HCAL Reconstruction: MC Correction Functions Update

Edmund Berry

Thursday, April 24, 2014

BROWN

Introduction

- Alexandre's talk describes a method for OOT PU corrections on data
- We would like to apply the same method for Monte Carlo
- Procedure:
  - Run Alexandre's ratio method on zero PU MC
  - Derive correction functions based on the pulse shape
  - Use the same definitions, fits, and methods as in data
  - Validate results on MC with OOT PU
- More details on the following slides

GEN-SIM datasets

- Consider two GEN-SIM datasets (no PU) at T1_US_FNAL:

Dataset	Production release
/MinBias_TuneZ2star_13TeV-pythia6/Summer13-START53_V7C-v1/GEN-SIM	CMSSW_5_3_10_patch2
/QCD_Pt-1800_TuneZ2star_13TeV_pythia6/Fall13-POSTLS162_V1-v1/GEN-SIM	CMSSW_6_2_0_patch1

- QCD_Pt-1800 dataset:
  - DAS link
  - 93453 ($\sim$ 100k) events, 95 files
  - HcalNoiseAnalyzer ntuples on FNAL EOS: /eos/uscms/store/user/eberry/QCD1800MC/
- MinBias dataset:
  - DAS link
  - 9999424 ($\sim$ 10M) events, 946 files
  - HcalNoiseAnalyzer ntuples on FNAL EOS: /eos/uscms/store/user/eberry/MinBiasMC/

No pileup processing (OLD and DONE):

- Need to process datasets to get `DIGI` and `RECO`
- Steps needed:

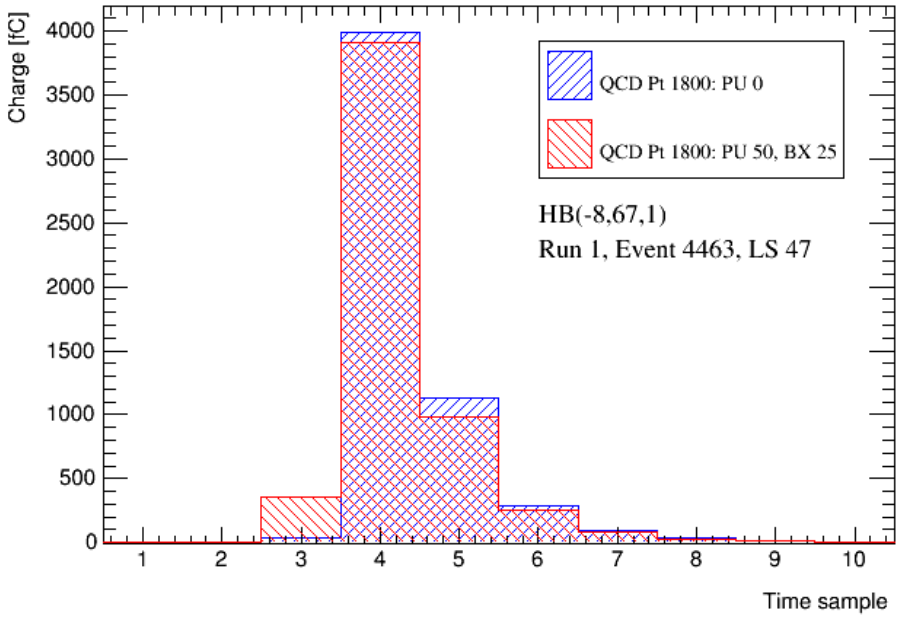
  `DIGI`, `L1`, `DIGI2RAW`, `HLT`, `RAW2DIGI`, `L1Reco`, `RECO`
- Then run `HcalNoiseAnalyzer` (updated for 62X)
  - HcalNoiseAnalyzer git page, Maintained by noise group?
  - Updated .cc file for 62X, E. Berry
- Use `CMSSW_6_2_8` to process `QCD_Pt-1800` dataset:
  - cmsDriver.py command
  - Final python cfg

With pileup processing (NEW and COMING)

- Need to overlay `QCD` with `MinBias`
- Use `MixingModule` in `CMSSW_6_2_8`
- Pileup scenario: `AVE_50_BX_25ns`
- Two stages:
  - 1) `DIGI`, `L1`, `DIGI2RAW`, `HLT`
  - 2) `RAW2DIGI L1Reco RECO`
- Stage 1: Done
  - cmsDriver.py command
  - Final python cfg
- Stage 2: Subset done
  - cmsDriver.py command
  - Final python cfg
- High PU is VERY CPU intensive: 2 minutes/event

Pileup vs. No Pileup pulse shape comparison: HB

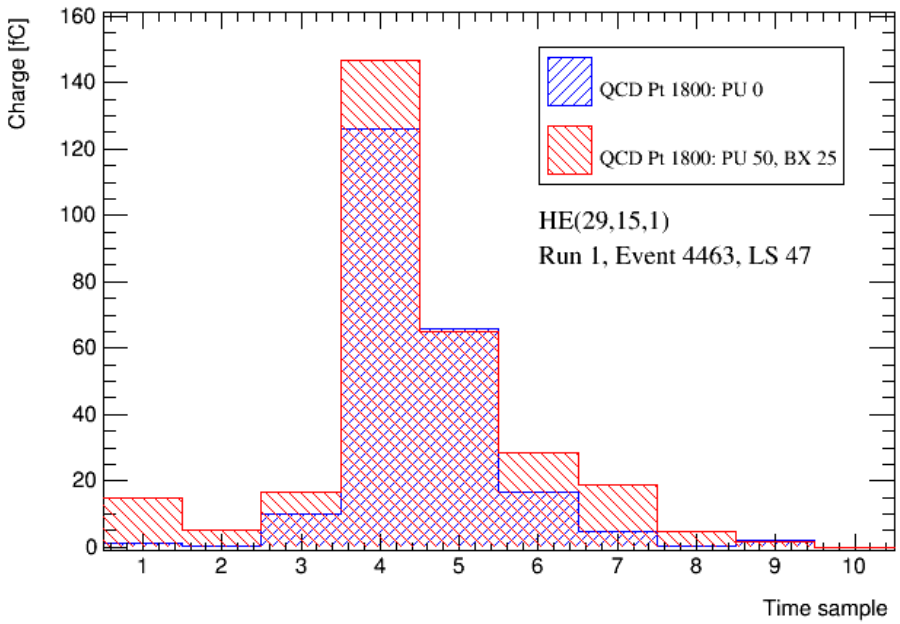
Example single DIGI comparison: HB



- As expected in TS3. Strangeness in TS4 + TS5.
- Bug in `MixingModule`? Investigating.
- More detailed results coming (HCAL DPG meeting)

Pileup vs. No Pileup pulse shape comparison: HE

Example single DIGI comparison: HE



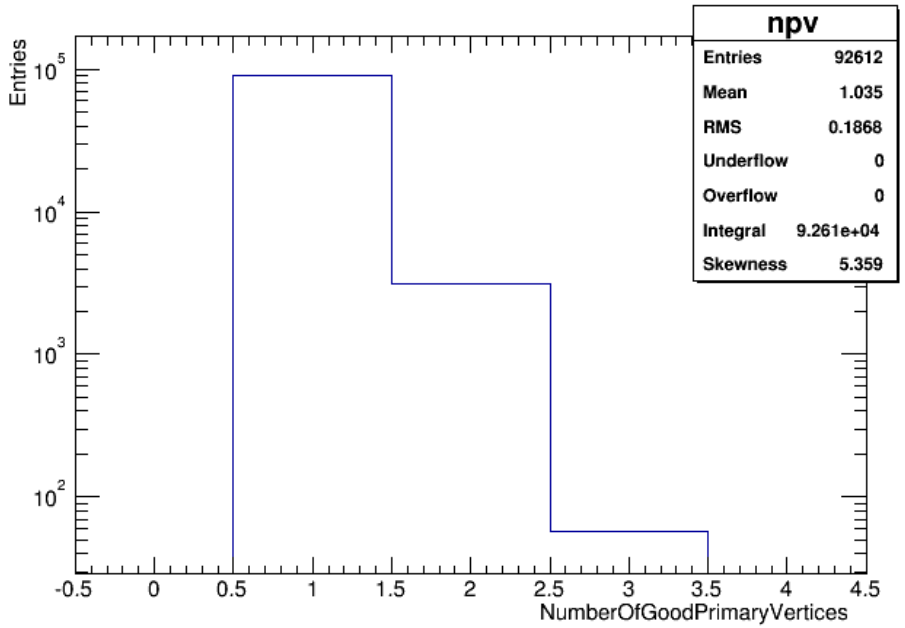
- More or less as expected
- More detailed results coming (HCAL DPG meeting)

Selecton

- Event selection:
  - No trigger requirement
  - No `OfficialDecision` requirement
  - `NumberOfGoodPrimaryVertices` > 0
- Channel selection:
  - Only `HBHE` considered
  - Rings: HB, HE: {17:20, 21:23, 24:25, 26:27, 28:28}
  - No channels in bad channels list
  - `RecHit` energy > 1 GeV
  - Charge > 5 fC
- Analyzer code:
  - Git page

N(vertex)

Number of primary vertices: QCD sample



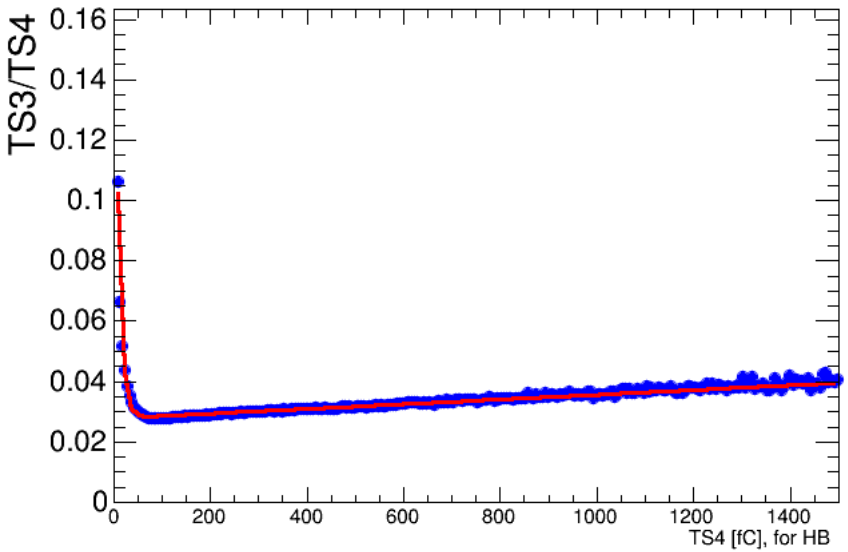
- 92612 events passing event selection
- Confirms no pileup, as expected

Definitions

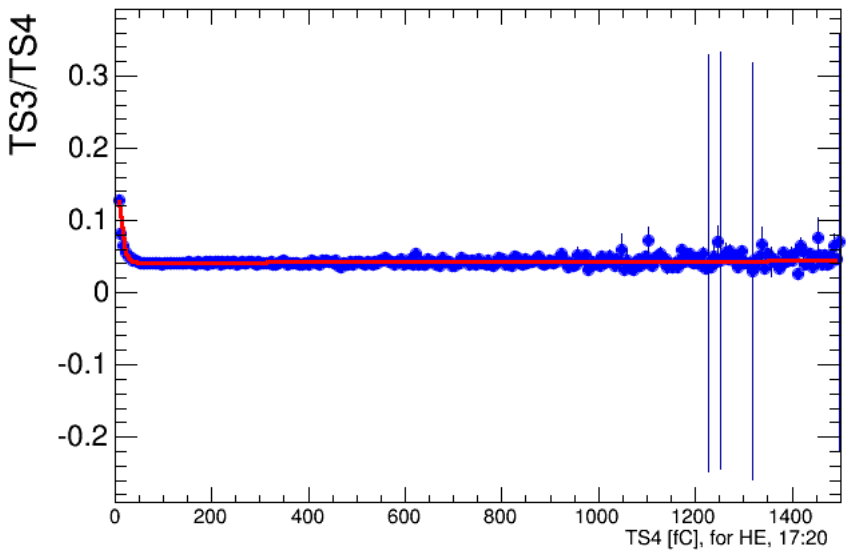
- The following plots show TProfile distributions
- One entry per HCAL digi in the ZS-collection
- x-axis corresponds to charge in TS4 [fC]
- y-axis corresponds to one of several charge ratios:
  - a_1: charge in TS3 [fC] / charge in TS4 [fC]
  - a1: charge in TS5 [fC] / charge in TS4 [fC]
  - a2: charge in TS6 [fC] / charge in TS4 [fC]
  - a3: charge in TS7 [fC] / charge in TS4 [fC]

a_1(TS4) in the QCD sample

a_1(TS4) in HB



a_1(TS4) in HE 17:20

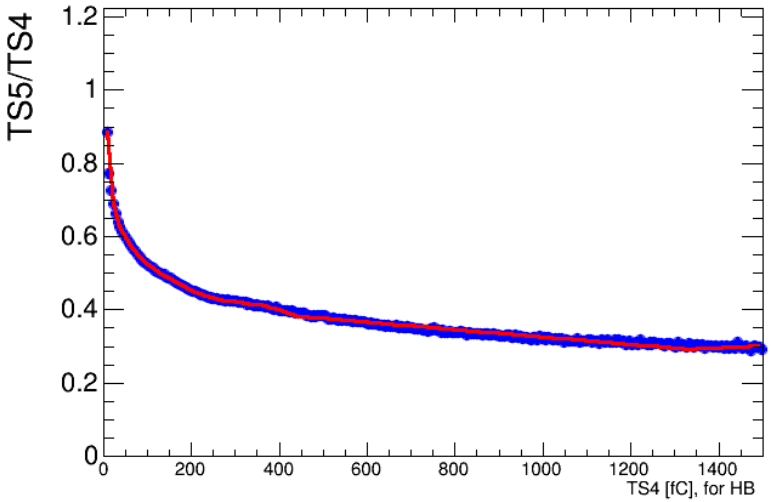


- Fit with exponential + polynomial:

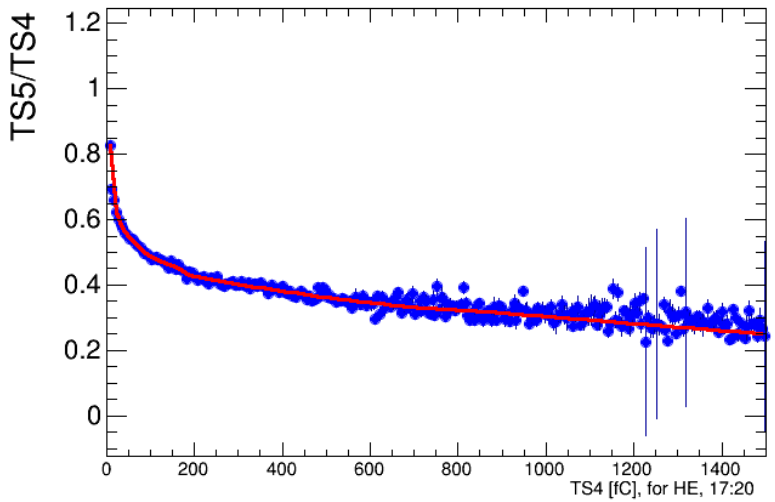
$$a\_1(TS4) = [0] + [1]\cdot TS4 + \mathrm{Exp}\left([2] + [3]\cdot TS4\right)$$

a1(TS4) in the QCD sample

a1(TS4) in HB



a1(TS4) in HE 17:20



- Fit with multiple polynomials (same shape as in data)
- Fit function describes the shape well
- Numeric results and data comparison next slide

a1(TS4) in the QCD sample: HB

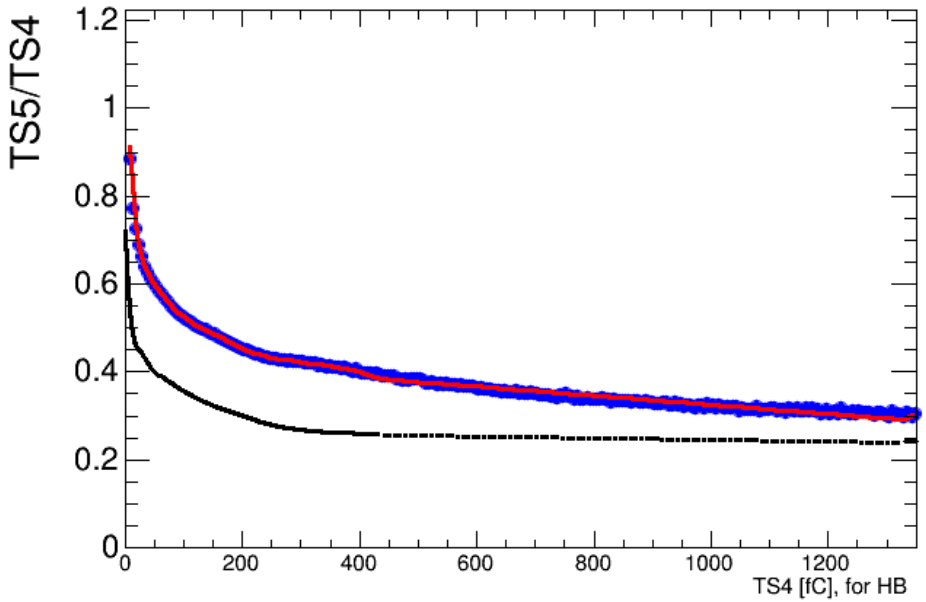
Variable	Region	MC value (unc. from fit)	Data value
a_0	TS4 $<$ 28 fC	1.4 ± 0.00028	0.73
a_1		$-0.12 \pm 2e-05$	-0.031
a_2		$0.0083 \pm 7.5e-07$	0.0012
a_3		$-0.00027 \pm 2.6e-08$	$-1.3e-05$
a_4		$3.4e-06 \pm 9.1e-10$	$-5.5e-08$
b_0	28 $\leq$ TS4 $<$ 60 fC	$0.84 \pm$ NA	0.42
b_1		$-0.0078 \pm 2e-05$	0.0069
b_2		$3.7e-05 \pm 2.3e-07$	-0.00033
b_3		$1.1e-06 \pm 3.3e-09$	$4.9e-06$
b_4		$-1.1e-08 \pm 5e-11$	$-2.4e-08$
c_0	60 $\leq$ TS4 $<$ 190 fC	$0.76 \pm$ NA	0.47
c_1		$-0.0042 \pm 5.2e-06$	-0.0015
c_2		$2.4e-05 \pm 2.2e-08$	$4.3e-06$
c_3		$-5.5e-08 \pm 1e-10$	$-4.7e-09$
d_0	190 $\leq$ TS4 $<$ 435 fC	$0.77 \pm$ NA	0.53
d_1		$-0.003 \pm 3.7e-06$	-0.0019
d_2		$9.1e-06 \pm 6.5e-09$	$4.6e-06$
d_3		$-9.7e-09 \pm 1.2e-11$	$-3.8e-09$
e_0	435 $\leq$ TS4 $<$ 1330 fC	$0.43 \pm$ NA	0.26
e_1		$-0.0001 \pm 5.1e-07$	$-1.9e-05$
f_0	1330 fC $\leq$ TS4	$0.43 \pm$ NA	0.24
f_1		$5.5e-05 \pm 5.9e-06$	$-3.6e-07$

a1(TS4) in the QCD sample: HE 17:20

Variable	Region	MC value (unc. from fit)	Data value
a_0	TS4 $<$ 23 fC	1.5 ± 0.00041	0.61
a_1		$-0.14 \pm 2e-05$	-0.0076
a_2		$0.0091 \pm 9.1e-07$	-0.00081
a_3		$-0.00024 \pm 4e-08$	$5.7e-05$
a_4		$1.7e-06 \pm 1.7e-09$	$-9.4e-07$
b_0	23 $\leq$ TS4 $<$ 65 fC	$0.71 \pm$ NA	0.4
b_1		$-0.0045 \pm 1.3e-05$	0.0068
b_2		$-1.6e-06 \pm 1.5e-07$	-0.00031
b_3		$7.7e-07 \pm 2.2e-09$	$4.5e-06$
b_4		$-5.2e-09 \pm 3.3e-11$	$-2.2e-08$
c_0	65 $\leq$ TS4 $<$ 190 fC	$0.76 \pm$ NA	0.46
c_1		$-0.0056 \pm 5.5e-06$	-0.0015
c_2		$3.8e-05 \pm 2.2e-08$	$4.9e-06$
c_3		$-9.4e-08 \pm 1.1e-10$	$-6.4e-09$
d_0	190 $\leq$ TS4 $<$ 850 fC	$0.47 \pm$ NA	0.48
d_1		$-0.00025 \pm 1.8e-06$	-0.0015
d_2		$2.8e-09 \pm 1.8e-09$	$3.3e-06$
d_3		$8.7e-11 \pm 2.1e-12$	$-2.5e-09$
e_0	850 $\leq$ TS4 $<$ 1640 fC	$0.41 \pm$ NA	0.26
e_1		$-0.00011 \pm 3.3e-06$	$-2.1e-05$
f_0	1640 fC $\leq$ TS4	$0.41 \pm$ NA	0.24
f_1		-830000.0 ± 2.0	0.0

a1(TS4) Data vs QCD MC

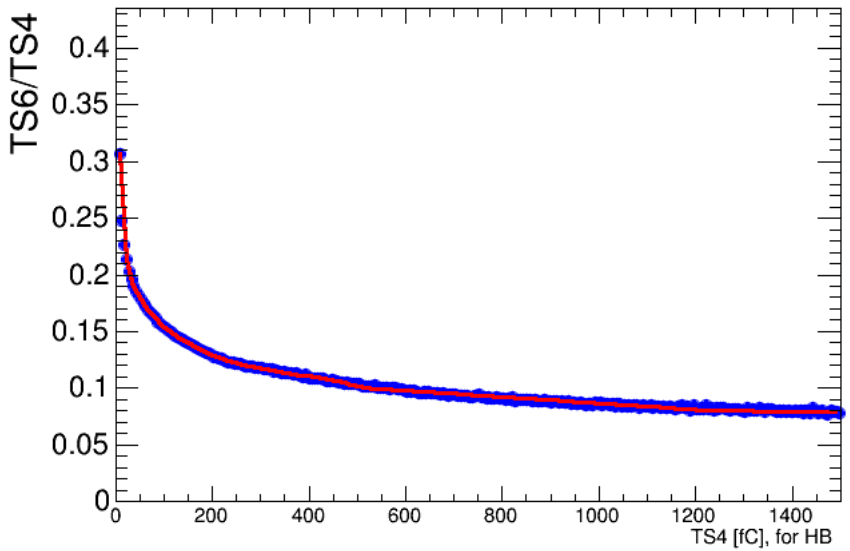
a1(TS4) Data vs Monte Carlo in HB



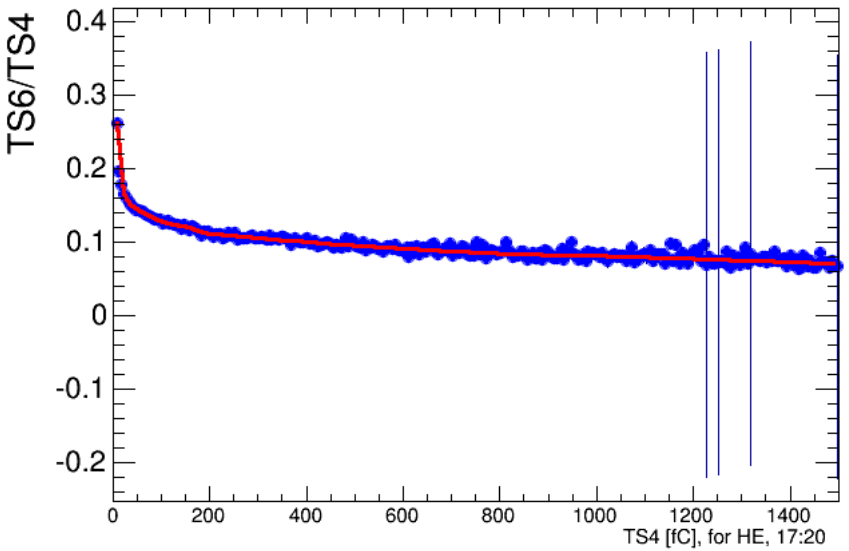
- Blue points: MC
- Red line: MC fit
- Black line: data fit (from Alexandre)

a2(TS4) in the QCD sample

a2(TS4) in HB



a2(TS4) in HE 17:20



a2(TS4) in the QCD sample: HB

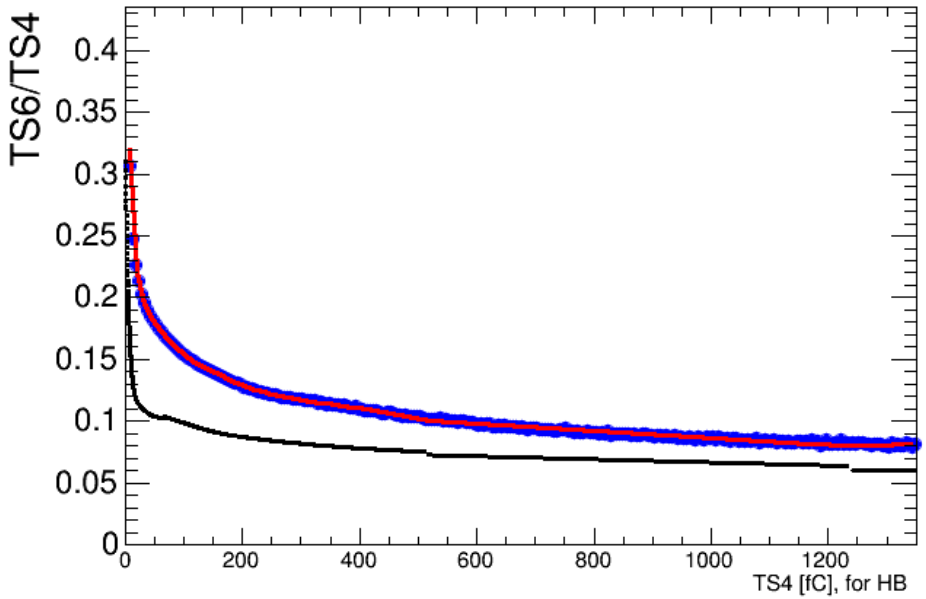
Variable	Region	MC value (unc. from fit)	Data value
a_0	TS4 $<$ 23 fC	0.48 ± 0.00011	0.31
a_1		$-0.03 \pm 7.6e-06$	-0.03
a_2		$0.00062 \pm 3.4e-07$	0.0017
a_3		$3.7e-05 \pm 1.4e-08$	$-4.5e-05$
a_4		$-1.3e-06 \pm 6e-10$	$4.4e-07$
b_0	23 $\leq$ TS4 $<$ 68 fC	$0.26 \pm$ NA	0.13
b_1		$-0.0028 \pm 4.1e-06$	-0.001
b_2		$1.9e-05 \pm 4.6e-08$	$1.2e-05$
b_3		$2.1e-07 \pm 6.2e-10$	$-6.7e-08$
b_4		$-2.7e-09 \pm 8.6e-12$	$3.8e-10$
c_0	68 $\leq$ TS4 $<$ 190 fC	$0.23 \pm$ NA	0.11
c_1		$-0.0013 \pm 1.6e-06$	$-5.3e-05$
c_2		$7.1e-06 \pm 6.6e-09$	$-1.1e-06$
c_3		$-1.5e-08 \pm 3e-11$	$3.7e-09$
d_0	190 $\leq$ TS4 $<$ 515 fC	$0.19 \pm$ NA	0.1
d_1		$-0.00049 \pm 7.9e-07$	-0.00011
d_2		$1.1e-06 \pm 1.2e-09$	$1.4e-07$
d_3		$-1e-09 \pm 2e-12$	$-7.2e-11$
e_0	515 $\leq$ TS4 $<$ 1240 fC	$0.11 \pm$ NA	0.079
e_1		$-2.8e-05 \pm 1.8e-07$	$-1.3e-05$
f_0	1240 fC $\leq$ TS4	$0.11 \pm$ NA	0.065
f_1		$-1.8e-06 \pm 8.5e-07$	$-3.8e-06$

a2(TS4) in the QCD sample: HE 17:20

Variable	Region	MC value (unc. from fit)	Data value
a_0	TS4 $<$ 23 fC	0.47 ± 0.00052	0.31
a_1		$-0.035 \pm 4.5e-05$	-0.027
a_2		$0.00077 \pm 2e-06$	0.0014
a_3		$4.4e-05 \pm 8.5e-08$	$-3.2e-05$
a_4		$-1.6e-06 \pm 3.5e-09$	$2.7e-07$
b_0	23 $\leq$ TS4 $<$ 68 fC	$0.19 \pm$ NA	0.15
b_1		$-0.0013 \pm 2.5e-05$	-0.0033
b_2		$3.4e-06 \pm 3e-07$	$7.8e-05$
b_3		$1.6e-07 \pm 4e-09$	$-9.3e-07$
b_4		$-1.3e-09 \pm 5.4e-11$	$4.5e-09$
c_0	68 $\leq$ TS4 $<$ 190 fC	$0.19 \pm$ NA	0.11
c_1		$-0.0013 \pm 1.1e-05$	-0.00025
c_2		$8.3e-06 \pm 4.4e-08$	$2.9e-07$
c_3		$-2e-08 \pm 2e-10$	$7.7e-10$
d_0	190 $\leq$ TS4 $<$ 1000 fC	$0.13 \pm$ NA	0.091
d_1		$-7.1e-05 \pm 2.3e-06$	-0.0001
d_2		$2e-08 \pm 2.6e-09$	$1.3e-07$
d_3		$8e-12 \pm 2.1e-12$	$-5.7e-11$
e_0	1000 $\leq$ TS4 $<$ 1380 fC	$0.1 \pm$ NA	0.065
e_1		$-2.1e-05 \pm 2.8e-06$	$-8.5e-06$
f_0	1380 fC $\leq$ TS4	$0.1 \pm$ NA	0.053
f_1		$-2.7e-05 \pm 1.3e-05$	0.0

a2(TS4) Data vs QCD MC

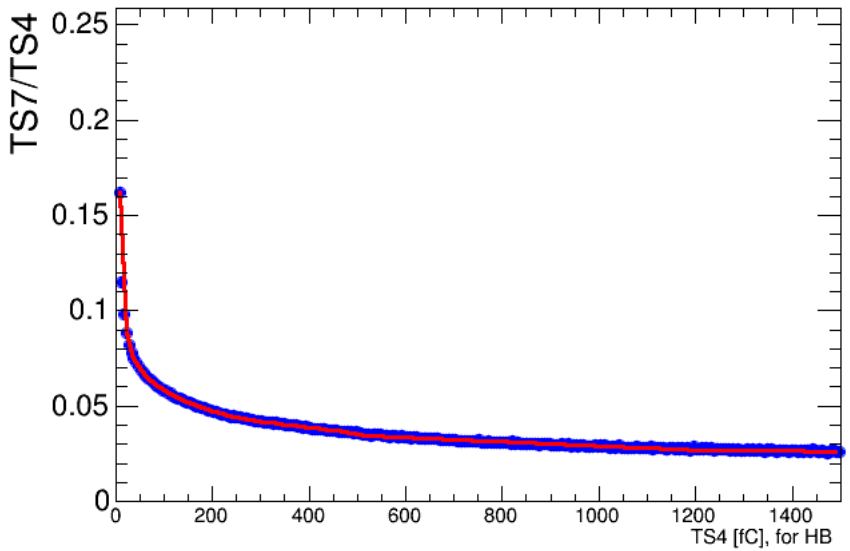
a2(TS4) Data vs Monte Carlo in HB



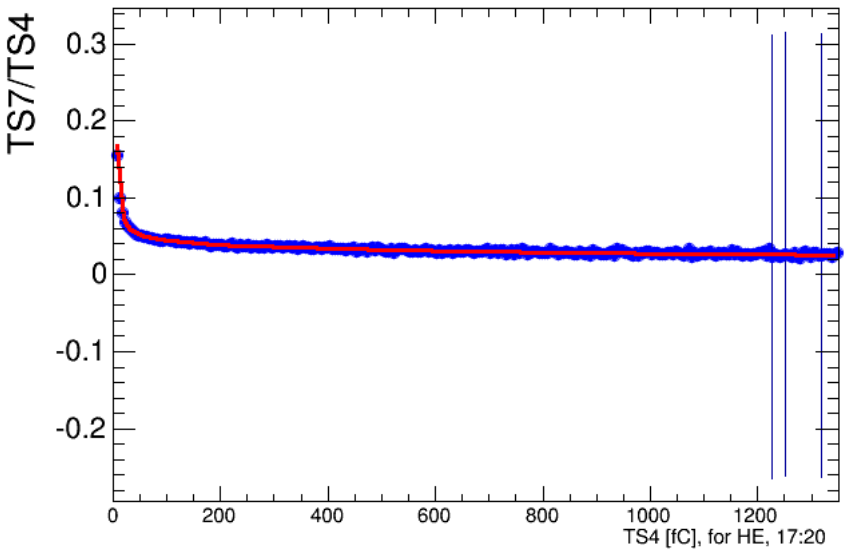
- Blue points: MC
- Red line: MC fit
- Black line: data fit (from Alexandre)

a3(TS4) in the QCD sample

a3(TS4) in HB



a3(TS4) in HE 17:20



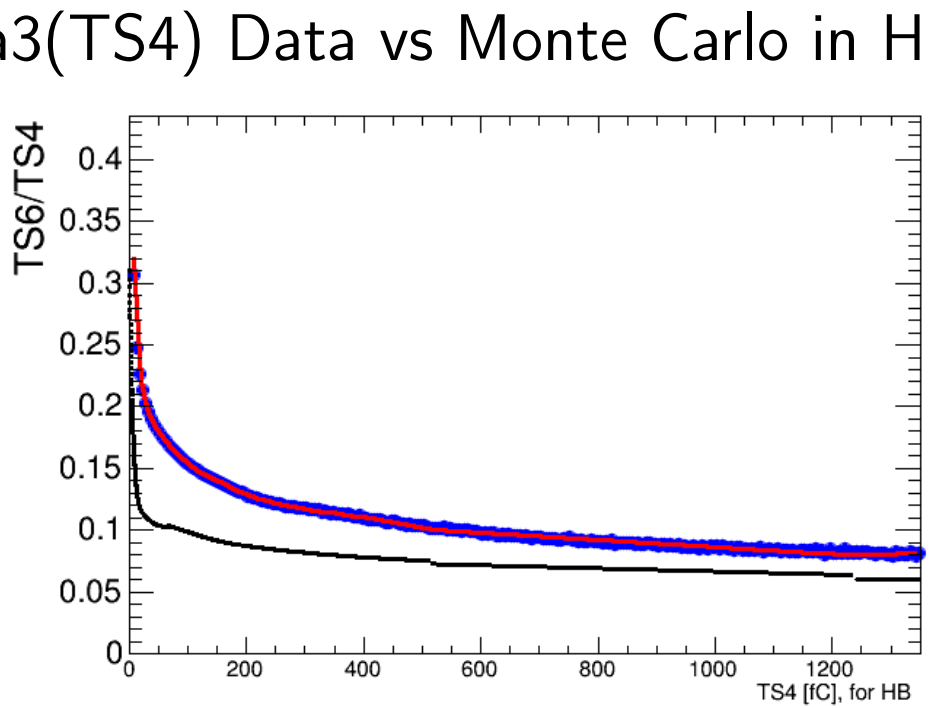
a3(TS4) in the QCD sample: HB

Variable	Region	MC value (unc. from fit)	Data value
a_0	TS4 $<$ 23 fC	0.48 ± 0.00011	0.31
a_1		$-0.03 \pm 7.6e-06$	-0.03
a_2		$0.00062 \pm 3.4e-07$	0.0017
a_3		$3.7e-05 \pm 1.4e-08$	$-4.5e-05$
a_4		$-1.3e-06 \pm 6e-10$	$4.4e-07$
b_0	$23 \leq$ TS4 $<$ 68 fC	$0.26 \pm$ NA	0.13
b_1		$-0.0028 \pm 4.1e-06$	-0.001
b_2		$1.9e-05 \pm 4.6e-08$	$1.2e-05$
b_3		$2.1e-07 \pm 6.2e-10$	$-6.7e-08$
b_4		$-2.7e-09 \pm 8.6e-12$	$3.8e-10$
c_0	$68 \leq$ TS4 $<$ 190 fC	$0.23 \pm$ NA	0.11
c_1		$-0.0013 \pm 1.6e-06$	$-5.3e-05$
c_2		$7.1e-06 \pm 6.6e-09$	$-1.1e-06$
c_3		$-1.5e-08 \pm 3e-11$	$3.7e-09$
d_0	$190 \leq$ TS4 $<$ 515 fC	$0.19 \pm$ NA	0.1
d_1		$-0.00049 \pm 7.9e-07$	-0.00011
d_2		$1.1e-06 \pm 1.2e-09$	$1.4e-07$
d_3		$-1e-09 \pm 2e-12$	$-7.2e-11$
e_0	$515 \leq$ TS4 $<$ 1240 fC	$0.11 \pm$ NA	0.079
e_1		$-2.8e-05 \pm 1.8e-07$	$-1.3e-05$
f_0	1240 fC $\leq$ TS4	$0.11 \pm$ NA	0.065
f_1		$-1.8e-06 \pm 8.5e-07$	$-3.8e-06$

a3(TS4) in the QCD sample: HE 17:20

Variable	Region	MC value (unc. from fit)	Data value
a_0	TS4 $<$ 23 fC	0.47 ± 0.00052	0.31
a_1		$-0.035 \pm 4.5e-05$	-0.027
a_2		$0.00077 \pm 2e-06$	0.0014
a_3		$4.4e-05 \pm 8.5e-08$	$-3.2e-05$
a_4		$-1.6e-06 \pm 3.5e-09$	$2.7e-07$
b_0	23 $\leq$ TS4 $<$ 68 fC	$0.19 \pm$ NA	0.15
b_1		$-0.0013 \pm 2.5e-05$	-0.0033
b_2		$3.4e-06 \pm 3e-07$	$7.8e-05$
b_3		$1.6e-07 \pm 4e-09$	$-9.3e-07$
b_4		$-1.3e-09 \pm 5.4e-11$	$4.5e-09$
c_0	68 $\leq$ TS4 $<$ 190 fC	$0.19 \pm$ NA	0.11
c_1		$-0.0013 \pm 1.1e-05$	-0.00025
c_2		$8.3e-06 \pm 4.4e-08$	$2.9e-07$
c_3		$-2e-08 \pm 2e-10$	$7.7e-10$
d_0	190 $\leq$ TS4 $<$ 1000 fC	$0.13 \pm$ NA	0.091
d_1		$-7.1e-05 \pm 2.3e-06$	-0.0001
d_2		$2e-08 \pm 2.6e-09$	$1.3e-07$
d_3		$8e-12 \pm 2.1e-12$	$-5.7e-11$
e_0	1000 $\leq$ TS4 $<$ 1380 fC	$0.1 \pm$ NA	0.065
e_1		$-2.1e-05 \pm 2.8e-06$	$-8.5e-06$
f_0	1380 fC $\leq$ TS4	$0.1 \pm$ NA	0.053
f_1		$-2.7e-05 \pm 1.3e-05$	0.0

a3(TS4) Data vs QCD MC



- Blue points: MC
- Red line: MC fit
- Black line: data fit (from Alexandre)

Conclusion

- Processed zero-pileup samples adequate for studies
- Preliminary results ready using Alexandre's method
  - Fit functions used for data model MC pulse shape well
  - Final fit parameters (i.e. pulse shapes) are significantly different between data and MC
- Working on validating results to put into CMSSW in time for 710