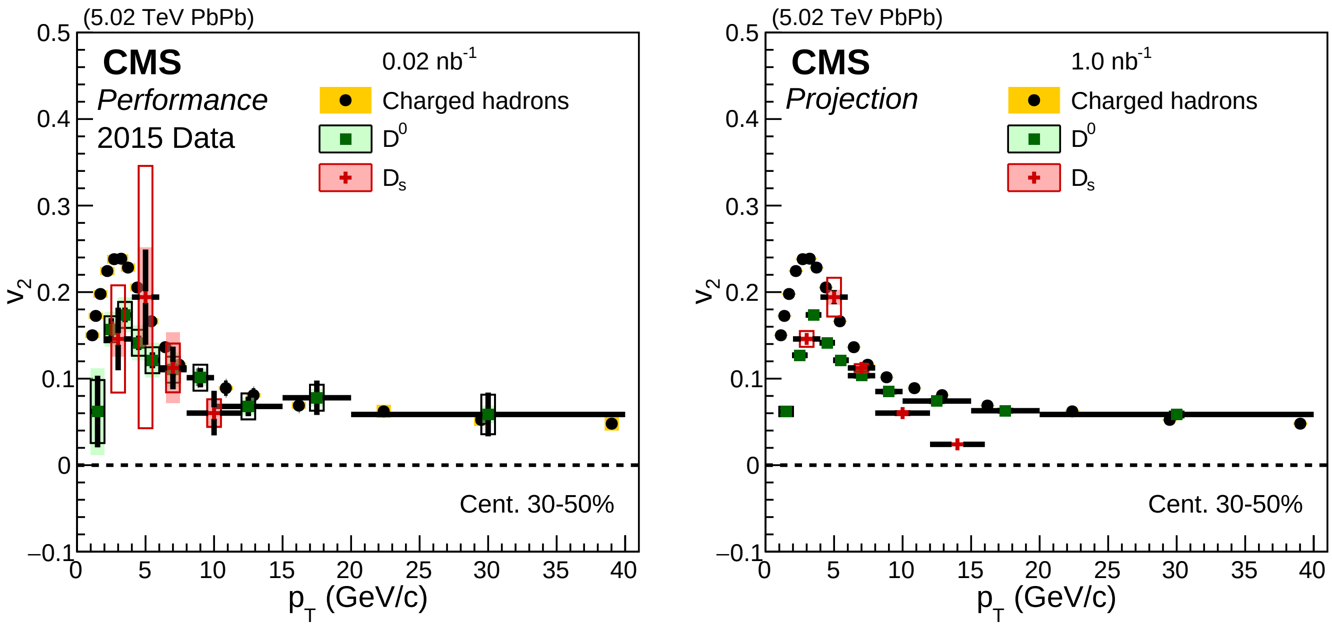

Figure 5: v_2 of charged particles, D^0 and D_s with 2015 data (left panel) and the statistics expected with L1 trigger rate upgrade (right panel).

from the ones that describe the data with medium-induced semi-hard gluon radiations. Figure 6 shows the performance of D^0 meson reconstruction as a function of the distance with respect to the jet axis ΔR. With the high statistics data taken in 2018 with jet trigger, high precision D^0-Jet correlation and D^0-hadron correlation could be performed for the first time (Figure 7).

2.4 Summary

The current and projected physics performance of various measurements are shown in Table 1. With the proposed work on the L1 trigger rate upgrade project, the kinematic range covered by CMS could be greatly increased and many more observables that are not yet accessible with 2015 data will become feasible for the first time with the data taken in 2018 and beyond.

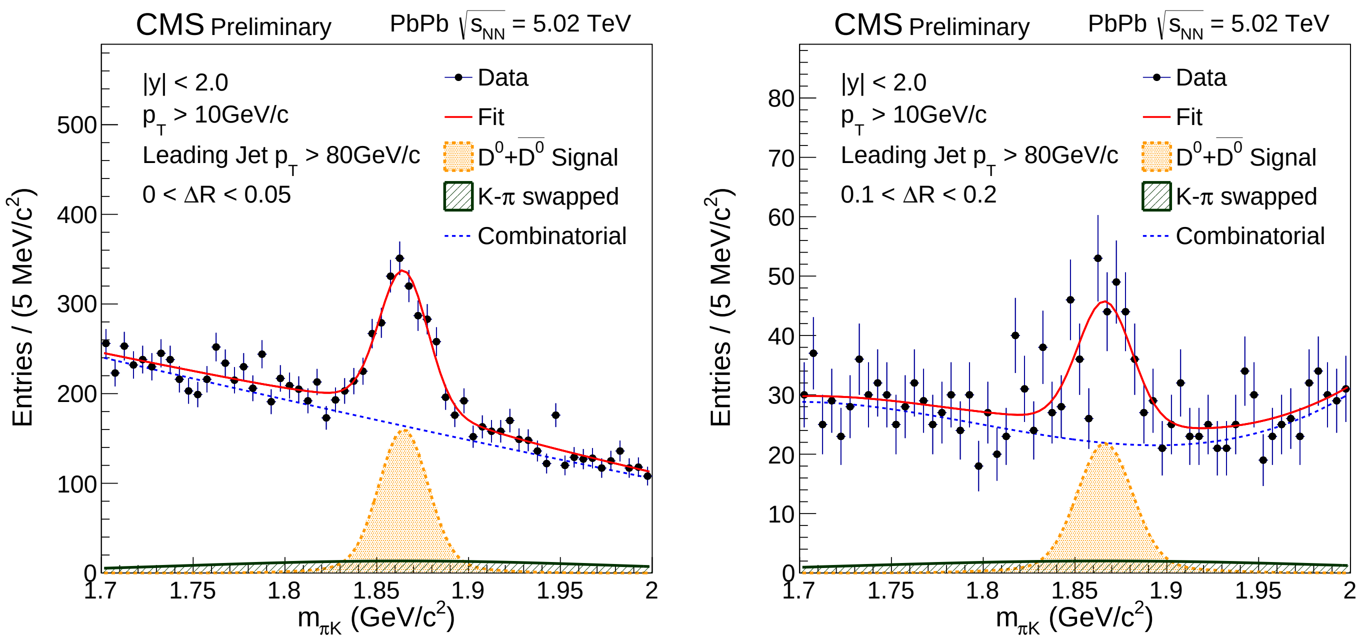

Figure 6: D^0 signal extracted in bins of distance with respect to the jet axis ΔR.

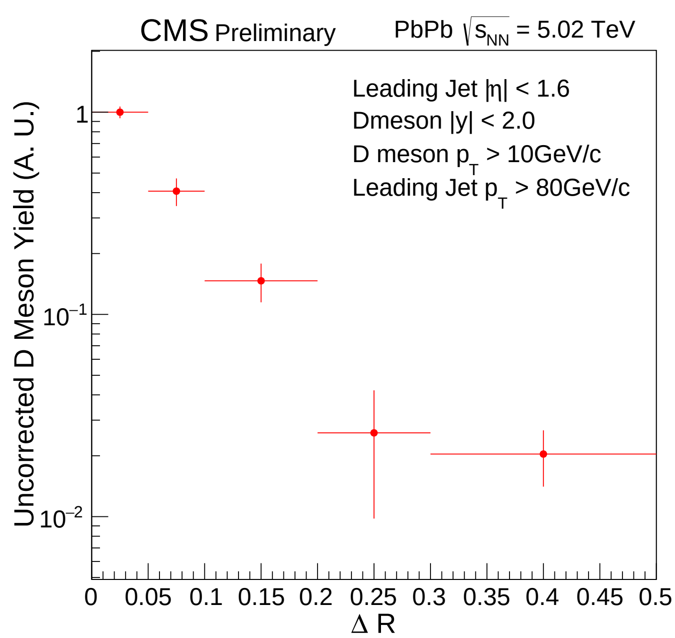

Figure 7: Extracted D^0 yield as a function of ΔR.

	Current 0.04 nb^{-1} + Present System		**2018 Pb-Pb 1 nb^{-1} + Upgraded System**	
Observables	p_T min	Statistical Uncertainty	p_T min	Statistical Uncertainty
$D^0 R_{AA}$	2	15%	0.5 - 1	10%
$D_s R_{AA}$	~ 4	20%	< 4	4%
$\Lambda_c R_{AA}$	10	$> 20\%$	< 10	4%
$B \to D R_{AA}$	6	20%	2	10%
$B^+(D^0\pi) R_{AA}$	Not accessible		~ 4	
Low p_T c and b jets	Not accessible		~ 30	
$D^0 v_2$ (= 0.06)	1	80%	0.5 - 1	18%
$D_s v_2$	Not accessible		~ 4	
$B \to D v_2$	Not accessible		~ 2	
$\Lambda_c v_2$	Not accessible		~ 6	

Table 1: Summary of the heavy flavor measurements with 2015 data and 2018 data with L1 trigger rate upgrade

3 Computing

At the time of writing it still being discussed with the LHC Program Coordinators the distributions of the luminosity that LHC could deliver among the 4 experiment. Two scenarios are considered for CMS: 1.2 nb^{-1} and 1.8 nb^{-1}. The physics program can be logically divided in two main blocks, the ẗriggered physics data collectionT̈hat is depending by the delivered luminosity and then M̈inimum Bias data collectionẗhat depends only from the LHC uptime.

The overall run will last 24 days (576 h) with 21 days (504 h) for delivering luminosity. The most probable scenario foresees and LHC uptime of 50 %. In order to collecte the requested six bilions Minimum bias events sample, a rate to disk of 6.5 kHz is needed. The physics triggered data rate is expected to be of 500 Hz for the 1.2 nb^{-1} scenario and of 750 Hz for the 1.8 nb^{-1} scenario.

The Minimum Bias event size is of 1 MB/event. Using the standard CMS RAW data content would require 6 PB of tape. In order to reduce the data volume to tape, it was envisaged the possibility for the Minimum Bias dataset to have a reduced RAW data content. In this format only the silicon pixel, silicon strip and Hadron Forward calorimeters information are kept. The reduced format allow a reduction of 35% on the event size. The default scenario for the 2018 Pb-Pb data-taking foresees that all physics triggers events will be collected with standard CMS RAW data content. A fraction of the Minimum Bias events (500 M) will be collected with standard CMS RAW data content. The remaining MB events (5.5 B) will have a reduced raw data content.

In order to optimize the tape occupancy, the products of the data reconstructions will not be saved. Only the Analysis Object Data (AOD) will be recorded on tape for both physics triggered and minimum bias datasets.

In table 3 the data volume produced considering the different scenarios of delivered luminosity are reported.

	Total RAW with MB RAW format (PB)	Total RAW with MB trk only (PB)
Luminosity 0.6 nb-1	6.6	4.7
Luminosity 1.2 nb-1	7.2	5.3
Luminosity 1.8 nb-1	7.9	6.0
	Total AOD (PB)	
Luminosity 0.6 nb-1	2.4	
Luminosity 1.2 nb-1	2.8	
Luminosity 1.8 nb-1	3.1	
	Total RAW + AOD (PB)	
Luminosity 0.6 nb-1	9.0	7.1
Luminosity 1.2 nb-1	10.0	8.1
Luminosity 1.8 nb-1	11.0	9.1
	Total RAW/day (TB)	
Luminosity 0.6 nb-1	313	223
Luminosity 1.2 nb-1	345	255
Luminosity 1.8 nb-1	377	287

The computing strat

4 Management Plan, Personnel, Schedule

5 Budget

References

[1] Roland G (CMS Collaboration) 2007 *J.Phys.* **G34** S733–736 (*Preprint* `nucl-ex/0702041`)

[2] Dokshitzer Y L and Kharzeev D 2001 *Phys.Lett.* **B519** 199–206 (*Preprint* `hep-ph/0106202`)