Data-driven efficiencies of the LHCb High Level Trigger

JAMIE GOODING
Technische Universität Dortmund

The trigger of the Large Hadron Collider beauty (LHCb) experiment underwent a major redesign between Runs 2 and 3 of the LHC. This redesign saw the Level 0 (L0) hardware trigger removed, with the High Level Trigger (HLT) receiving information at the full 30 MHz detector readout rate. Amongst the changes required to process data at this rate, were the full deployment of the first HLT level (HLT1) to GPUs and extensive use of a streamlined event model at the second HLT level (HLT2). To understand the HLT performance, efficiencies of the selection algorithms used in each stage must be calculated. These efficiencies are calculated using the data-driven TISTOS method, a tag-and-probe approach which evaluates events based on the candidates required to form the trigger decision. This contribution discusses the upgraded LHCb trigger and TISTOS approach, before examining the calculation trigger efficiencies in Run 3. In particular, the contribution focuses on recent studies on the validation of TISTOS, preliminary results from data taken by LHCb in 2024 and developments in tooling for performing these calculations.

The upgrade of the LHCb trigger removed the L0 trigger, which had been employed in Runs 1 and 2 to reduce the data rate from 30 MHz to 1 MHz for further processing by the HLT. The L0 trigger had become a bottleneck in the efficiency of selecting decay channels with hadronic final states, inhibiting increases in the yields of these channels at the higher luminosities of Run 3 and beyond. To accommodate the increase of the HLT1 input rate to 30 MHz, a dedicated software framework (Allen) was developed to run efficiently on GPUs. This framework handles the decoding of subdetector information, reconstruction of tracks and vertices, and selection of events. The full, offline-quality reconstruction of events is then performed in HLT2. The Turbo event model, in which only the candidates relevant to the trigger decision are saved, is used to reduce event sizes and thus increase the number of events which can be saved.

The efficiency of a selection is typically defined as the fraction of events passing the selection. However, this is not possible for trigger efficiencies, as only events selected by the trigger are recorded. Instead, the TISTOS method was developed as a data-driven approach to calculating trigger efficiencies, in which two samples are defined. Events requiring the signal candidate for a trigger decision are labelled as "triggered on signal" (TOS). Events requiring any other reconstructed candidate in the event for a trigger decision are labelled "triggered independent of signal" (TIS). The subsample in which events are labelled both TIS and TOS is then known as the TISTOS subsample. Efficiencies can be calculated from these subsamples by a tag-and-probe method, taking TIS and TISTOS events as tag and probe, respectively.

The TISTOS approach has been applied to calculate trigger efficiencies in key channels of interest in data taken by LHCb in 2024. These channels cover decays of b-hadrons to final states containing muons, electrons and light hadrons, and of c-hadrons to final states containing light hadrons. Efficiencies in 2024 data are evaluated against equivalent efficiencies in Run 2 and in Monte Carlo simulated samples of decays in conditions representative of 2024 data-taking. Additionally, a software package, `TriggerCalib`, has been developed to provide a central tool for calculating trigger efficiencies from the TISTOS information in samples.