LHCb Topological Trigger Reoptimization

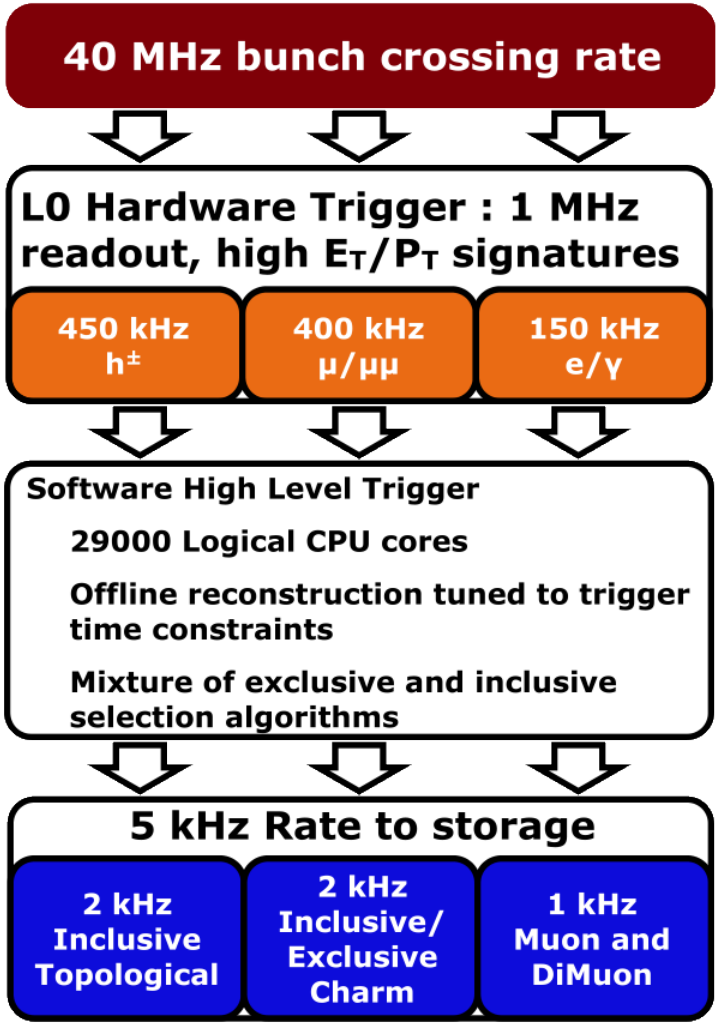
Philip Iten[1], Tatiana Likhomanenko[2,3],
Egor Khairullin[2], Andrey Ustyuzhanin[2,3], Mike Williams[1]

[1] Massachusetts Institute of Technology, US
[2] Yandex School of Data Analysis, RU
[3] NRC "Kurchatov Institute", RU

April 14, 2015

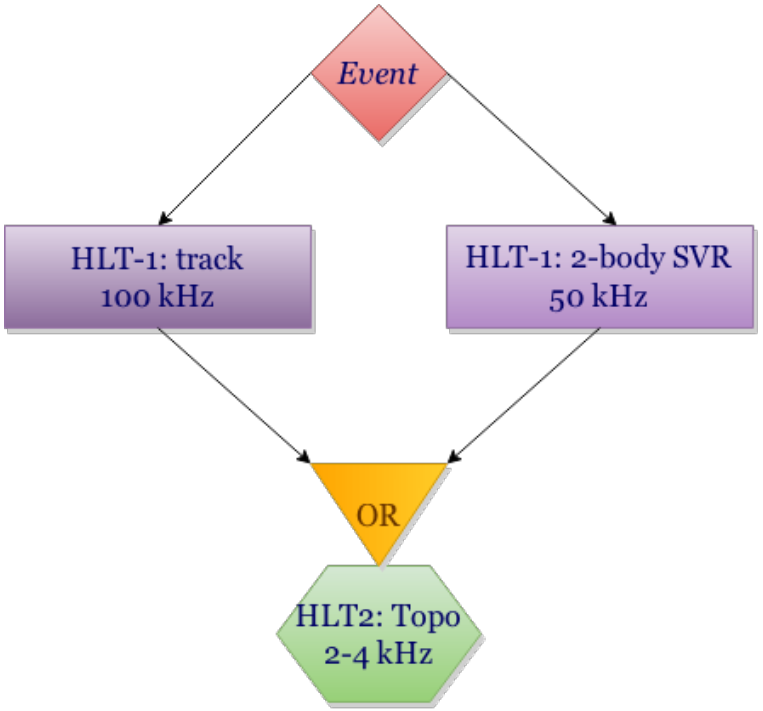
What is Topological Trigger?

- Generic trigger for decays of beauty and charm hadrons
- It designed to be inclusive trigger line to efficiently select any B decay with at least 2 charged daughters
- Look for 2, 3, 4 track combinations in a wide mass range
- Designed to efficiently select decays with missing particles
- Use fast-track fit to improve signal efficiency and minbias rejection



Goal: improve topological trigger efficiency for Run-2

Run-2 HLT Scheme



What tells us an event contains interesting physics?

- A combination of displacement from PV and high PT

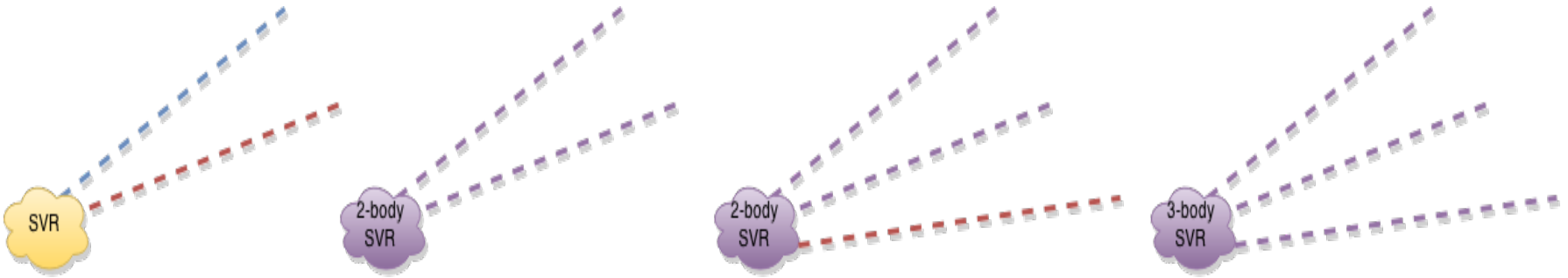
Run-2 strategy:

- HLT-1 track is looking for either one super high PT or high displacement track
- HLT-1 2-body SVR classifier is looking for two tracks making a vertex
- HLT-2 improved topo classifier uses full reconstructed event to look for 2, 3, 4 and more tracks making a vertex

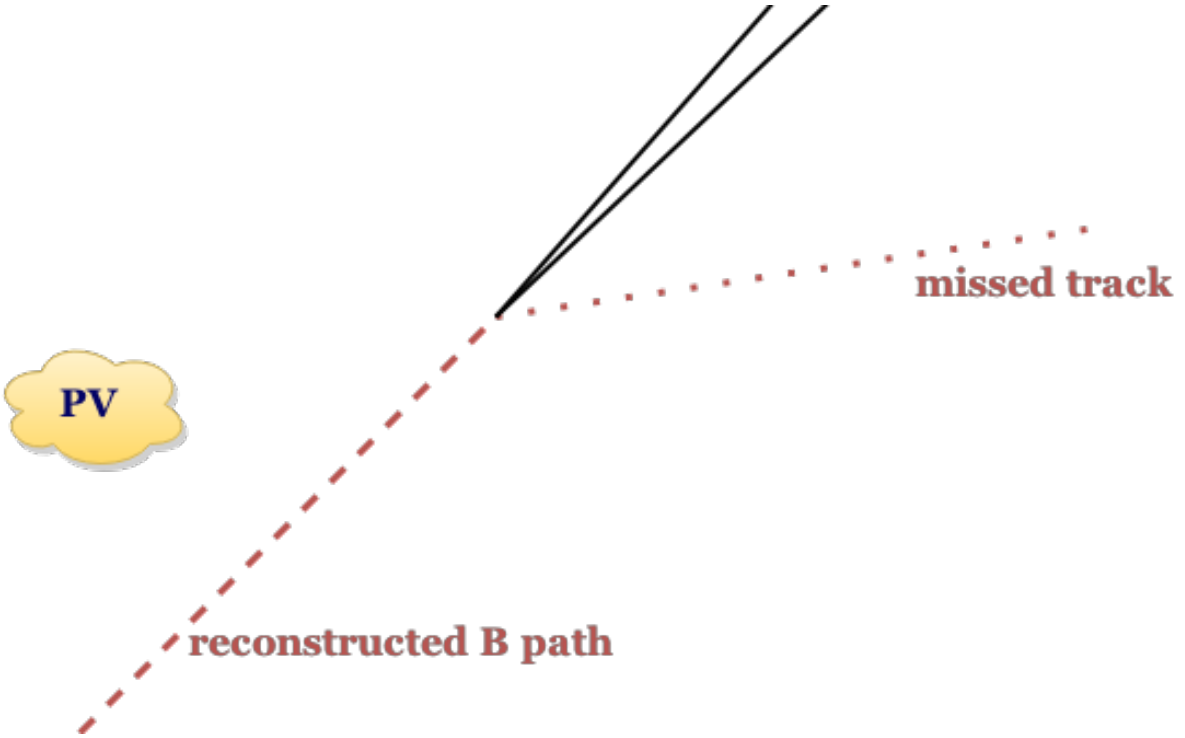
NOTE: tracking thresholds are quite different in Run-1 and Run-2

N-body Tracks

- Two, three or four tracks are combined to form a SVR
- Each secondary vertex in Monte Carlo data is preselected in such way, that all tracks must be matched to particles from the signal decay (true match preselection)

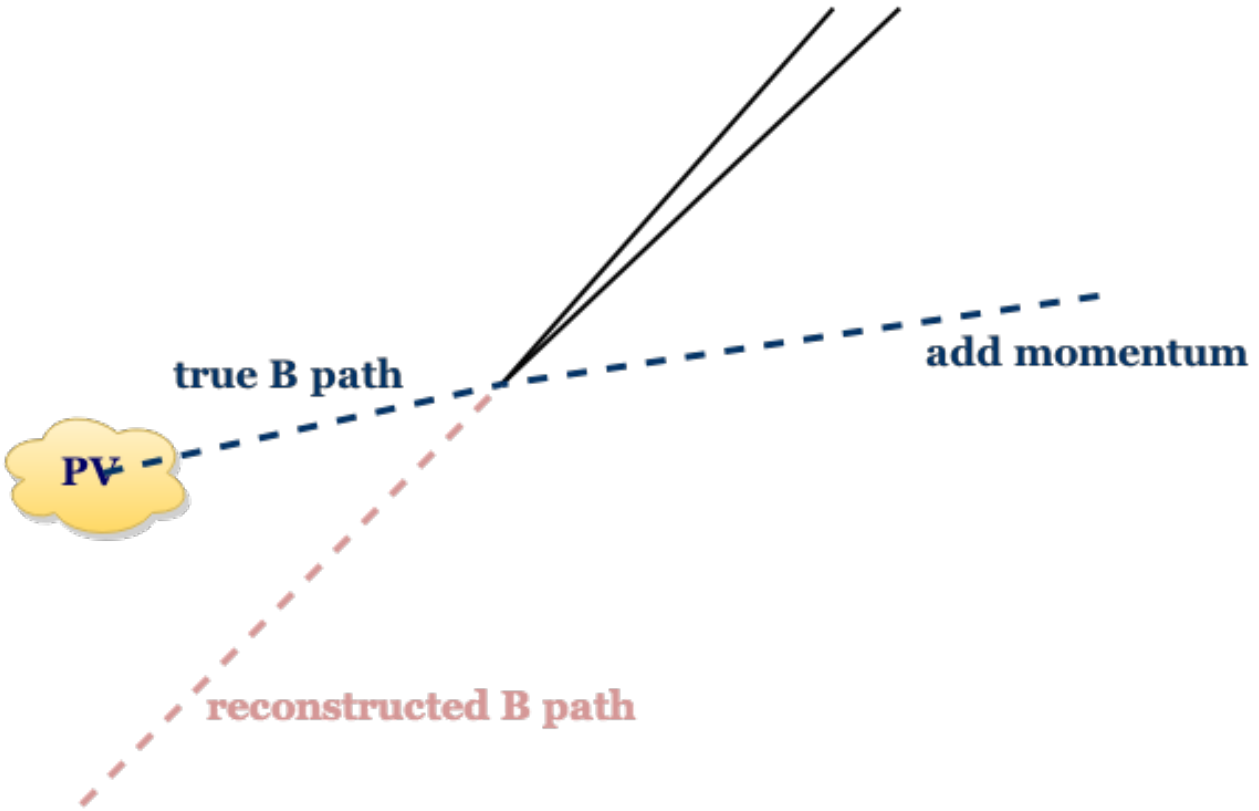


Ommision of Daughters



The trigger is designed to allow for the omission of one or more daughters when forming the trigger candidate.

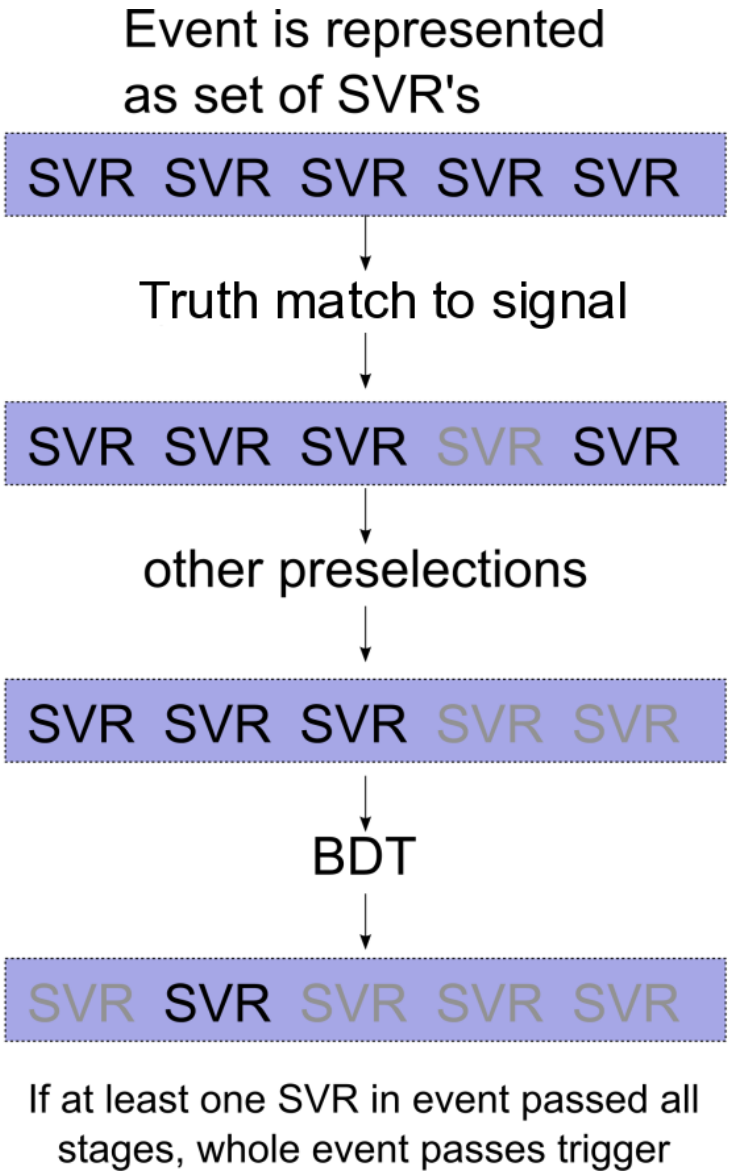
Ommision of Daughters



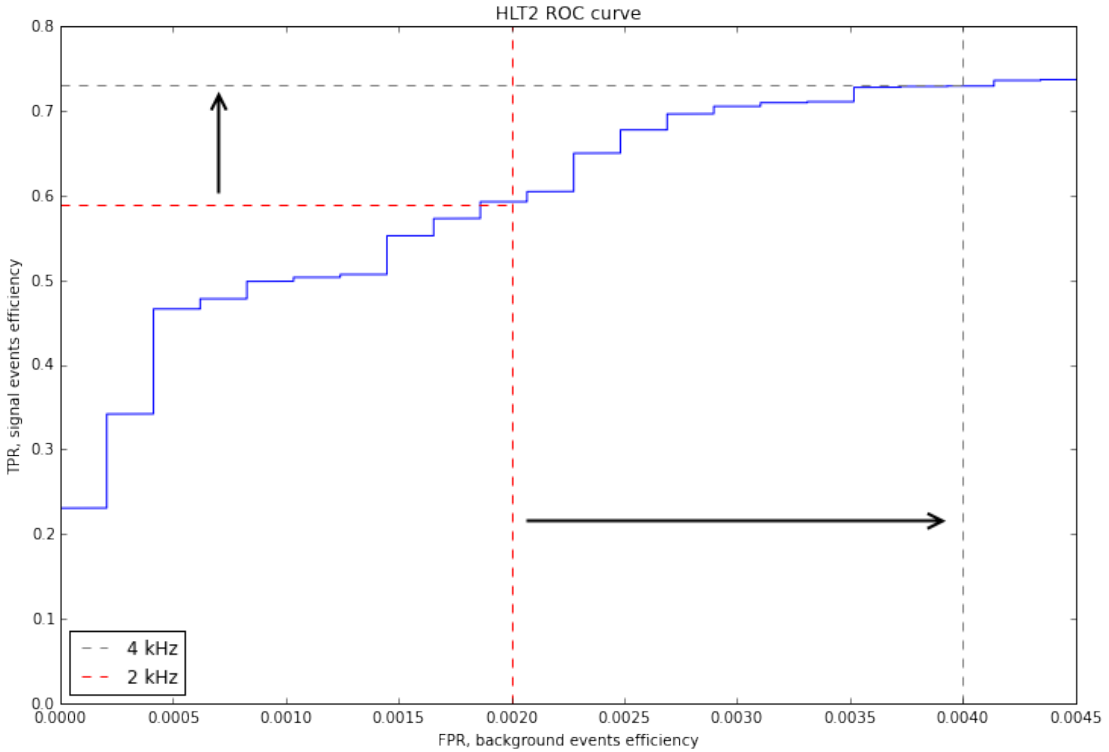
$$m_{corrected} = \sqrt{m^2 + \left|p'_{Tmissing}\right|^2} + \left|p'_{Tmissing}\right|$$

Machine Learning Specific Problem: data structure

- Signal samples are simulated 13-TeV B decays of various topologies
- Background sample is generic Pythia 13-TeV proton-proton collisions
- Most events have many secondary vertices SVRs (not all events have an SVR)

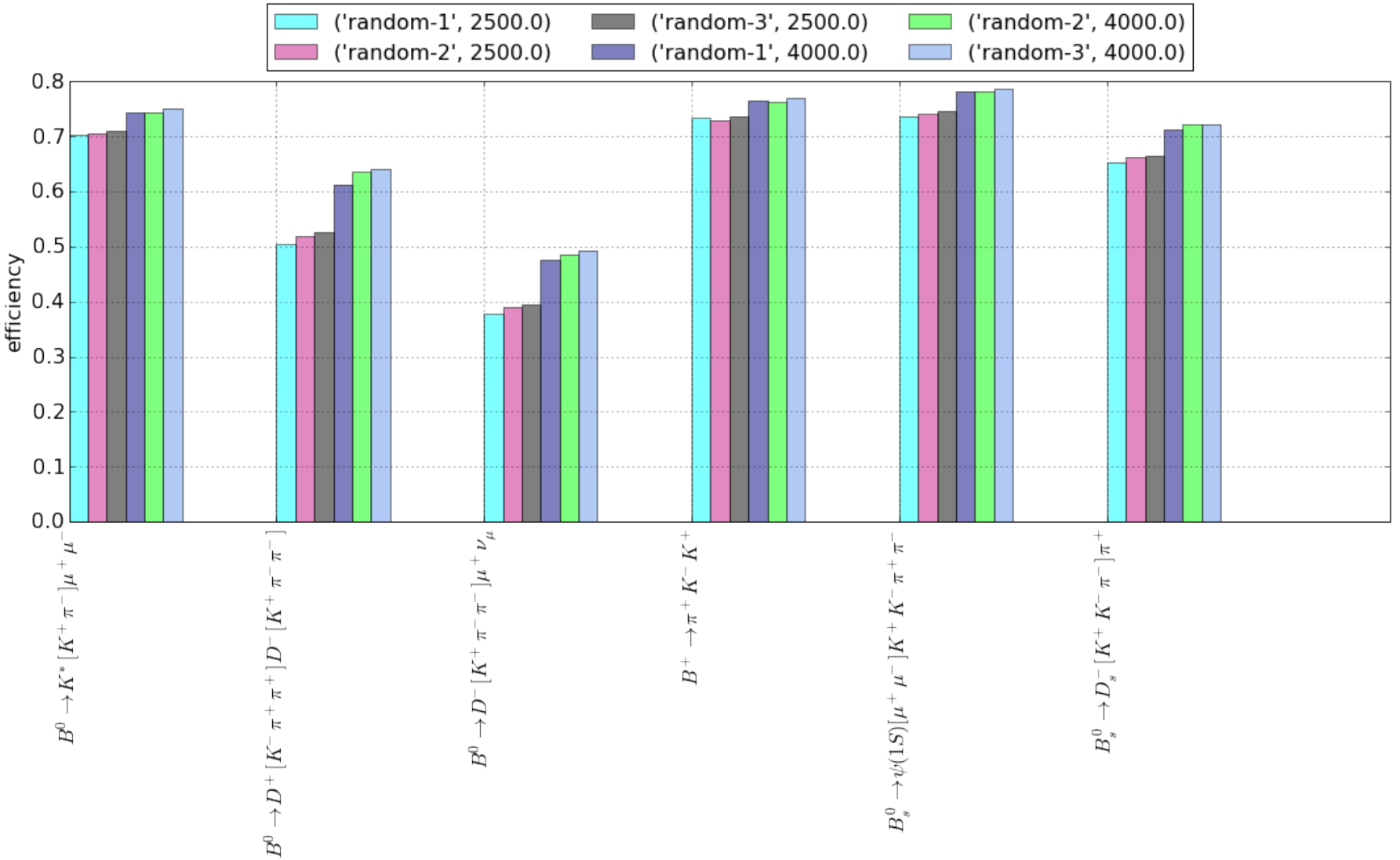


Machine Learning Specific Problem: FOM

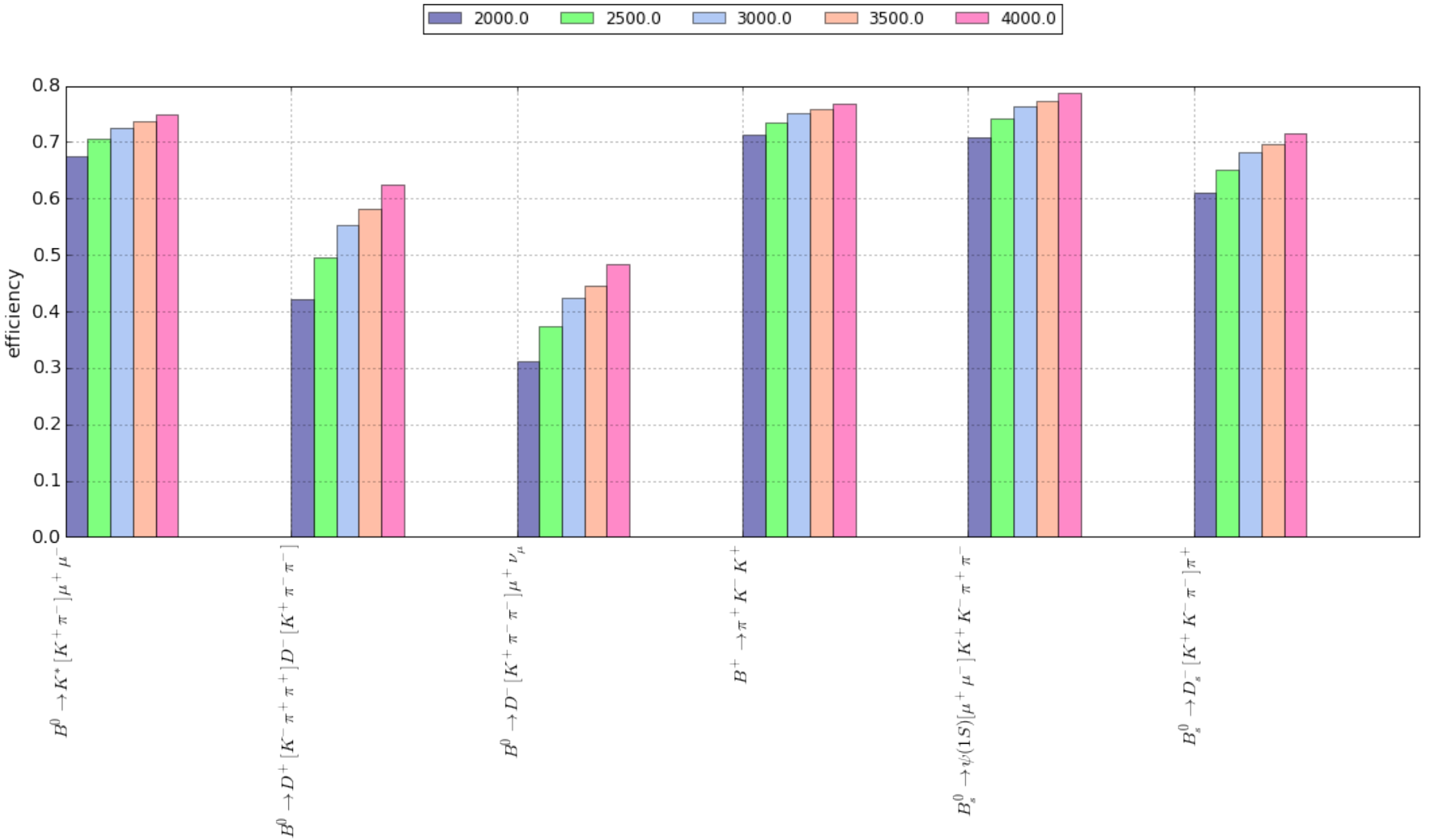


- FOM is the over all efficiency, calculated for passed events, not SVRs
- Output rate must be limited
- Restriction is imposed on background events efficiency FPR = 0,2% (corresponds to 2 kHz)

HLT2: threshold unstability



HLT2: efficiency vs output rate



Online Processing: BBDT vs Post-prunning

Bonsai BDT (BBDT):

- Used in Run-1 for online processing
- Features hashing before training by yourself
- Convert decision trees to n-dimentional table making it essentially infinitely fast
- Predict operation takes one reading from this table

But:

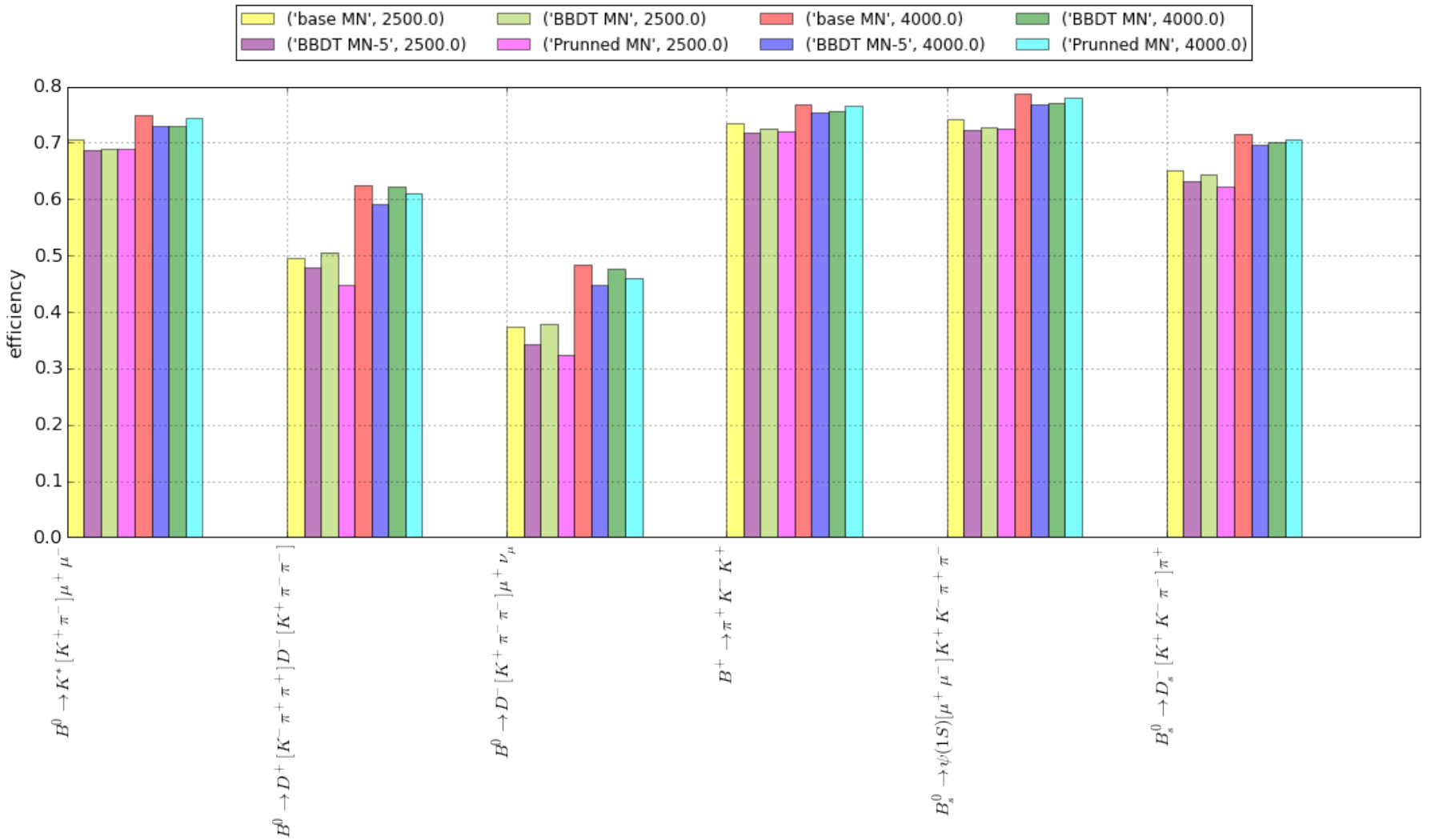
- We are limited in the table size (or count of bins for each feature)
- Discretization reduces efficiency

MatrixNet (MN)

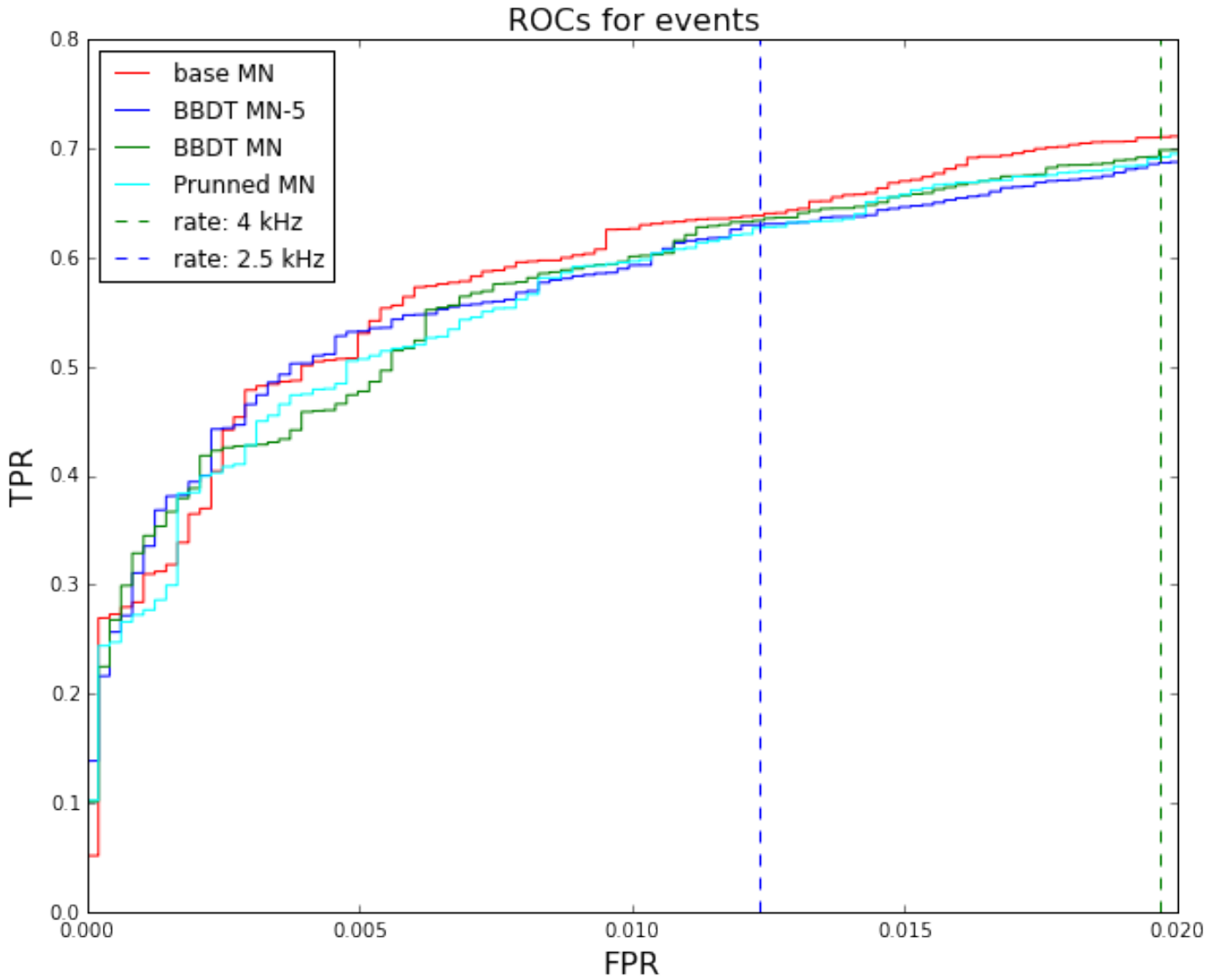
post-prunning:

- Another strategy for online processing
- Features also hashing with amount count of bins for each variable
- Post-prunning of the decision trees to speedup prediction operation (less count of trees)
- Online predict event by all trees

BBDT vs Post-prunning Efficiencies



BBDT vs Post-prunning Efficiencies: ROCs



Current Status: Run-1 vs Run-2

Ratio of Run-2 over Run-1 for HLT2/HLT1 efficiencies

mode	2.5 kHz	4. kHz
$B^0 \to K^*[K^+\pi^-]\mu^+\mu^-$	1.64	1.72
$B^+ \to \pi^+K^-K^+$	1.59	1.65
$B_s^0 \to D_s^-[K^+K^-\pi^-]\mu^+\nu_\mu$	1.14	1.47
$B_s^0 \to \psi(1S)[\mu^+\mu^-]K^+K^-\pi^+\pi^-$	1.62	1.71
$B_s^0 \to D_s^-[K^+K^-\pi^-]\pi^+$	1.46	1.52
$B^0 \to D^+[K^-\pi^+\pi^+]D^-[K^+\pi^-\pi^-]$	1.40	1.86

Note that the denominator is reconstructible with $PT(B) > 2$ GeV, $\tau(B) > 0.2$ ps.

Summary

1. New HLT scheme in Run-2: sophisticated HLT1 (classifier) and HLT2-Topo
2. Overall (HLT2/HLT1) efficiency improvement: 15-60% for 2.5 kHz (50-80% for 4 kHz) vs Run-1
3. Timing comparison of MatrixNet BBDT vs post-pruning is in progress
4. Looking forward to data taking!

Thank you for attention!

Backup

HLT1-track and HLT1 2-body SVR preselections

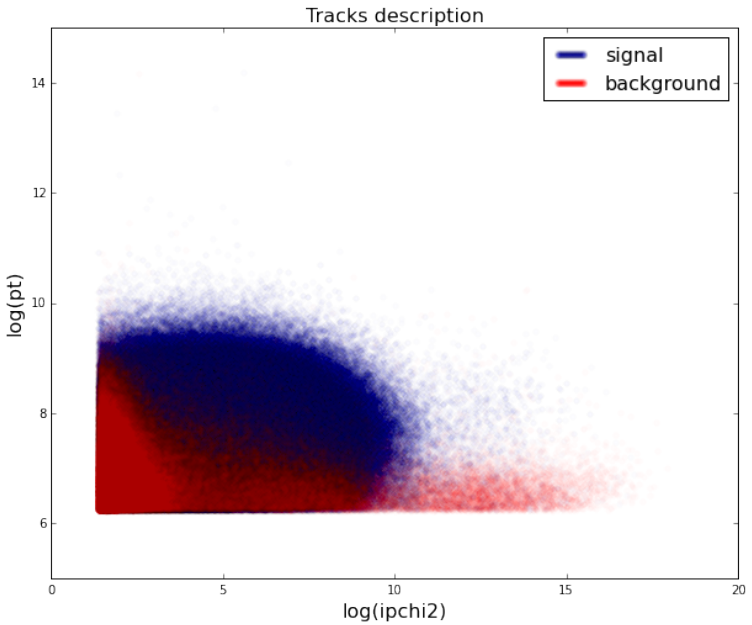
1 track:

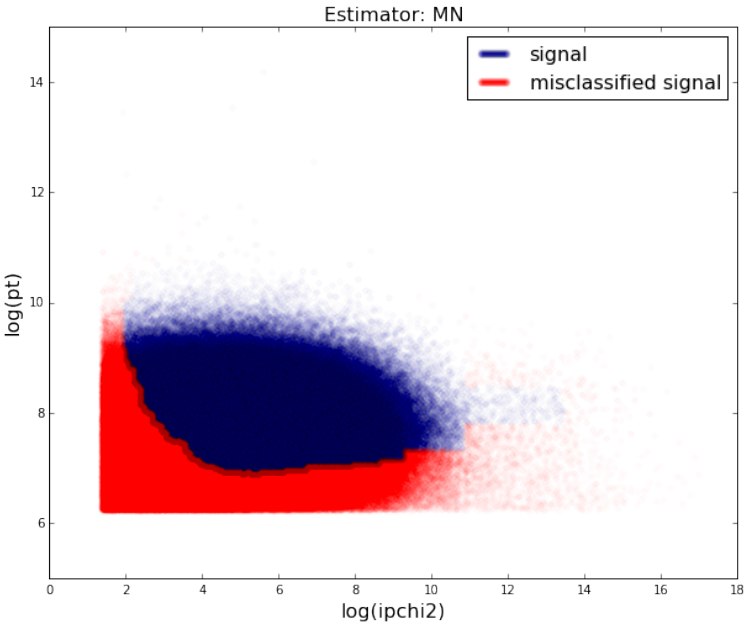
- Tracks preselections:
  - $PT > 500$ MeV;
  - $IP_{\chi^2} > 4$;
  - $track_{\chi^2}/ndof < 3$;
- BDT uses PT, IP_{χ^2}
- Output rate 100 kHz

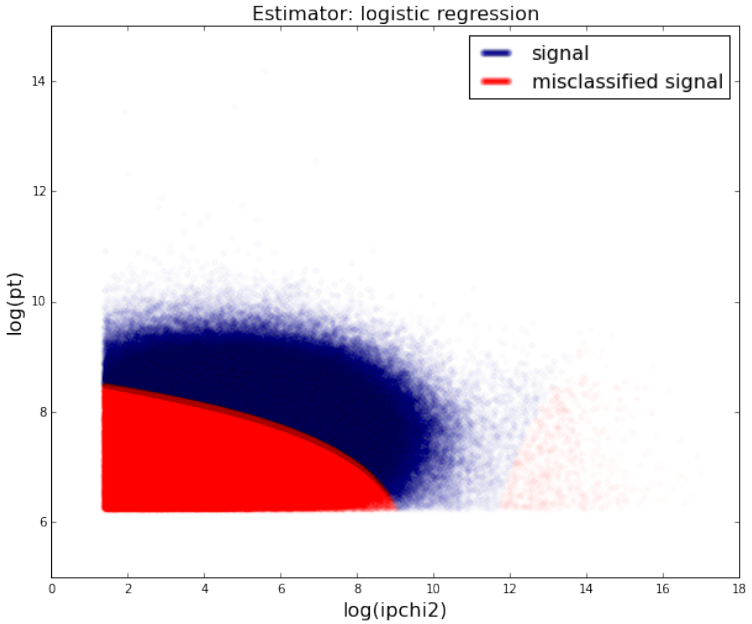
2-body SVR:

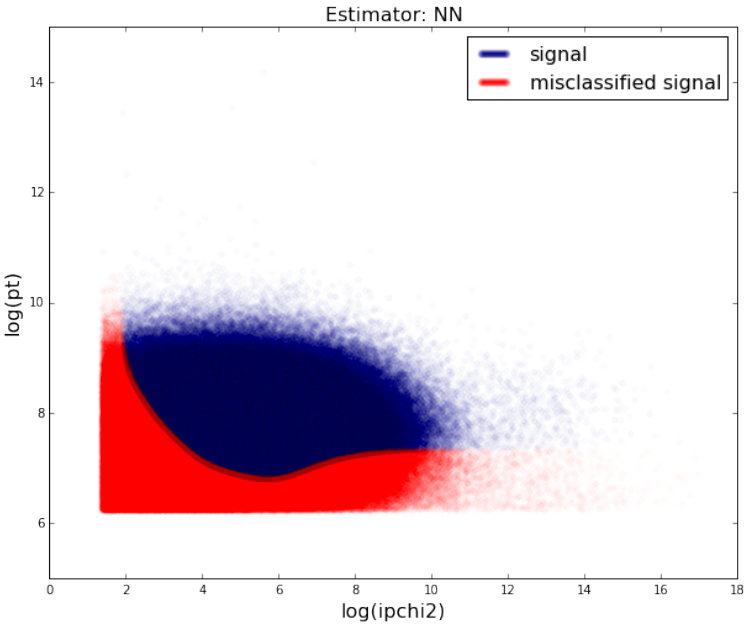
- Tracks preselections:
  - $PT > 500$ MeV;
  - $IP_{\chi^2} > 4$;
  - $track_{\chi^2}/ndof < 2.5$;
- SVR preselections:
  - $PT > 2$ GeV;
  - $vertex_{\chi^2} < 10$;
  - $1 < MCOR$ GeV;
  - $2 < \eta < 5$ (PV to SVR)
- Don't use MCOR in BDT (from a systematics perspective)
- BDT variables: sum PT, $vertex_{\chi^2}$, FD_{χ^2}, N(tracks with $IP_{\chi^2} < 16$)
- Output rate 50 kHz

HLT1-track: decision boundary

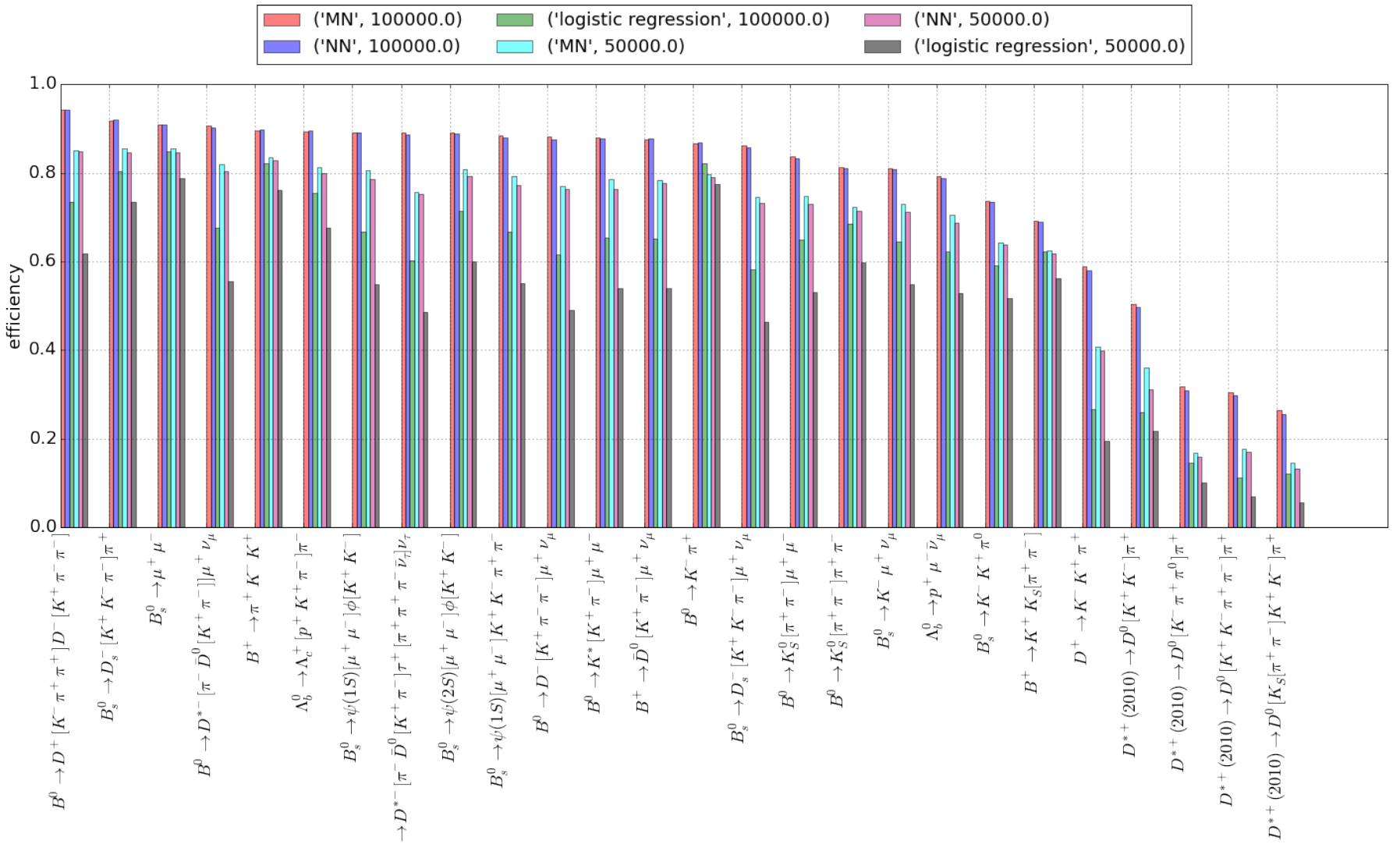




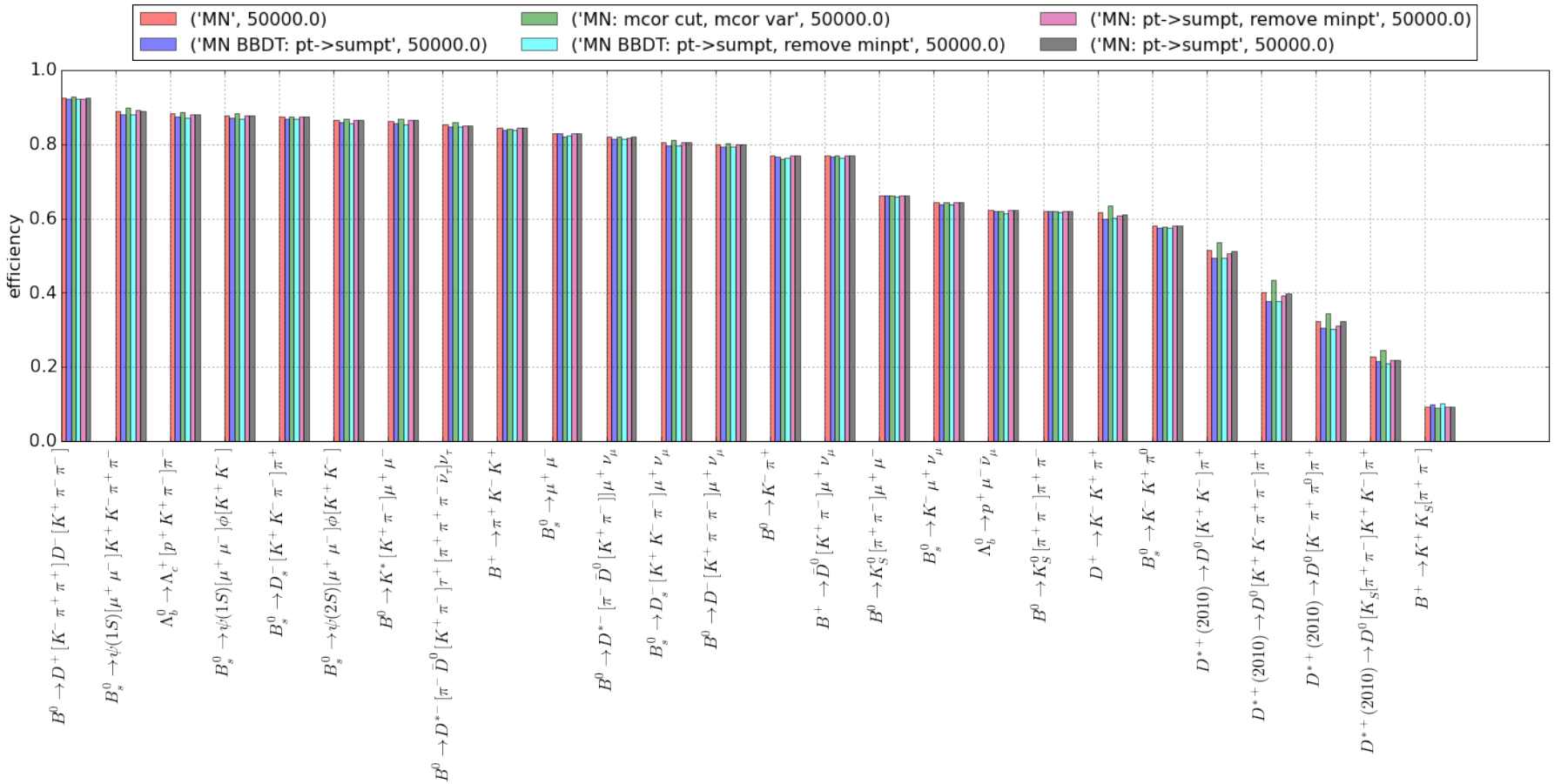




HLT1-track: MN vs NN vs logistic



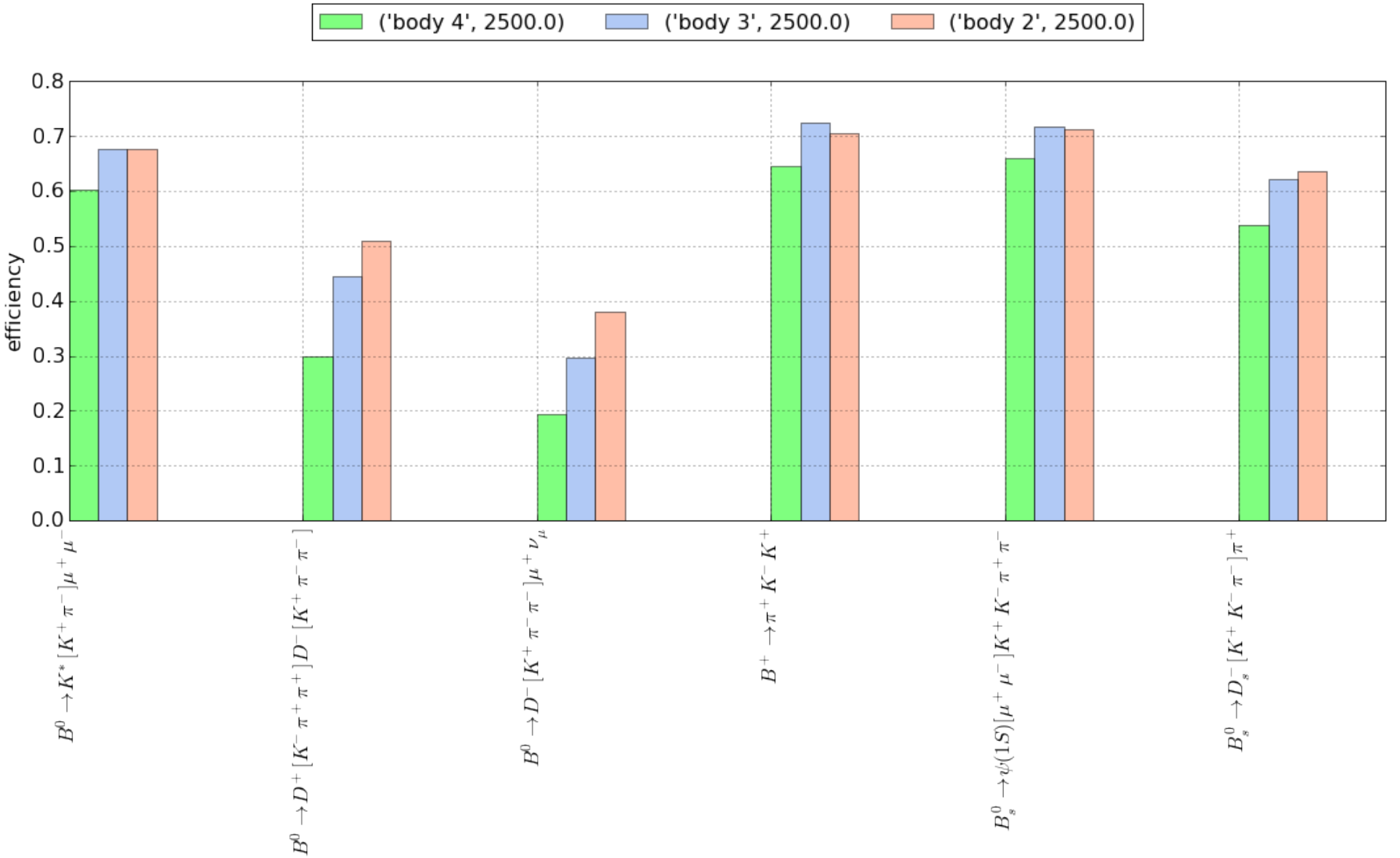
HLT1-SV


('MN', 50000.0)
('MN BBDT: pt->sumpt', 50000.0)
('MN: mcor cut, mcor var', 50000.0)
('MN BBDT: pt->sumpt, remove minpt', 50000.0)
('MN: pt->sumpt, remove minpt', 50000.0)
('MN: pt->sumpt', 50000.0)
efficiency
1.0
0.8
0.6
0.4
0.2
0.0
$B^0 \to D^+ [K^- \pi^+ \pi^+] D^- [K^+ \pi^- \pi^-]$
$B_s^0 \to \psi(1S) [\mu^+ \mu^-] K^+ K^- \pi^+ \pi^-$
$\Lambda_b^0 \to \Lambda_c^+ [p^+ K^+ \pi^-] \pi^-$
$B_s^0 \to \psi(1S) [\mu^+ \mu^-] \phi [K^+ K^-]$
$B_s^0 \to D_s^- [K^+ K^- \pi^-] \pi^+$
$B_s^0 \to \psi(2S) [\mu^+ \mu^-] \phi [K^+ K^-]$
$B^0 \to K^* [K^+ \pi^-] \mu^+ \mu^-$
$B^0 \to D^{*-} [\pi^- \bar{D}^0 [K^+ \pi^-] \tau^+ [\pi^+ \pi^+ \pi^- \bar{\nu}_\tau] \nu_\tau$
$B^+ \to \pi^+ K^- K^+$
$B_s^0 \to \mu^+ \mu^-$
$B^0 \to D^{*-} [\pi^- \bar{D}^0 [K^+ \pi^-]] \mu^+ \nu_\mu$
$B_s^0 \to D_s^- [K^+ K^- \pi^-] \mu^+ \nu_\mu$
$B^0 \to D^- [K^+ \pi^- \pi^-] \mu^+ \nu_\mu$
$B^0 \to K^- \pi^+$
$B^+ \to \bar{D}^0 [K^+ \pi^-] \mu^+ \nu_\mu$
$B^0 \to K_S^0 [\pi^+ \pi^-] \mu^+ \mu^-$
$B_s^0 \to K^- \mu^+ \nu_\mu$
$\Lambda_b^0 \to p^+ \mu^- \bar{\nu}_\mu$
$B_s^0 \to K_S^0 [\pi^+ \pi^-] \pi^+ \pi^-$
$D^+ \to K^- K^+ \pi^+$
$B_s^0 \to K^- K^+ \pi^0$
$D^{*+}(2010) \to D^0 [K^+ K^-] \pi^+$
$D^{*+}(2010) \to D^0 [K^+ K^- \pi^+ \pi^-] \pi^+$
$D^{*+}(2010) \to D^0 [K^- \pi^+ \pi^0] \pi^+$
$D^{*+}(2010) \to D^0 [K_S [\pi^+ \pi^-] K^+ K^-] \pi^+$
$B^+ \to K^+ K_S [\pi^+ \pi^-]$

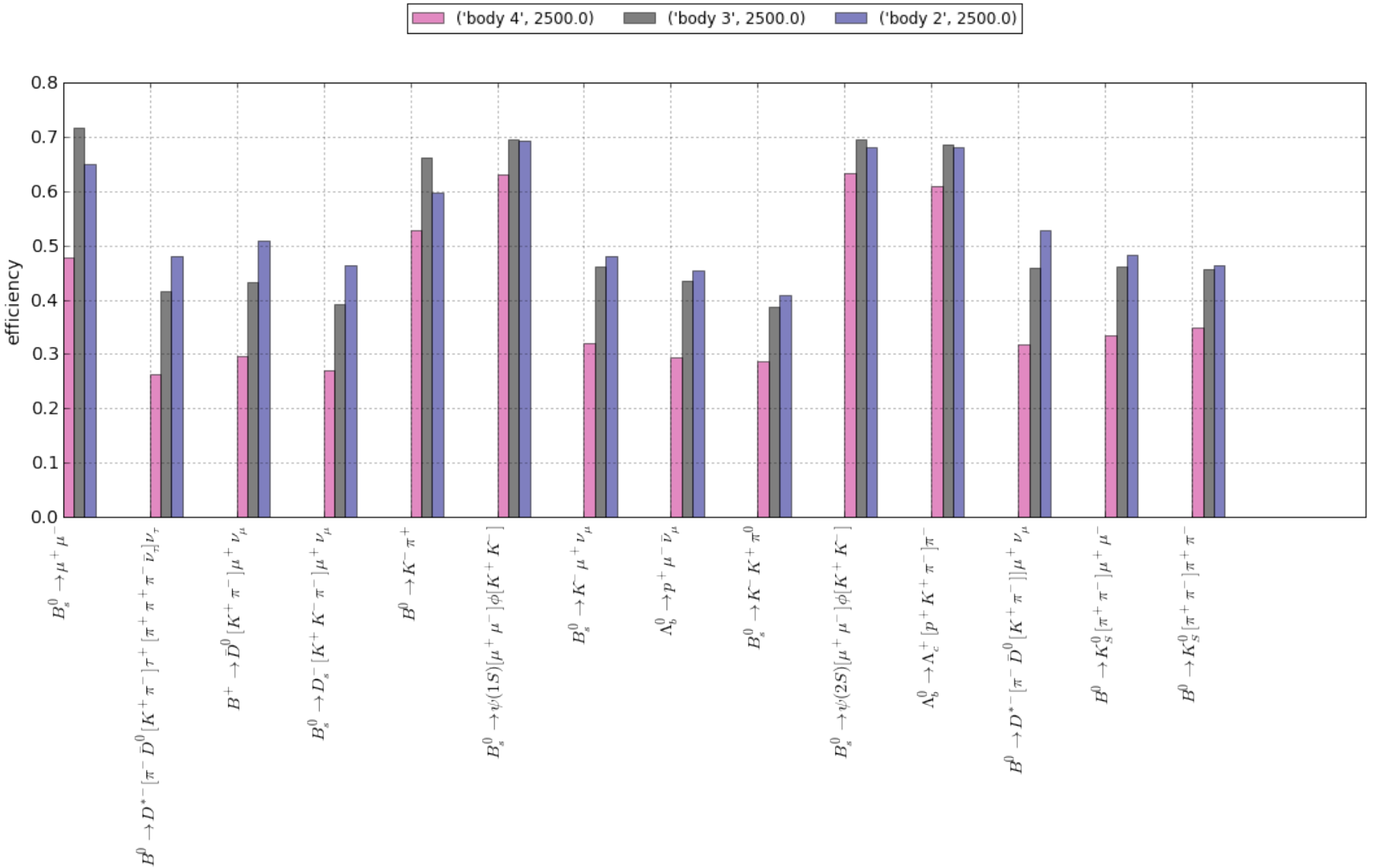
HLT2 preselections

- The same preselections as for 2-body SVR
- Changed track $PT > 200$ MeV
- Added $MCOR < 10$ GeV
- Added $N(\text{tracks with } IP_{\chi^2} < 16) < 2$
- Used any min PT
- BDT variables: n, $MCOR$, sum PT, $vertex_{\chi^2}$, η, FD_{χ^2}, min PT, IP_{χ^2}, $N(\text{tracks with } IP_{\chi^2} < 16\ )$, $N(tracks)$
- Output rate 2-4 kHz

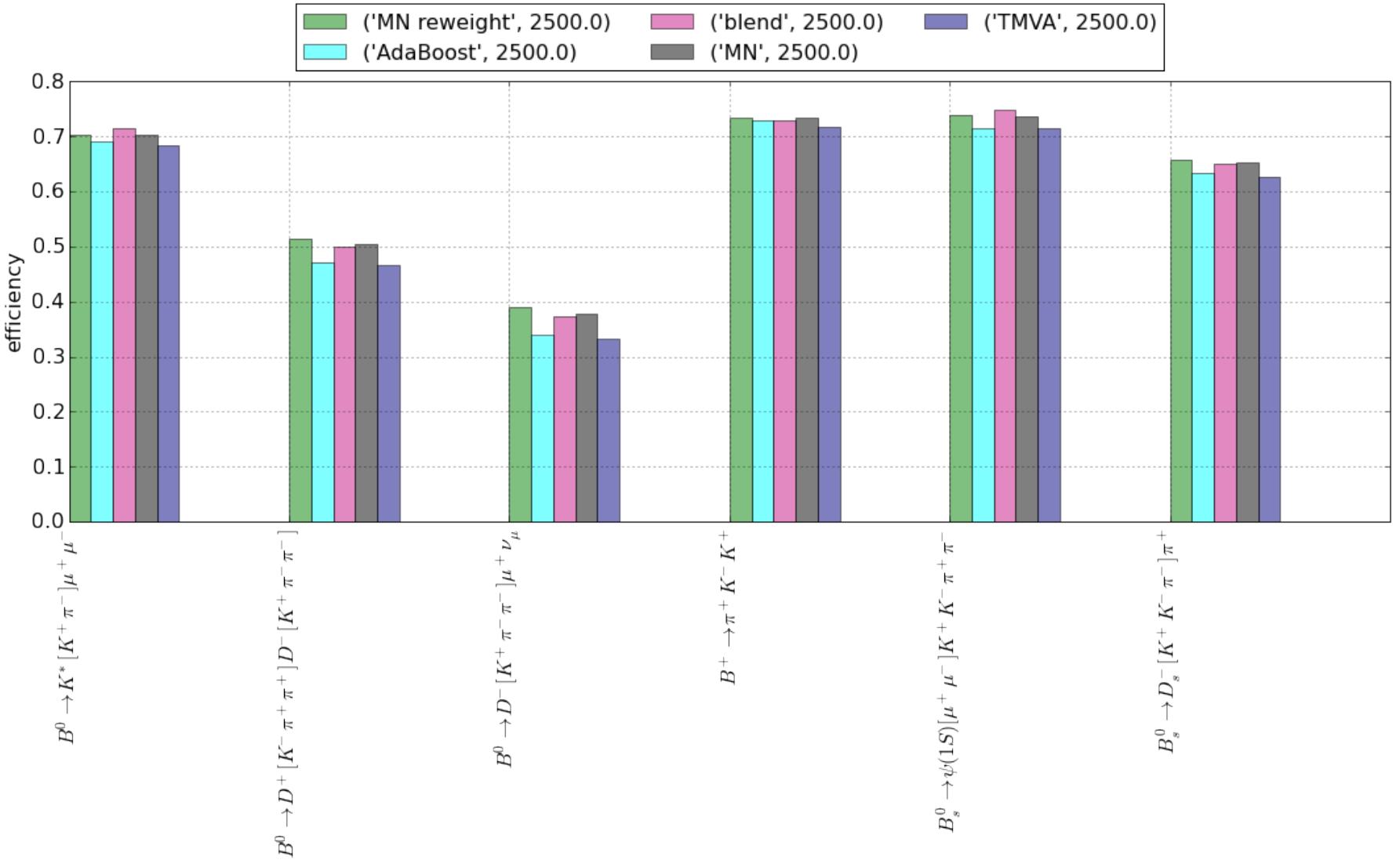
HLT2: n-bodies comparison



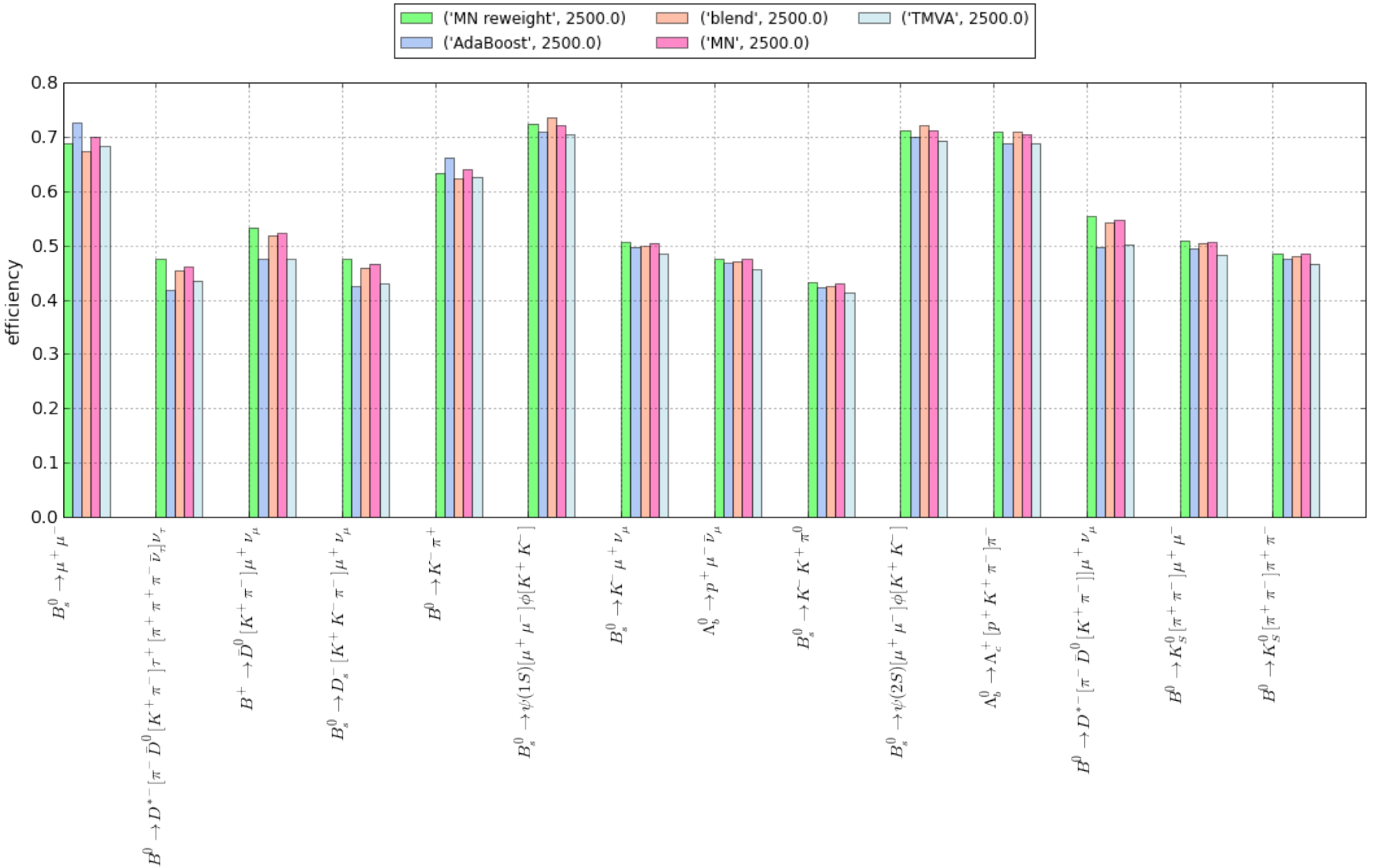
HLT2: n-bodies comparison for other modes



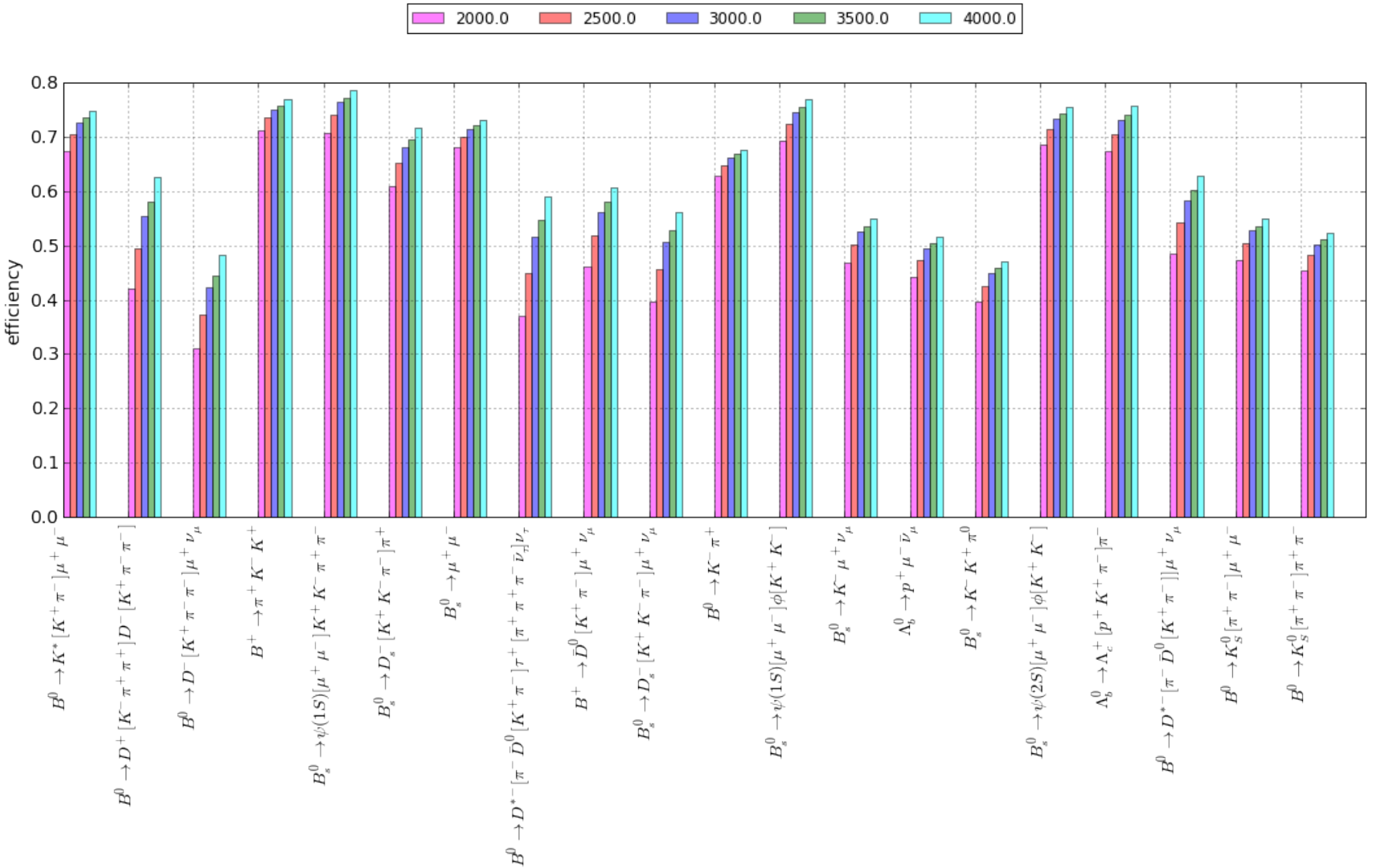
HLT2: models comparison



HLT2: models comparison for other modes



HLT2: efficiency vs output rate for other modes



BBDT vs Post-prunning efficiencies for other modes