Bremsstrahlung showers

take a brake, take a shower

To discuss briefly:

1. Shower likelihood estimation
2. Hypothesis testing
2. Shower fitting for true track and mc-hits
3. Shower fitting for reco track and hits

Timon van Dieren

Reconstruction meeting May 7, 2020

Some remarks

- Events to be discussed are from ORCA 115 lines atmospheric muons with energies ~100 GeV – 1 TeV

```
input_file = "/data/km3net/mc/ORCA115_2019_muon/KM3NeT_ORCA_115_20m_9m/v5.0/reco/
mcv5.0.mupage_1G.sirene.jte.jchain.aashower.99821.root"
```

- Problem of this mc simulation file is energy reconstruction : limited to 1-100 GeV. For reco track I correct for this by applying a distance correction to the true energy.

```
dist_energy_correction =  (track_reco.pos – track_mc.pos).len()
… # check whether the distance is positive or negative
energy_func = JPHYSICS.JGeane_t(a, b)
…
 track_reco.E = energy_func.getE(track_mc.E, dist_energy_correction)
```

- For mc-hits: don’t forget scaling factor

Shower likelihood estimation (1)

PDFs from

```
DIRECT_LIGHT_FROM_MUON            =  1,    //!< direct    light from muon
SCATTERED_LIGHT_FROM_MUON         =  2,    //!< scattered light from muon

DIRECT_LIGHT_FROM_EMSHOWERS       =  3,    //!< direct    light from EM showers
SCATTERED_LIGHT_FROM_EMSHOWERS    =  4,    //!< scattered light from EM showers

DIRECT_LIGHT_FROM_DELTARAYS       =  5,    //!< direct    light from delta-rays
SCATTERED_LIGHT_FROM_DELTARAYS    =  6,    //!< scattered light from delta-rays

SCATTERED_LIGHT_FROM_MUON_5D      = 12,    //!< scattered light from muon

DIRECT_LIGHT_FROM_EMSHOWER        = 13,    //!< direct    light from EM shower
SCATTERED_LIGHT_FROM_EMSHOWER     = 14,    //!< scattered light from EM shower
```

https://common.pages.km3net.de/jpp/JPDFTypes_8hh_source.html

mc-hits

$$\frac{dP}{dt}(t) = \sum_{i=1,2,5,6,13,14} \frac{dP_i}{dt}(t)$$

For 1st hits

$$g(t) = \frac{dP}{dt}(t)\frac{e^{-n(t)}}{1-e^{-N}}$$

$$n(t) = \int_{T_{min}}^{t} \frac{dP}{dt'}dt'$$

$$N = \int_{T_{min}}^{T_{max}} \frac{dP}{dt'}dt'$$

Shower likelihood estimation (2)

Wanted: shower distance, shower energy

Method: maximizing likelihood. We need both hit and no-hit PMTs, by looking at hit PMTs only the “best-fit” would have infinite energy.

MC-hits: hits are discrete, stochastic process. L = multiplication of all Poisson probabilities to have k_i hits at PMT i with expected number of hits N_i . This likelihood neglects the timing information. This could be implemented by multiplying by dP/dt(t_i) for all the hits.

Hits: we want estimator both incorporating hit/no-hit information and timing (THANK YOU AART!)

For speeding up computations: only look at hits and pmts with distance of closest approach < d_threshold (I use R<80 m)

---

mc-hits

$$\ln L = \sum_{PMTs} \ln \left(\frac{N_i^{k_i} e^{-N_i}}{k_i!} \right)$$

Poisson

For 1st hits

$$\ln L_{1st} = \sum_{hit\ PMTs} \log \left(1 - e^{-N_i}\right) + \log \left(\frac{dP}{dt}(t_i) \frac{e^{-n_i(t_i)}}{1 - e^{-N_i}} \right) + \sum_{no-hit\ PMTs} -N_i$$

Probability at least 1 hit

Probability that *given a hit occurred* it took place at time t_i

Probability no hit

Intermezzo: hypothesis tests (1)

Why would there be an intermezzo without music!? https://open.spotify.com/album/4no6O3D7VpSoyf6mzM1FZJ?si=Cln0MW9RRtasunlMfc0ifQ

Two tests:

1. Null hypothesis: no shower.
Calculate *1 – Poisson_cdf(hits, <npe>)* where <npe> is for muon/delta only

2. Likelihood ratio test
Calculate 1 – chi2_cdf(x, dof=2) where
x = -2 * (ln_L_no_shower - ln_L_shower_fit)

Intermezzo: hypothesis tests (2)

Why would there be an intermezzo without music!? https://open.spotify.com/album/4no6O3D7VpSoyf6mzM1FZJ?si=Cln0MW9RRtasunlMfc0ifQ

Results for true track, mc-hits (by doing a scan along distances for a shower energy of 20 GeV and then take the maximum ln(L)

evt_nr	E_muon [GeV]	<npe> no_shwr	#MC-hits	chi2 test	ln_L_no_shwr	ln_L_fit	ratio test	3 sigma	5 sigma
1	354.658	450	624	0.0	-3667.7	-3591.7	0.0	yes	yes
2	188.249	370	394	0.105359538	-2260.9	-2260.7	0.88241364	no	no
3	184.685	125	138	0.116519144	-789.4	-789.5	1.0	no	no
4	530.014	417	492	0.000143898	-2618.7	-2604.0	3.94e-07	yes	no
5	142.317	436	463	0.098435012	-2047.9	-2047.9	1.0	no	no
6	226.68	380	438	0.001756902	-2465.8	-2466.1	1.0	no	no
7	655.869	70	110	2.862e-06	-676.1	-672.2	0.020421285	yes	no
8	742.269	188	262	1.48e-07	-1234.2	-1230.7	0.03065644	yes	yes
9	176.474	543	544	0.468565595	-3114.2	-3115.4	1.0	no	no
10	152.11	395	458	0.000833668	-2546.2	-2544.6	0.20933041	yes	no
11	205.757	560	586	0.129313551	-3075.9	-3076.1	1.0	no	no
12	98.771	318	238	0.999998467	-1795.2	-1795.4	1.0	no	no
13	726.116	592	1350	0.0	-6438.6	-5855.7	0.0	yes	yes
14	111.576	379	318	0.999307813	-2266.5	-2266.6	1.0	no	no
15	431.83	34	54	0.000442053	-463.7	-458.2	0.004094578	yes	no
16	515.087	505	619	4.03e-07	-2845.0	-2845.4	1.0	yes	no
17	141.042	234	256	0.068159675	-1497.7	-1497.8	1.0	no	no
18	357.436	323	391	0.000121814	-2316.6	-2315.0	0.210766712	yes	no
19	688.5	63	179	0.0	-945.8	-840.9	0.0	yes	yes
20	9849.701	11	210	0.0	-852.0	-787.3	0.0	yes	yes
21	107.358	166	76	1.0	-791.1	-789.3	0.156174725	no	no

Note: for ln_L_fit we use a shower of energy 20 GeV at a range of distances along the track, and then pick the distance corresponding to
maximum ln(L) to be the 'fit'.

Intermezzo: hypothesis tests (3)

Why would there be an intermezzo without music!? https://open.spotify.com/album/4no6O3D7VpSoyf6mzM1FZJ?si=Cln0MW9RRtasunlMfc0ifQ

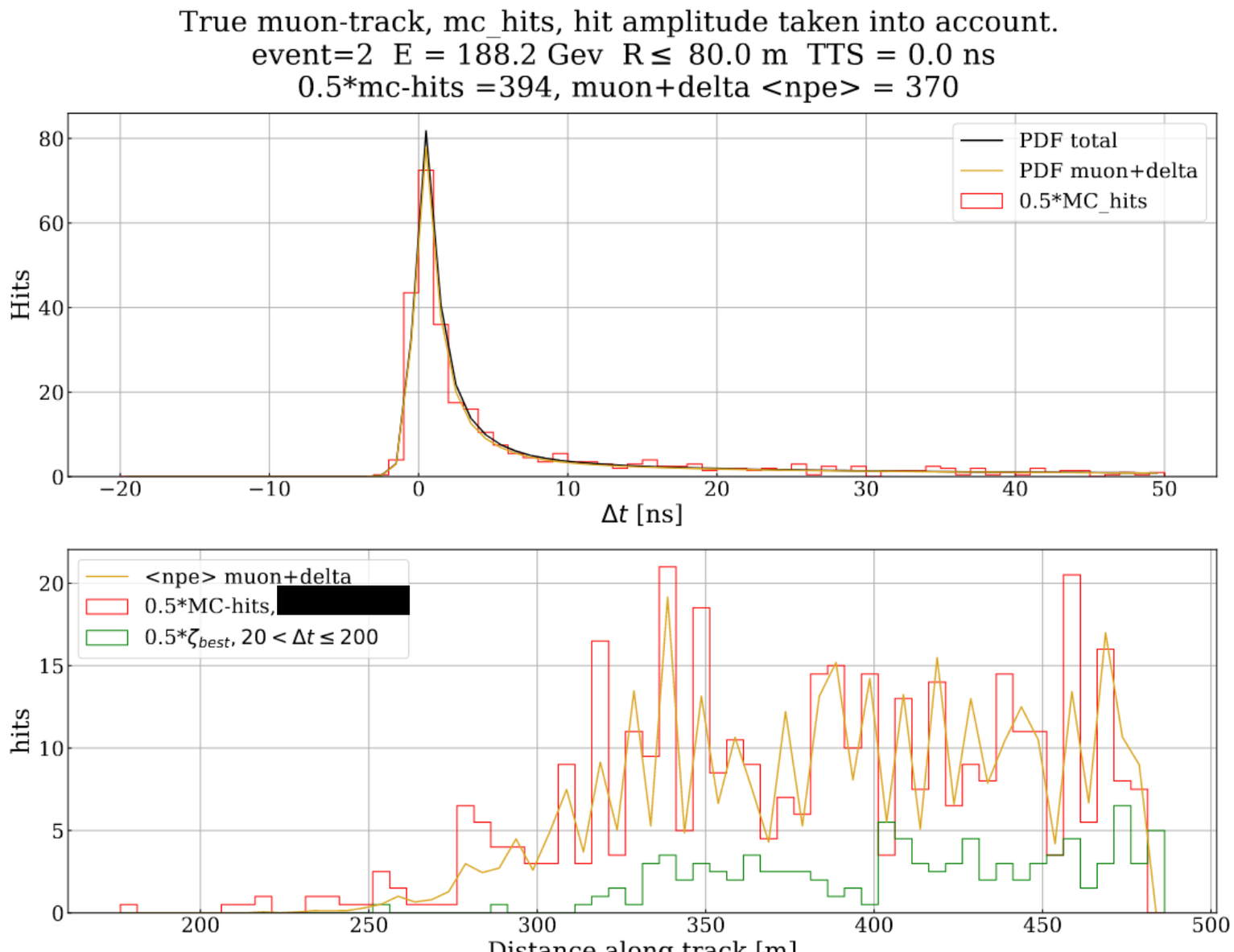
Results for reco track, hits (by doing a scan along distances for a shower energy of 20 GeV and then take the maximum ln(L)

evt_nr	E_mc [GeV]	E_reco [GeV]	<npe>	no_shwr	#hits	chi2 test	ln_L_no_shwr	ln_L_fit	ratio test	3 sigma	5 sigma
1	354.658	259.090513428	586		685	3.1237e-05	-6094.7	-6085.0	1.1391e-05	yes	no
2	188.249	96.5507206274	439		391	0.989903148	-3467.1	-3467.5	1.0	no	no
3	184.685	33.1746458531	171		170	0.514017267	-1525.9	-1526.0	1.0	no	no
4	530.014	421.699827646	477		612	1e-09	-5345.2	-5321.9	0.0	yes	yes
5	142.317	57.7971633069	421		417	0.562520522	-3388.4	-3389.1	1.0	no	no
6	226.68	120.497244064	452		461	0.318106027	-4054.4	-4054.9	1.0	no	no
7	655.869	516.887276337	106		103	0.593550195	-1035.1	-1034.7	0.37133274	no	no
8	742.269	540.665685328	220		237	0.119429936	-2080.4	-2079.9	0.30454107	no	no
9	176.474	81.8841246849	649		584	0.994766967	-5255.8	-5255.6	0.552017114	no	no
10	152.11	74.5225035265	499		774	0.0	-6836.8	-6823.1	1.66e-07	yes	yes
11	205.757	98.6592748540	623		592	0.891241117	-5092.3	-5092.4	1.0	no	no
12	98.771	21.1350707530	422		324	0.999999642	-3277.5	-3277.4	0.615563616	no	no
13	726.116	557.118249485	616		1446	0.0	-12279.3	-12092.9	0.0	yes	yes
14	111.576	33.2644162273	466		412	0.994060193	-3917.6	-3918.0	1.0	no	no
15	431.83	299.157929521	99		160	8e-09	-1720.5	-1717.5	0.015024423	yes	yes
16	515.087	369.525080258	547		568	0.179259973	-4780.7	-4782.8	1.0	no	no
17	141.042	30.2246679666	294		288	0.616759142	-2659.9	-2660.4	1.0	no	no
18	357.436	226.518371276	394		426	0.054889041	-3857.1	-3856.6	0.320949108	no	no
19	688.5	476.621140965	107		149	4.562e-05	-1376.7	-1346.0	0.0	yes	no
20	9849.701	8932.26209214	50		204	0.0	-2019.5	-1972.0	0.0	yes	yes
21	107.358	20.2880842994	236		152	0.999999997	-1778.0	-1777.3	0.227173642	no	no

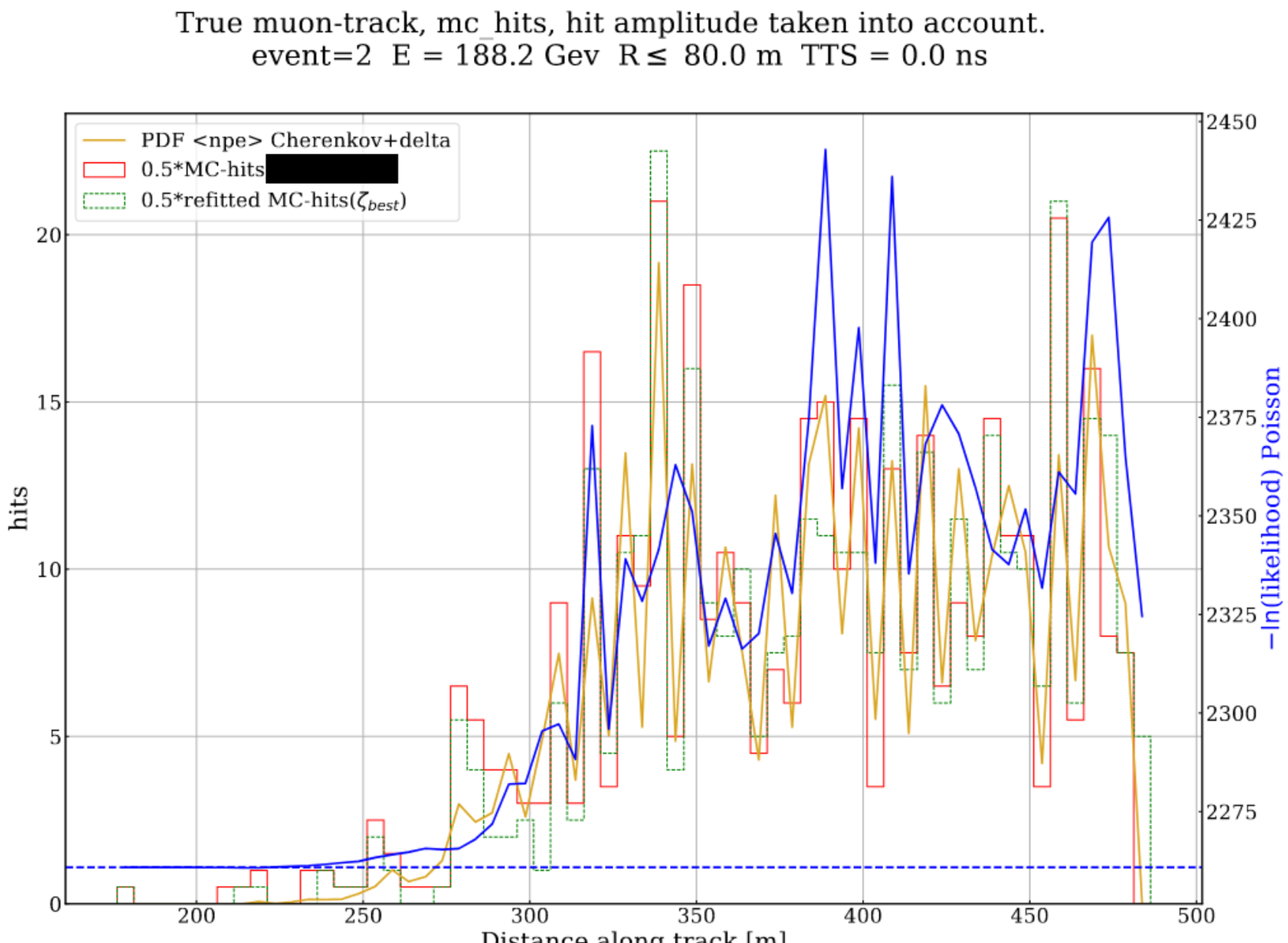
round_fit
p3_hits
-05-07 03:27:04
1-05-07 03:27:04

Note: for ln_L_fit we use a shower of energy 20 GeV at a range of distances along the track, and then pick the distance corresponding to
maximum ln(L) to be the 'fit'.
#

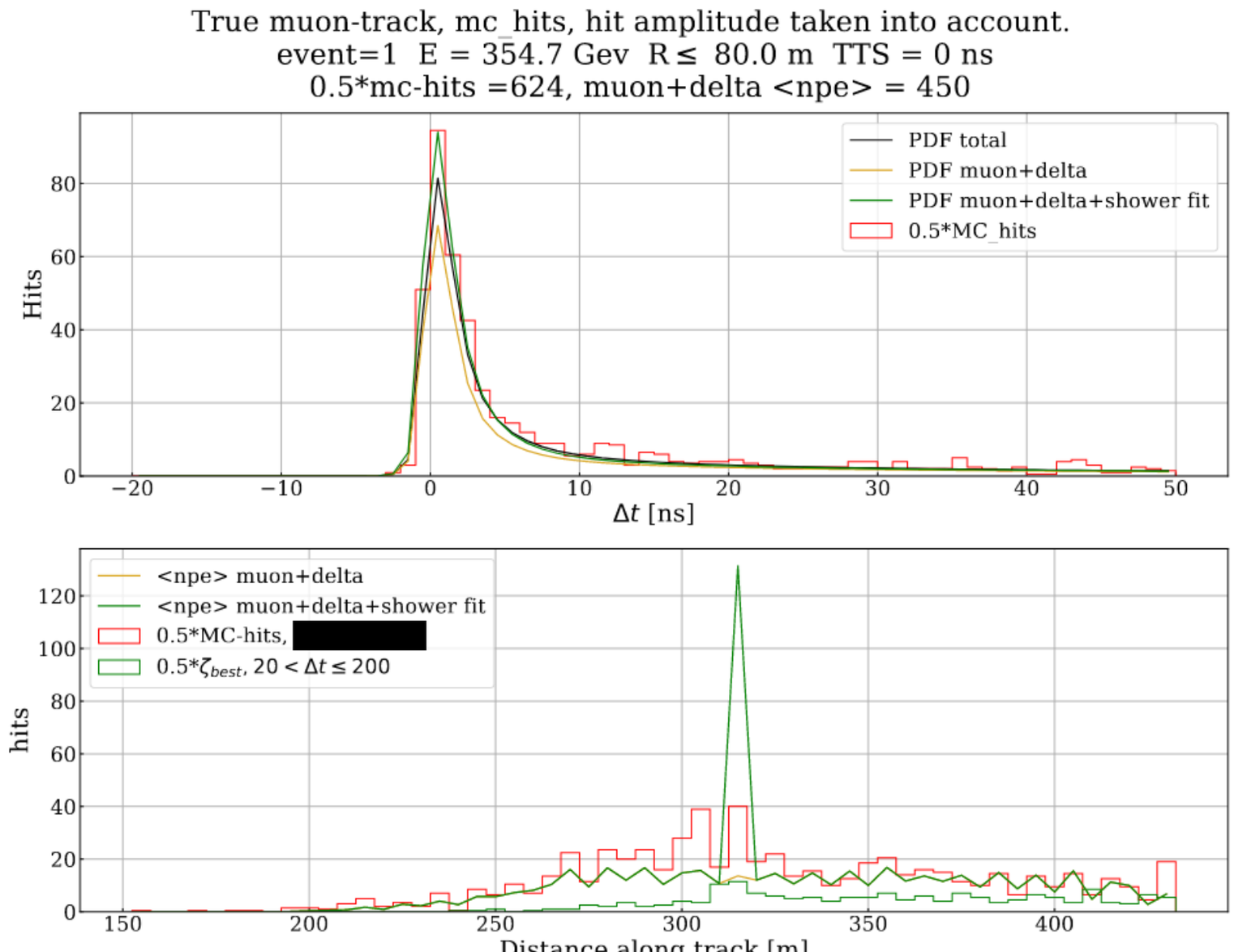
MC-hits: no shower example (1)



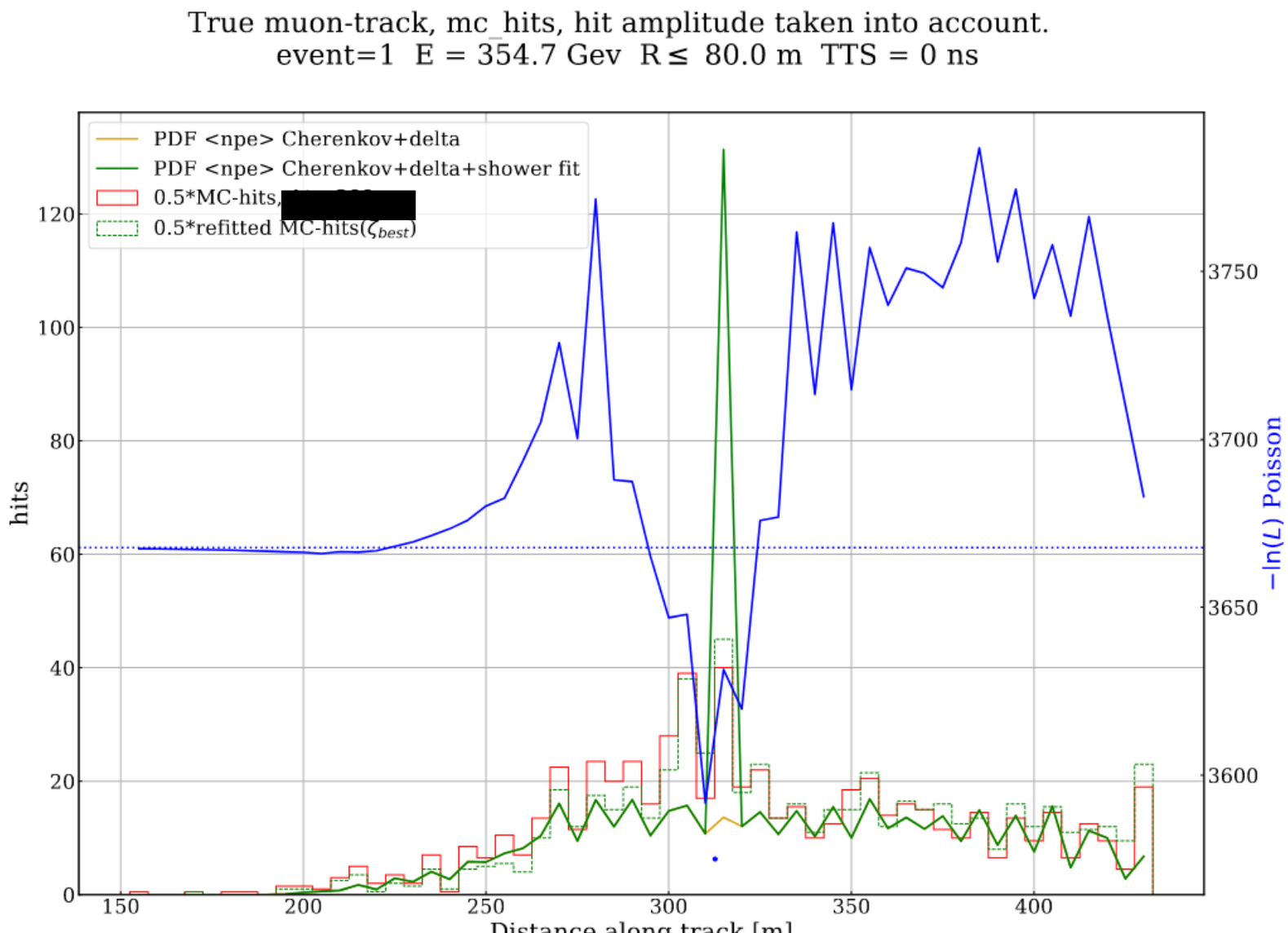
MC-hits: no shower example (2)



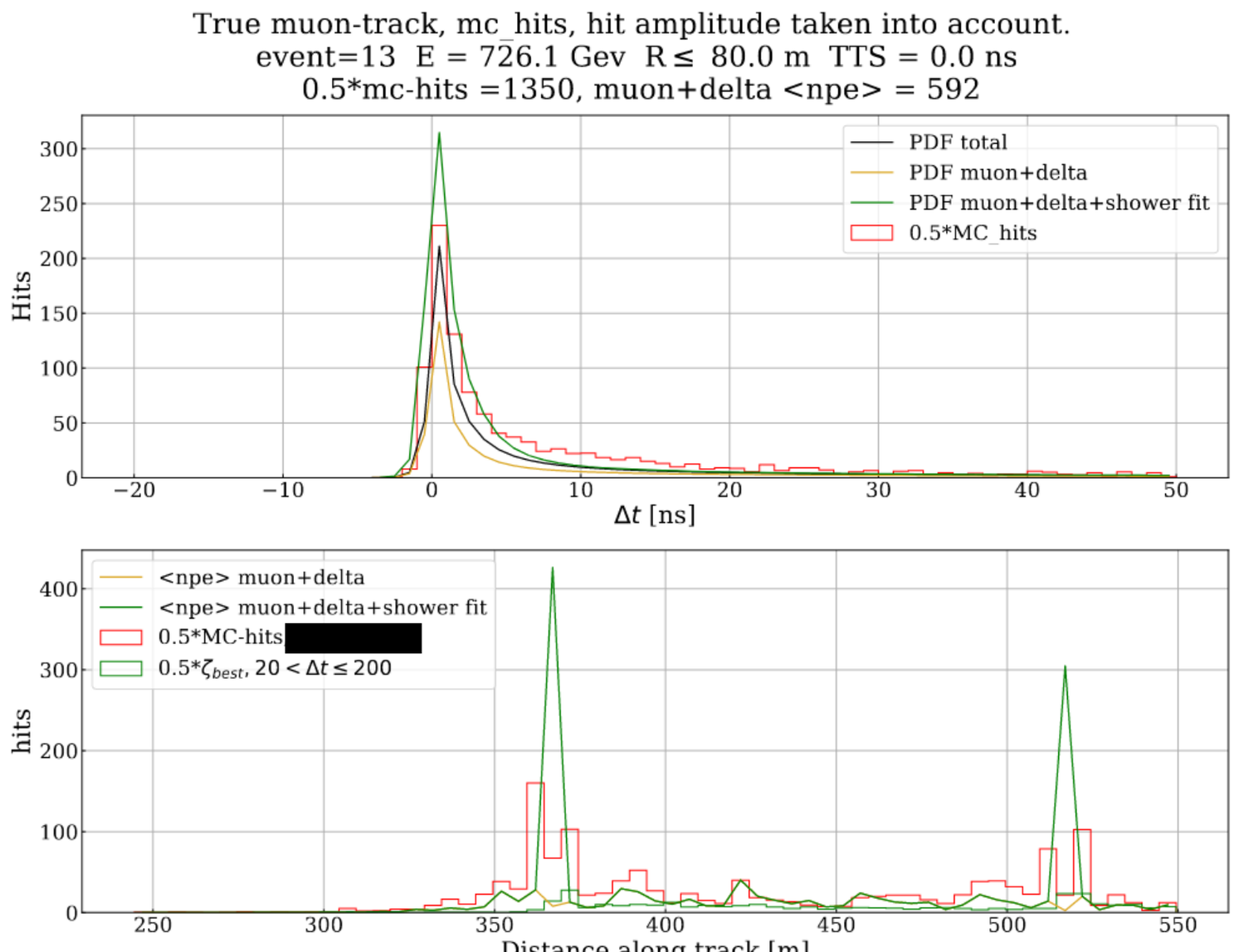
MC-hits: shower example (1)



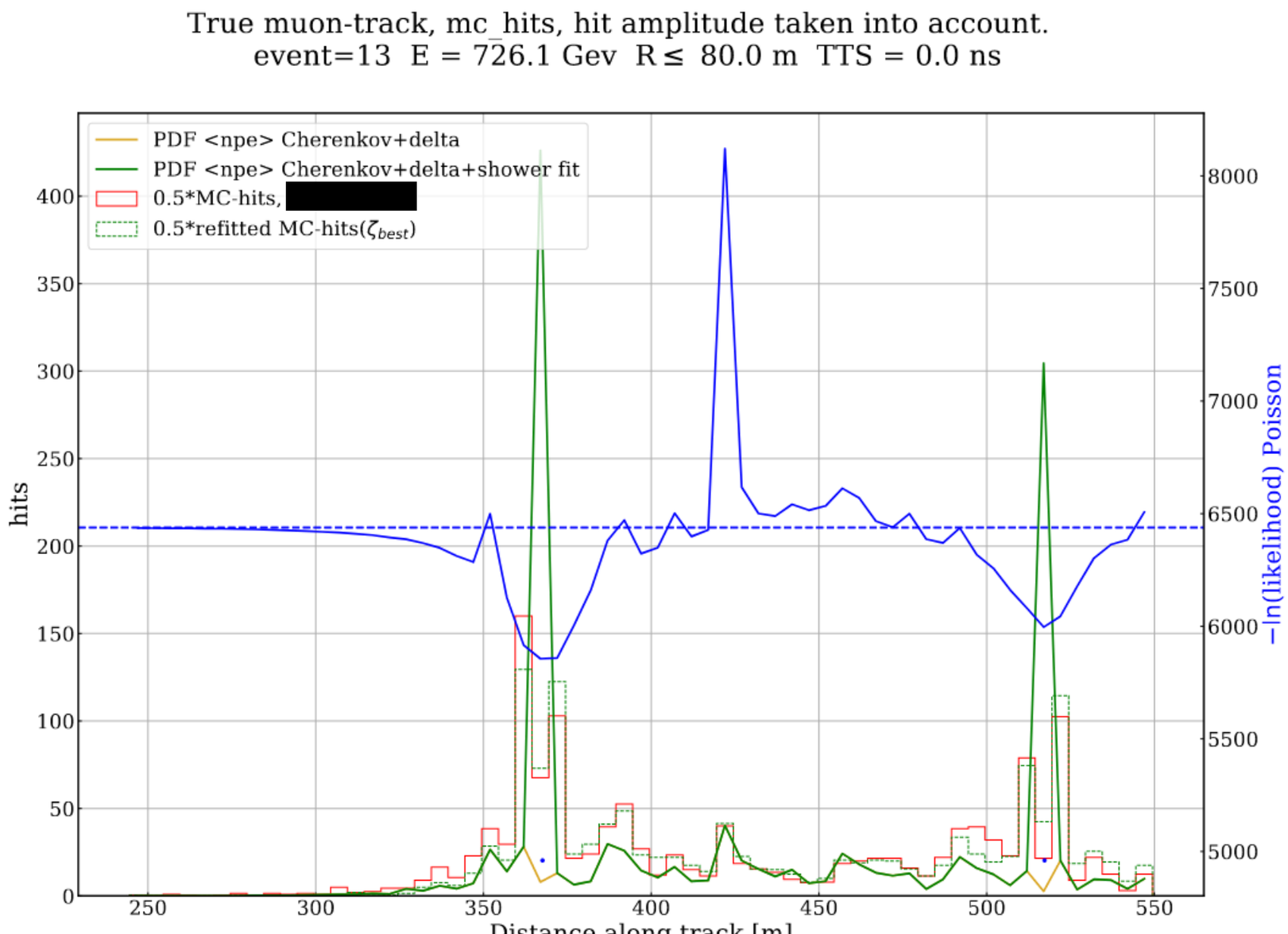
MC-hits: shower example (2)



MC-hits: shower example (3)



MC-hits: shower example (4)



MC-hits: shower fits

evt_nr	E_muon [GeV]	5 sigma H0	<npe> MIP	MC-hits	<npe> fit	ln_L_no_shwr	ln_L_fit	E_shwr [GeV]	d_shwr [m]	<npe> shwrs
1	354.658	yes	450	624	568	-3667.7	-3575.0	[16.3]	[312.7]	[118]
4	530.014	no	417	492	448	-2618.7	-2600.8	[12.3]	[265.1]	[31]
10	152.11	no	395	458	422	-2546.2	-2534.8	[7.9]	[260.7]	[27]
13	726.116	yes	592	1350	1312	-6438.6	-4960.4	[77.6, 55.7]	[367.6, 517.3]	[418, 302]

evt_nr	E_muon [GeV]	5 sigma H0	<npe> MIP	MC-hits	<npe> fit	ln_L_no_shwr	ln_L_fit	E_shwr [GeV]	d_shwr [m]	<npe> shwrs
8	742.269	yes	188	262	190	-1234.2	-1234.1	[0.7]	[425.3]	[1]
19	688.5	yes	63	179	137	-945.8	-824.9	[40.3]	[495.7]	[74]

I Fitted some events which have 3 or 5 sigma rejection of null hypothesis: no showers

Energies are within GeV range. Visual check of fitted distances looks nice.

Even found event with two showers (for muon with E_mc = 726.1 GeV)

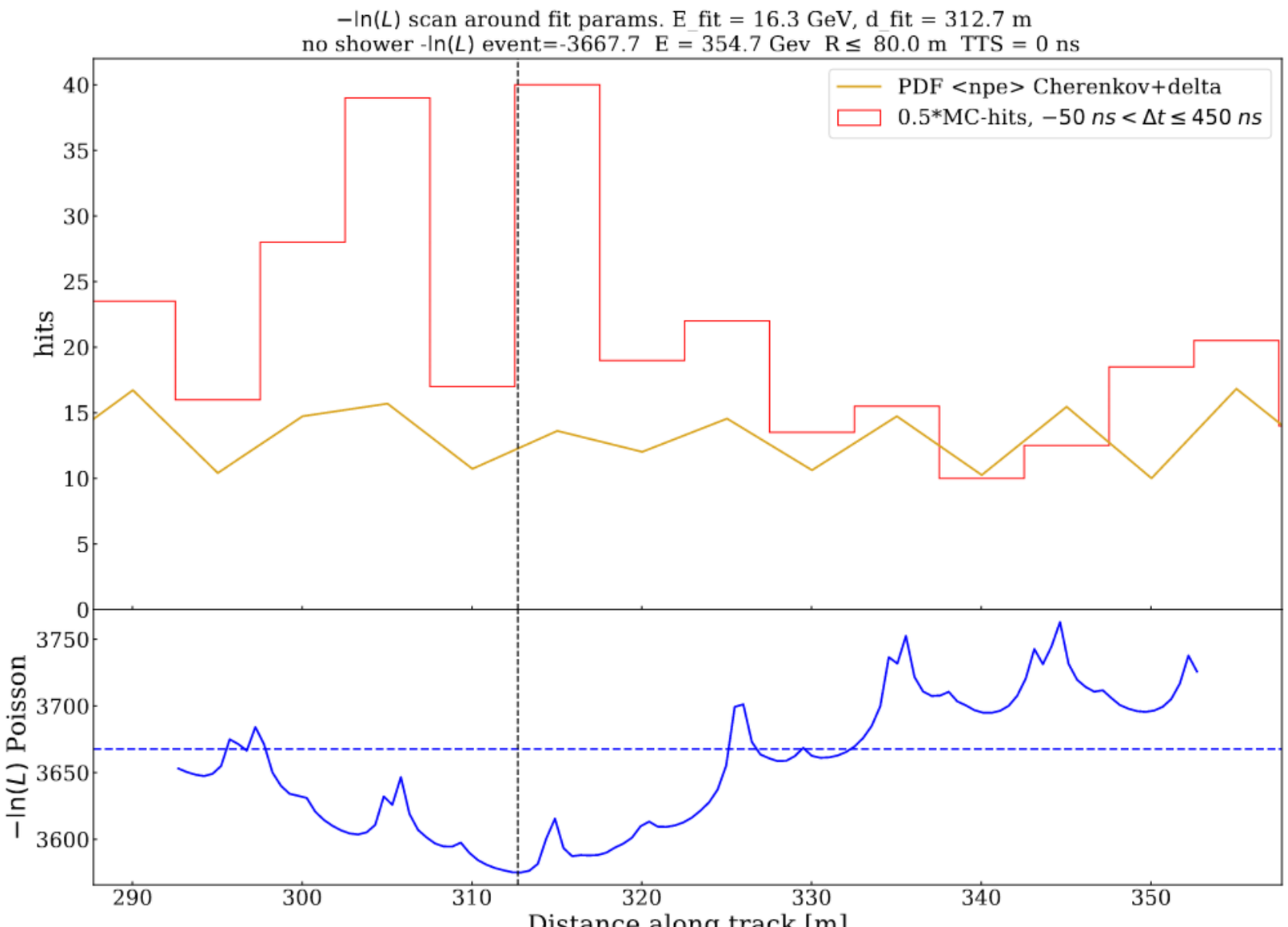
Note: fit for evt_nr = 8 is not statistically significant based on the likelihood ratio test.

MC-hits: ln(L) scan around fit (1)

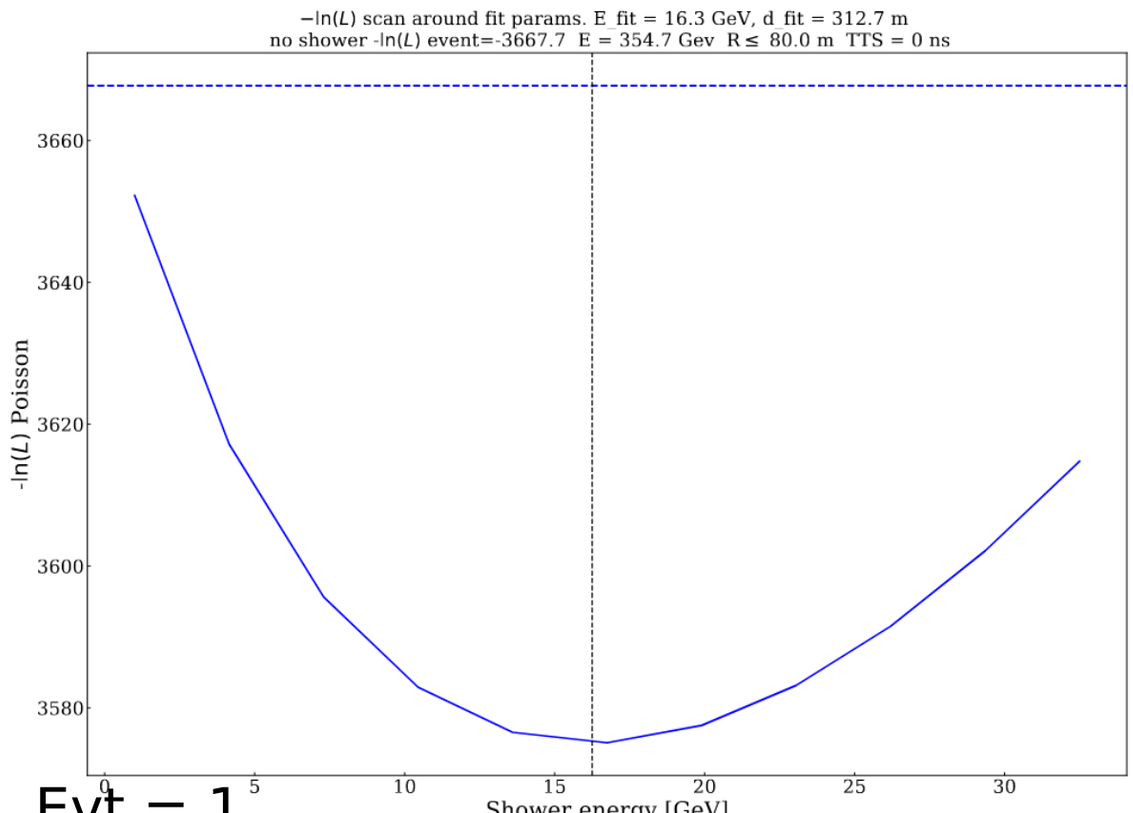
Evt = 1
Fix energy

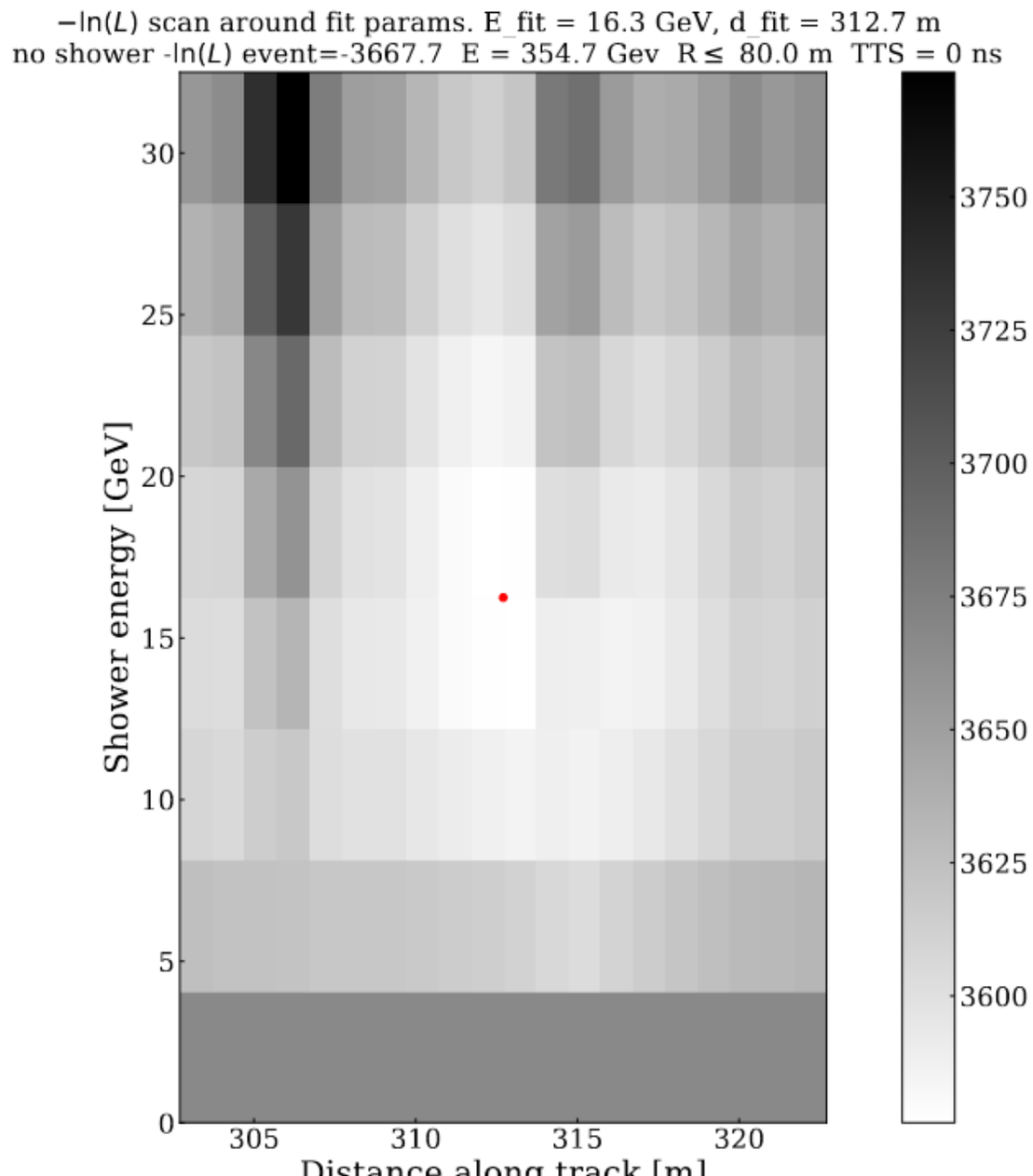
E_muon = 354.7 GeV

Why all these local minima? Maybe it's a geometrical effect of the detector → work in progress...



MC-hits: ln(L) scan around fit (2)





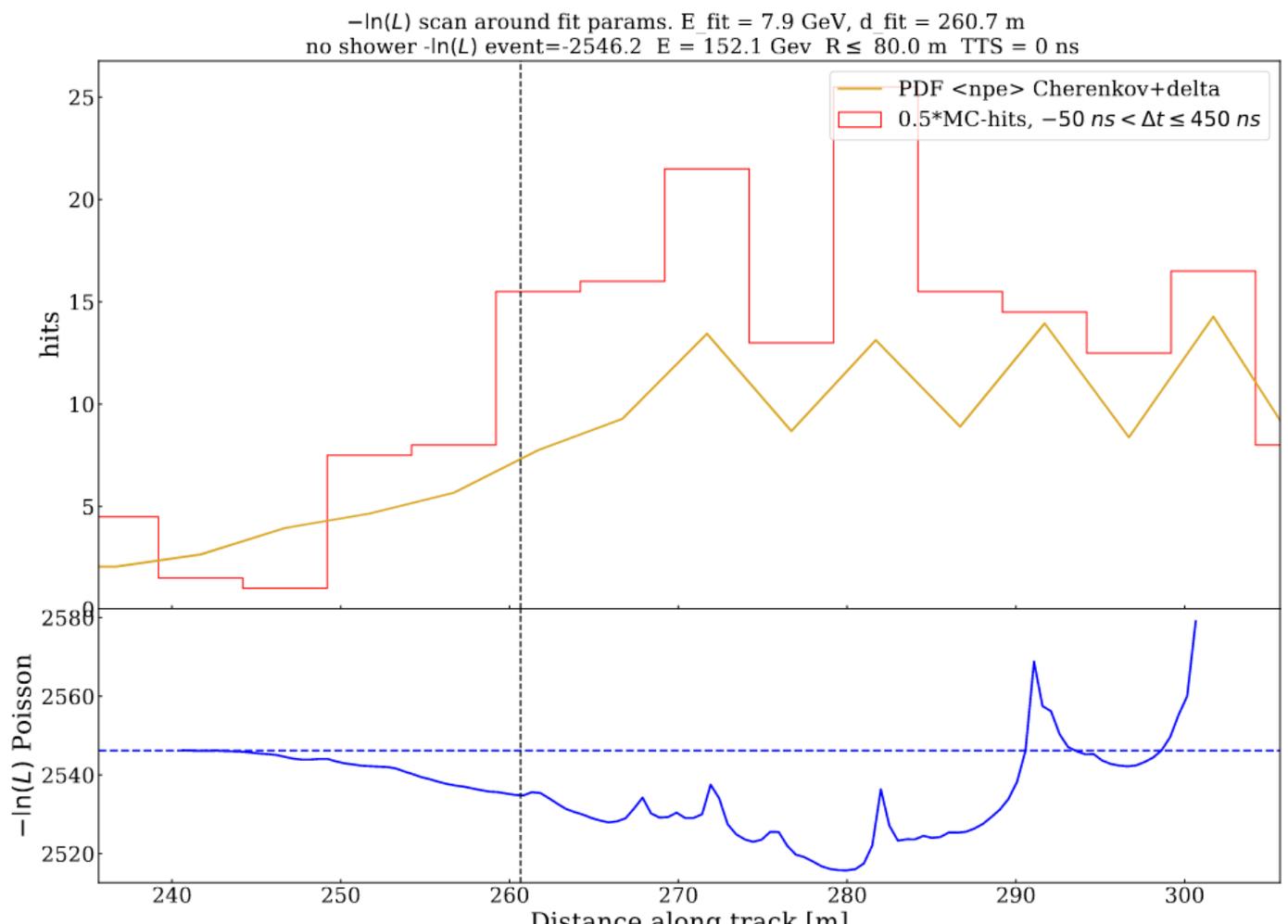
Evt = 1
Fix distance

E_muon =
354.7 GeV

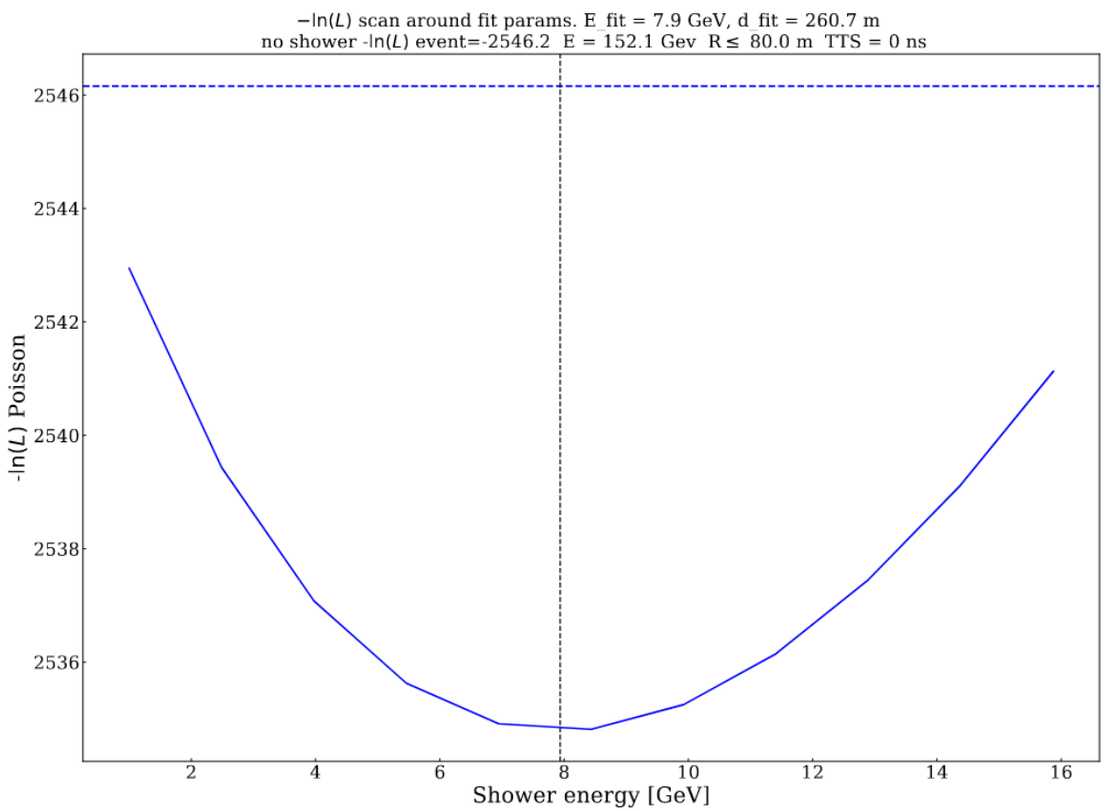
MC-hits: ln(L) scan around fit (3)

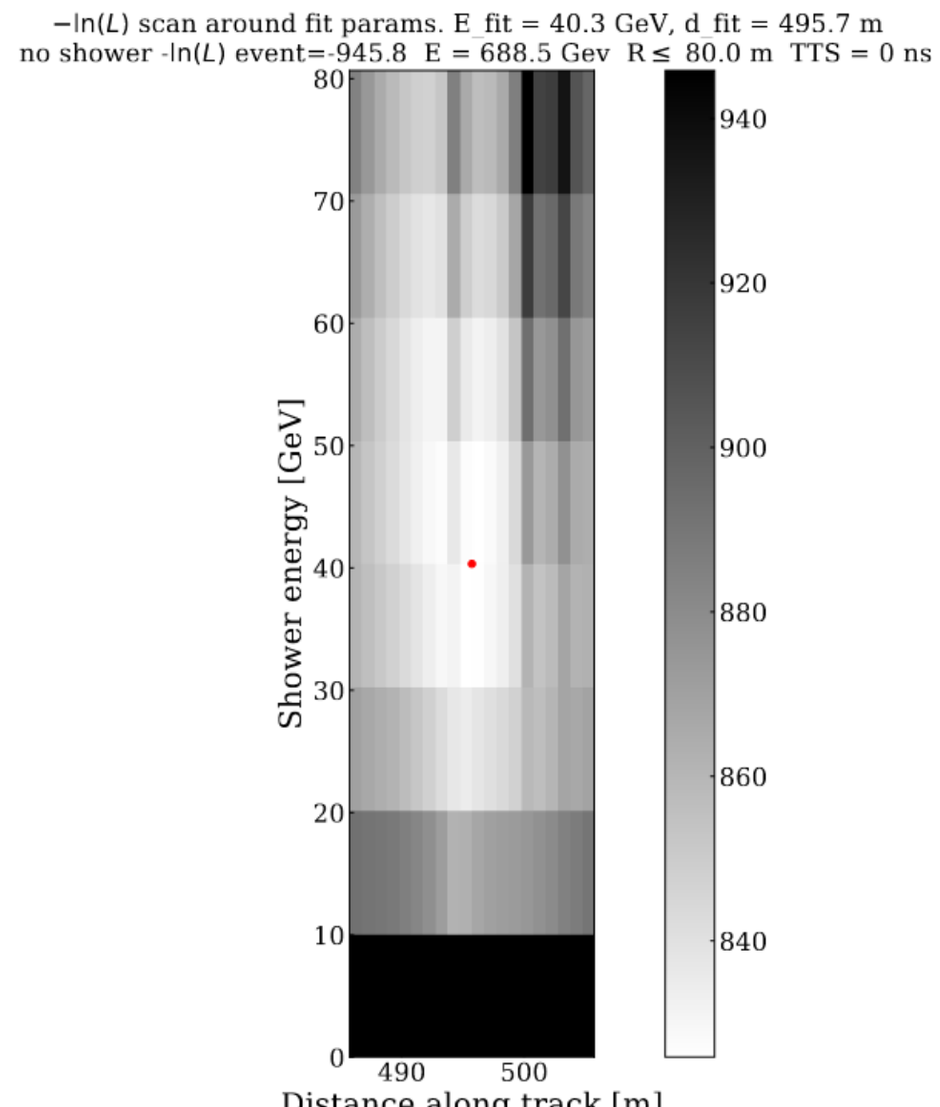
Evt = 10
Fix energy

Muon energy = 152.1 GeV



MC-hits: ln(L) scan around fit (4)





Evt = 10
Fix distance

Muon energy = 152.1 GeV

Moving on to reco track + hits

Corrections:

- TTS = 1.3 ns
- Time slewing correction (using JPP slew function)
- 1st hit probability
- Background

$$\frac{dP}{dt}(t) = \sum_{i=1,2,5,6,13,14} \frac{dP_i}{dt}(t) + f_b$$

If multiple hits at same PMT: take first

Only take hits within -50 < t_res < 450 ns (PDF limits!)

Hits: shower fits

evt_nr	E_muon [GeV]	E_reco [GeV]	5 sigma H0	<npe> no_shwr	#hits	<npe> fit	ln_L_no_shwr	ln_L_fit	E_shwr [GeV]	d_shwr [m]	<npe> shwrs
1	354.658	259.09	no	586	685	687	-6094.7	-6078.7	[8.8]	[60.0]	[66]
4	530.014	421.7	yes	477	612	589	-5345.2	-5321.2	[24.1]	[21.1]	[59]
10	152.11	74.52	yes	499	774	599	-6836.8	-6823.4	[14.1]	[19.3]	[58]
13	726.116	557.12	yes	616	1446	1373	-12279.3	-11854.9	[70.0, 50.0]	[32.6, 181.4]	[381, 273]

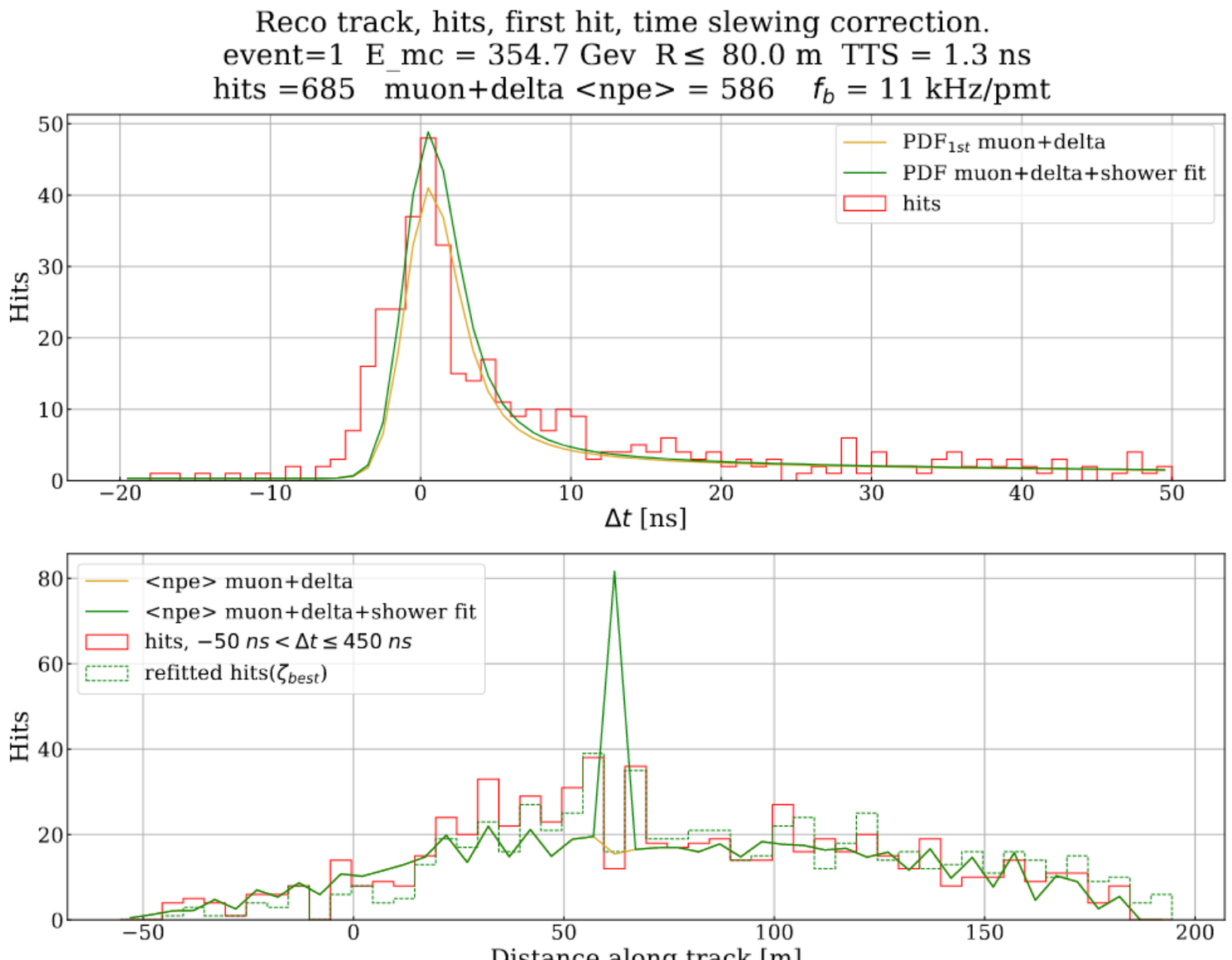
evt_nr	E_muon [GeV]	E_reco [GeV]	5 sigma H0	<npe> no_shwr	#hits	<npe> fit	ln_L_no_shwr	ln_L_fit	E_shwr [GeV]	d_shwr [m]	<npe> shwrs
19	688.5	476.62	no	107	149	164	-1376.7	-1344.5	[28.2]	[57.2]	[55]

Just for comparison I include the shower fits for MC-hits below

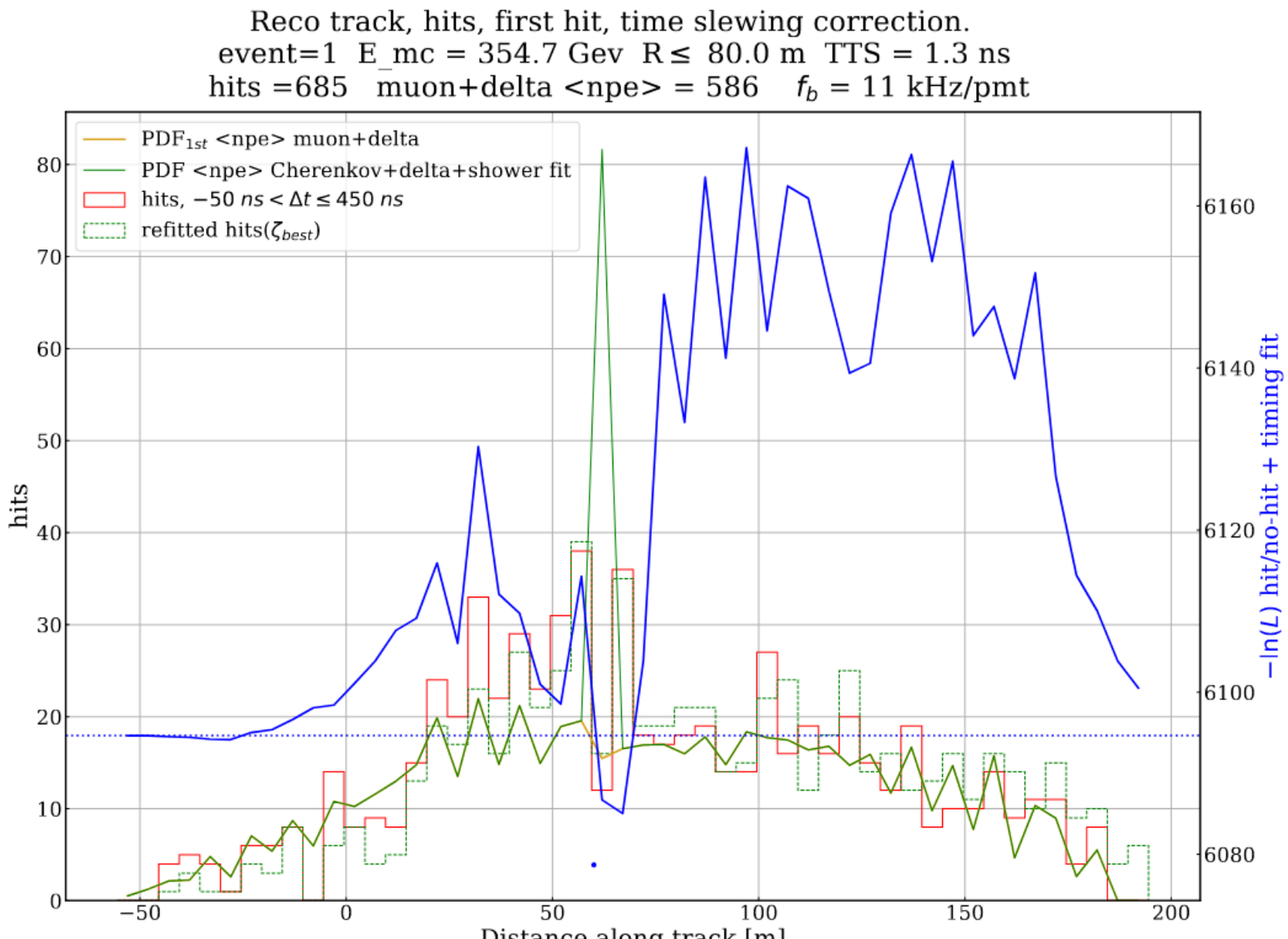
evt_nr	E_muon [GeV]	5 sigma H0	<npe> MIP	MC-hits	<npe> fit	ln_L_no_shwr	ln_L_fit	E_shwr [GeV]	d_shwr [m]	<npe> shwrs
1	354.658	yes	450	624	568	-3667.7	-3575.0	[16.3]	[312.7]	[118]
4	530.014	no	417	492	448	-2618.7	-2600.8	[12.3]	[265.1]	[31]
10	152.11	no	395	458	422	-2546.2	-2534.8	[7.9]	[260.7]	[27]
13	726.116	yes	592	1350	1312	-6438.6	-4960.4	[77.6, 55.7]	[367.6, 517.3]	[418, 302]

evt_nr	E_muon [GeV]	5 sigma H0	<npe> MIP	MC-hits	<npe> fit	ln_L_no_shwr	ln_L_fit	E_shwr [GeV]	d_shwr [m]	<npe> shwrs
8	742.269	yes	188	262	190	-1234.2	-1234.1	[0.7]	[425.3]	[1]
19	688.5	yes	63	179	137	-945.8	-824.9	[40.3]	[495.7]	[74]

Hits: shower example (1)



Hits: shower example (2)



Hits: two ‘problems’ (1)

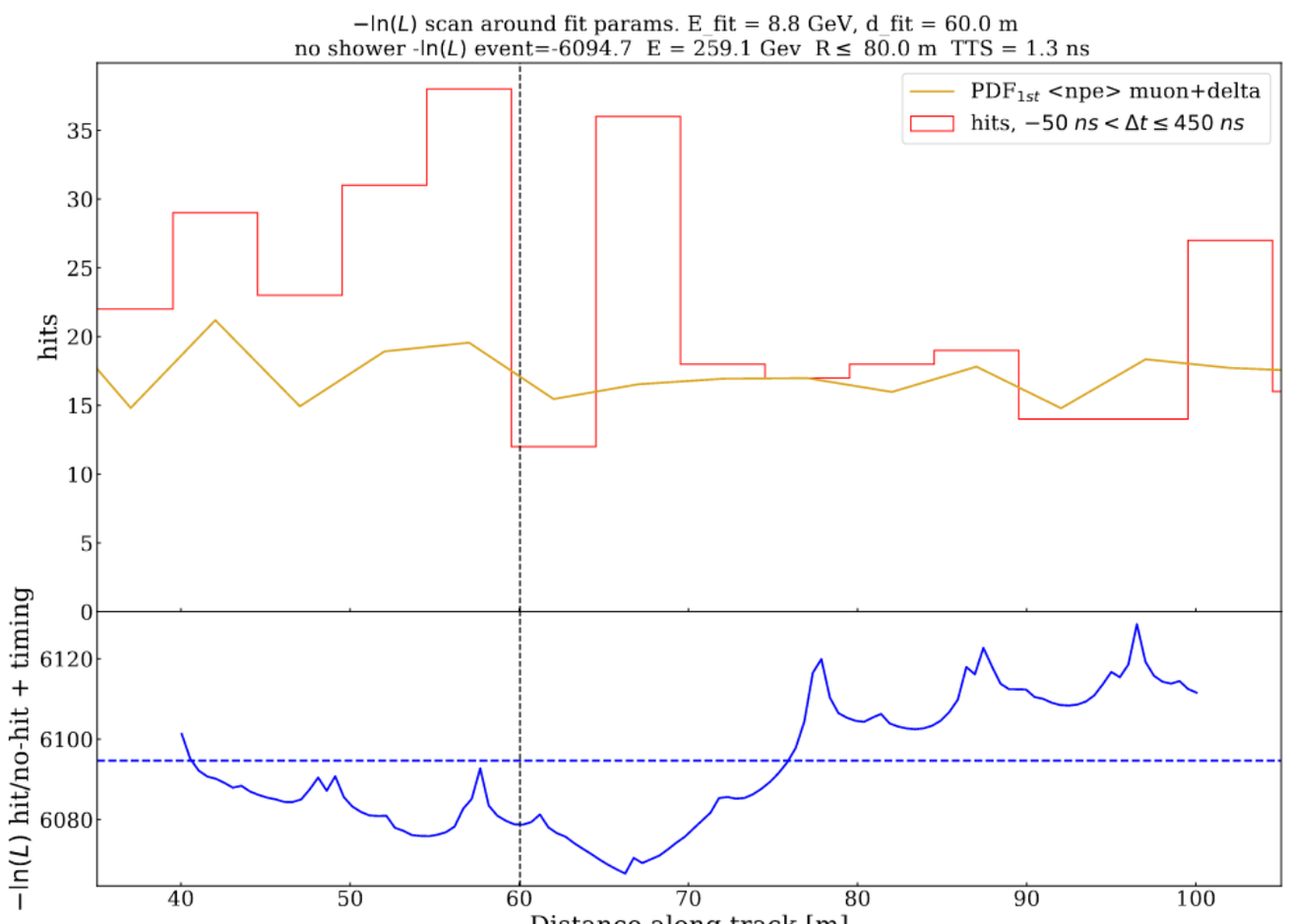
First problem: local minima. This is a coding issue. Using scipy.minimize, now looking at scipy.minimize.shgo or scipy.minimize.basinhopping.

Event =1

Reco track, hits, E_reco

Energy fixed

(muon E_mc = 354.7 GeV)

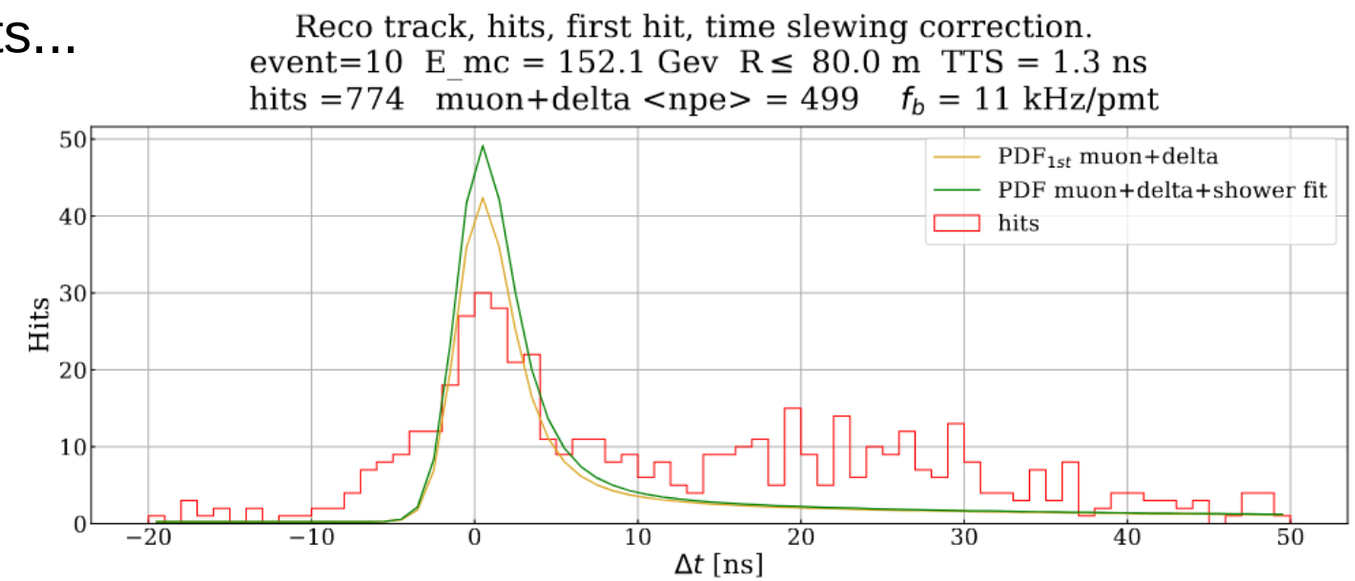


Hits: two ‘problems’ (2)

Second problem: time residuals and #hits...

Am I missing hits or is it blurring?

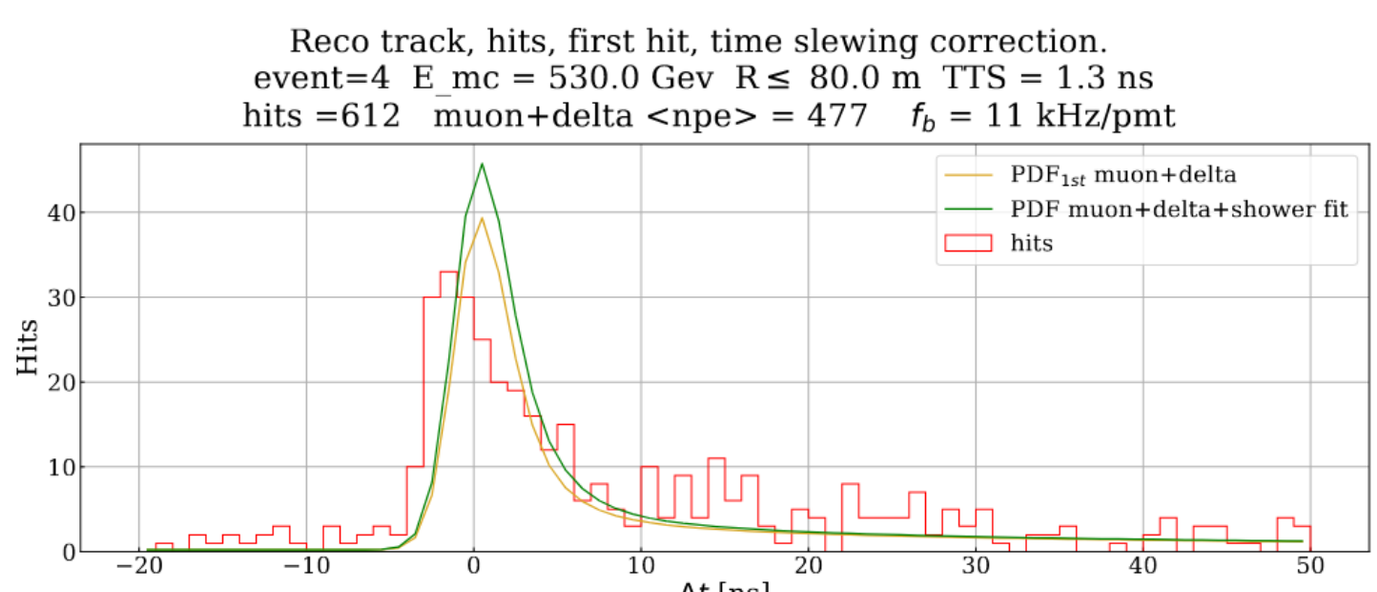
What could cause this:
Triggers? PMT simulation?
Wrong slewing correction?



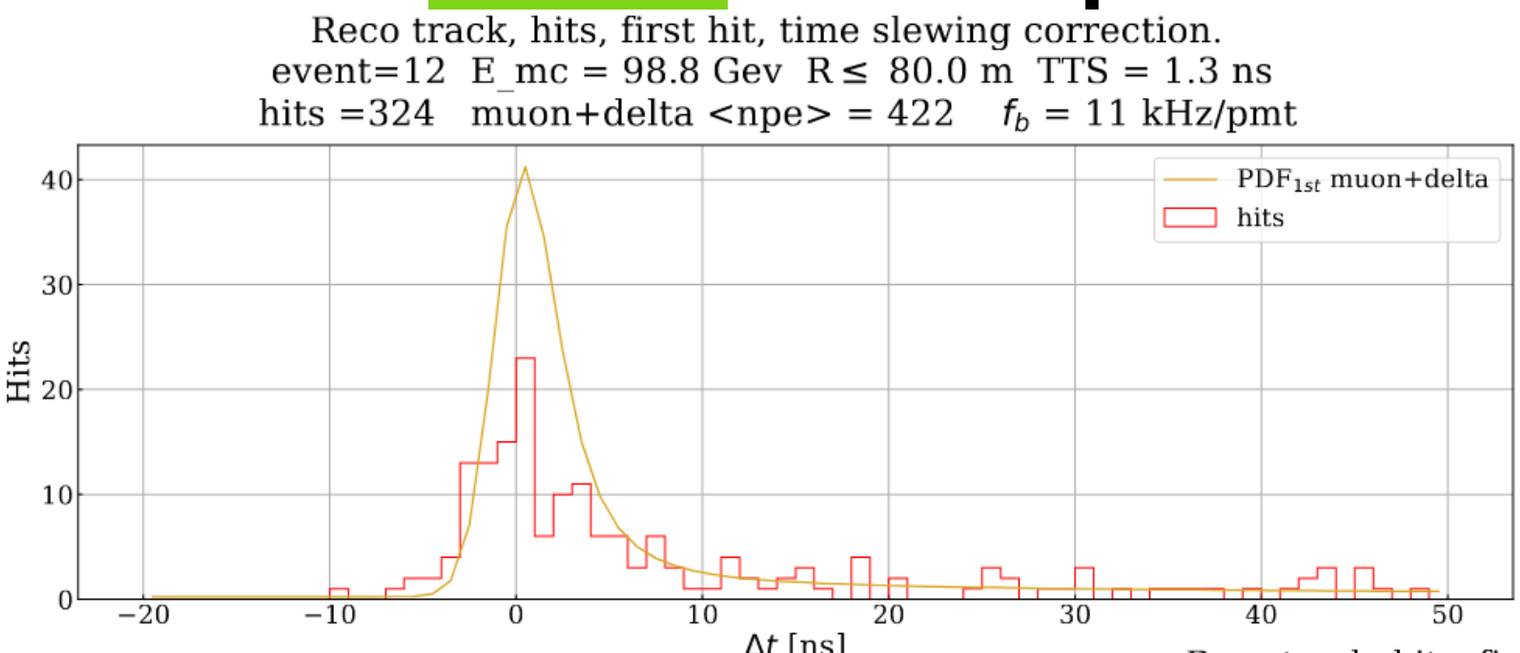
Possibly related issue:

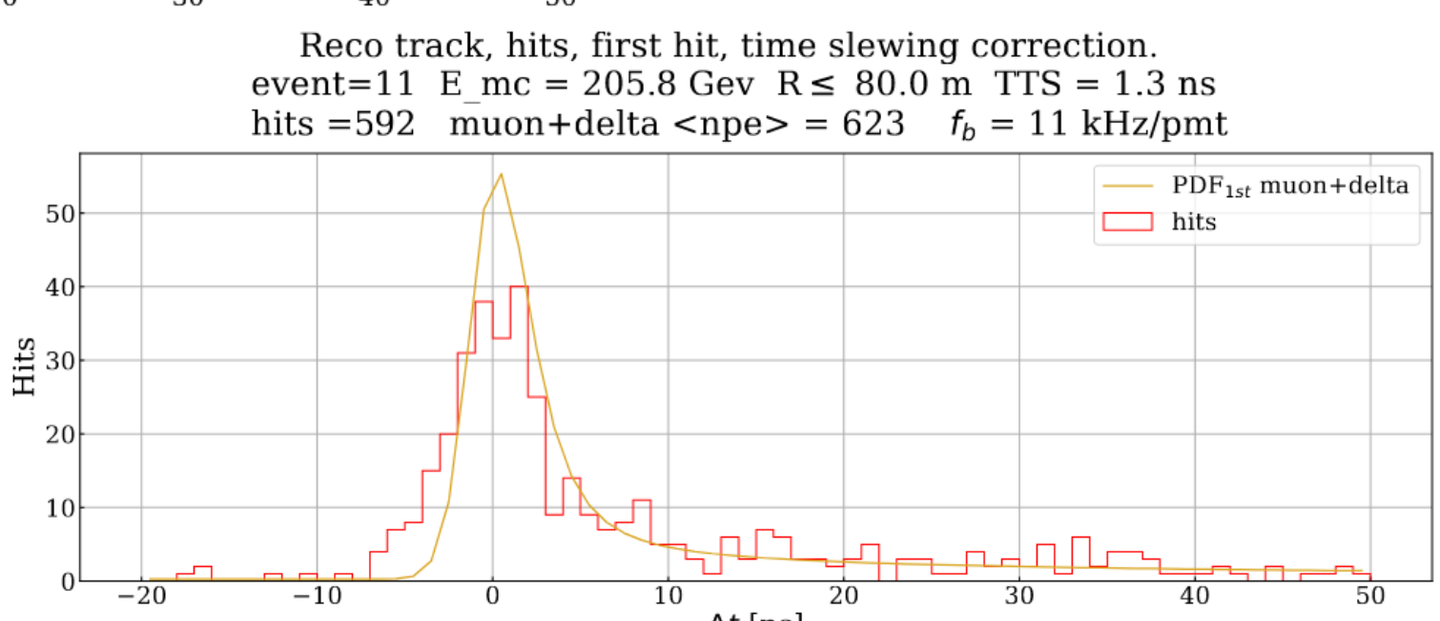
for hits, reco track in 10/21 the #hits is smaller than <npe> for muon+delta only

For mc-hits, true track in 3/21 the #MC-hits is smaller than <npe> for muon+delta only



Hits: two ‘problems’ (3)





Conclusion

- Started with shower fitting for both mc-hits, true track and hits, reco track. Shower energy has same order of magnitude but varying by a factor of ~2.
- For -ln(L) minimization: have to improve avoiding local-minima
- Time residuals and #hits for hits, reco track don't look too great. May due to some blurring or missing effects that suppress the #hits?