Karlsruhe Institute of Technology

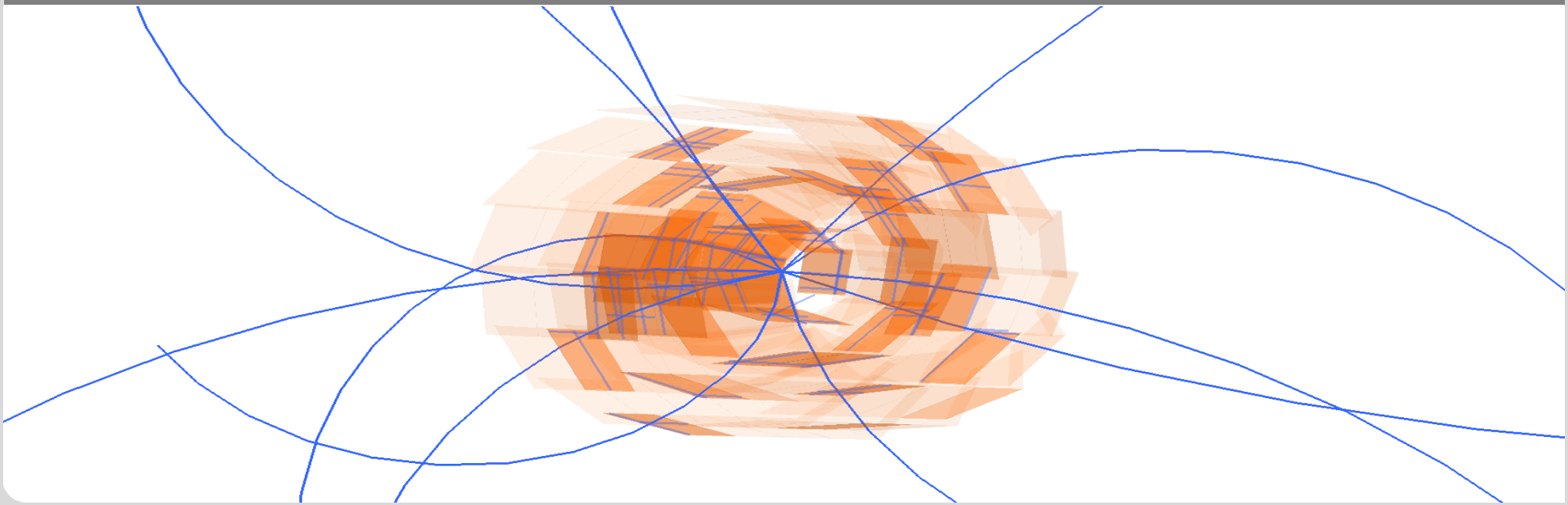
Belle II

VXDTF2 MVA QE: Looking for the Hit Efficiency

Online Tracking Meeting

Sebastian Racs | 23rd February 2018

INSTITUT FÜR EXPERIMENTELLE TEILCHENPHYSIK (ETP)



What's up with the Hit Efficiency

- Drop of the VXDTF2's Hit Efficiency by 5 % reported by Eugenio
- Was traced back to the activation of the VXDTF2's MVA QualityEstimation, but also happens before
- Problem was not observed originally when the VXDTF2 MVA was developed
- Hit Efficiency was not checked anymore directly before activation

$\Rightarrow$ Sub-optimal combination of 2 simultaneous changes

Figures of merit

MVA	Find. Eff.	Hit Eff.	Hit Purity	Fake Rate	Clone Rate
✗	0.9199	**0.8990**	0.9674	0.06558	0.00036
✓	0.9314	**0.8611**	0.9684	0.05930	0.00059

Simulation and evaluation

- On master state bbe0a3b1 (13.02.18)
- 15k $\Upsilon(4S)$ events with official phase 3 Bkg overlay 15th Campaign
- MVA with default weight (without timing) from master

Drop observed on current state

- Ca. 4 % drop in **Hit** Efficiency, 1 % increase in **Finding** Efficiency

$\rightarrow$ Why was this not seen before?

Figures of merit

Subsets	MVA	Find. Eff.	Hit Eff.	Hit Purity	Fake Rate	Clone Rate
✓	✗	0.9199	0.8990	0.9674	0.06558	0.00036
✓	✓	0.9314	0.8611	0.9684	0.05930	0.00059
✗	✗	0.8783	**0.9046**	0.9661	0.06794	0.00019
✗	✓	0.8811	**0.9045**	0.9691	0.06239	0.00019

Current state vs. during development

- Module `AddVXDTrackCandidateSubSets` was introduced and activated as part of the 2 Step Candidate Selection shortly before the MVA
- Reduces Hit Efficiency while increasing Finding Efficiency
- MVA amplifies this effect

→ What about training a new weight file?

Figures of merit

Subsets	MVA	Find. Eff.	Hit Eff.	Hit Purity	Fake Rate	Clone Rate
✓	✗	0.9199	0.8990	0.9674	0.06558	0.00036
✓	✓	0.93**14**	**0.8611**	**0.9684**	0.05930	0.00059
✗	✗	0.8783	0.9046	0.9661	0.06794	0.00019
✗	✓	0.8811	0.9045	0.9691	0.06239	0.00019
✓	new train.	0.93**45**	**0.7712**	**0.9899**	0.06047	0.00099

Training new Sample

- Trained with 100k $\Upsilon(4S)$ events with official phase 3 Bkg overlay 15th Campaign
- Trades off a very high Hit **Purity** for a big further drop in Hit **Efficiency**

Efficiencies by p_t Profile

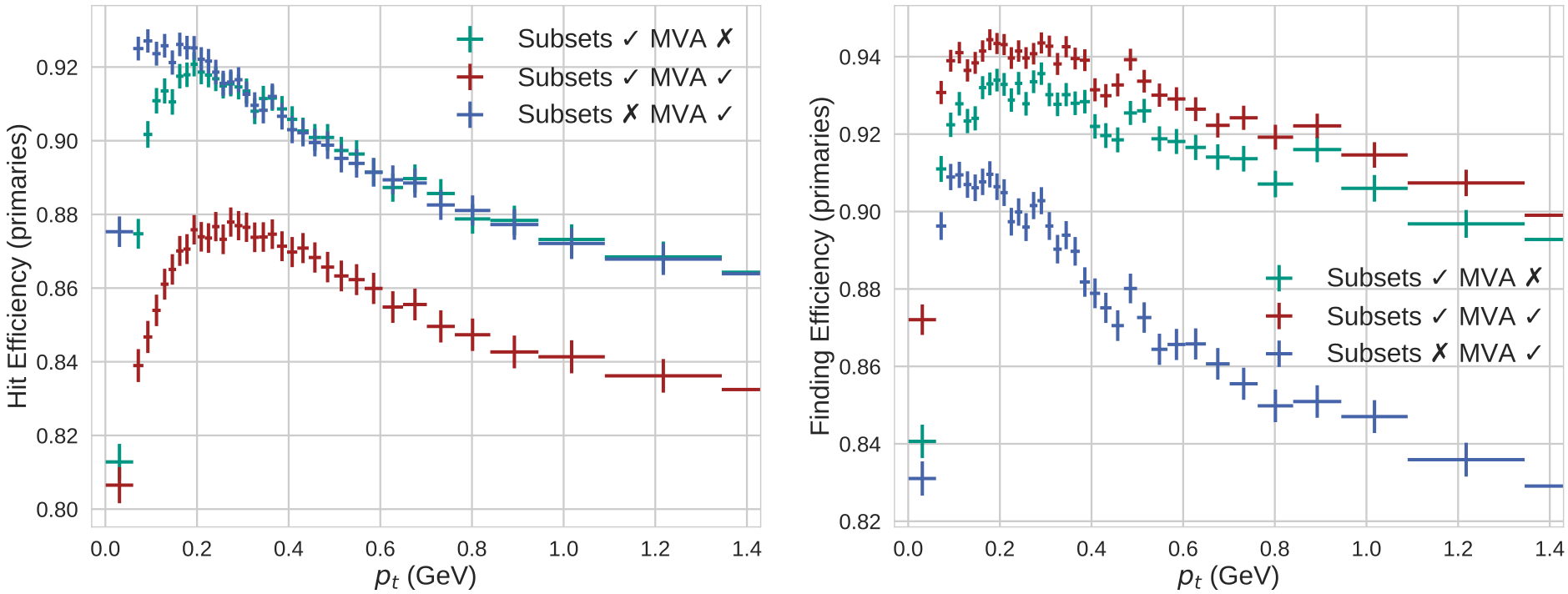


Figure: Hit Efficiency and Finding Efficiency by p_t Profile

- Using the Subsets and MVA together (current default) gives the best Finding Efficiency but worst Hit Efficiency over the whole p_t range

Reminder: Cutting on the QualityIndicator

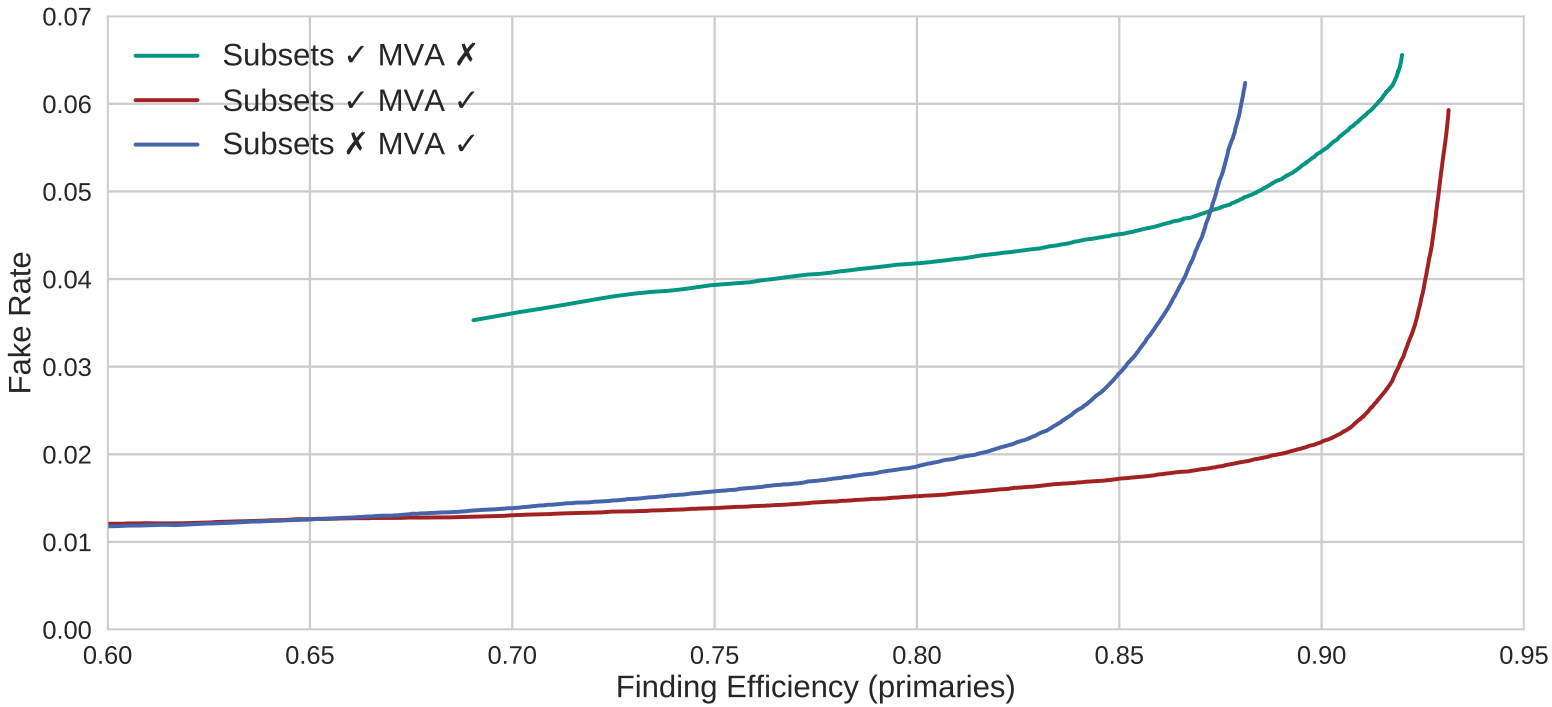


Figure: Fake Rate vs. Finding Efficiency for Cuts on the QualityIndicator

Discussion

- What trade-off do we want between finding efficiency, hit efficiency and hit purity?
- Is there another/better way to resolve clusters overlaps in VXDTF2?
- It might be possible to recover clusters that should be part of 2 true tracks $\Rightarrow$ I will have a look at this
- Complete/Final CKF Setup not yet active $\Rightarrow$ will probably find long tracks in SVD

$\Rightarrow$ We can do long-term meaningful studies of the MVA methods once the setup is stable