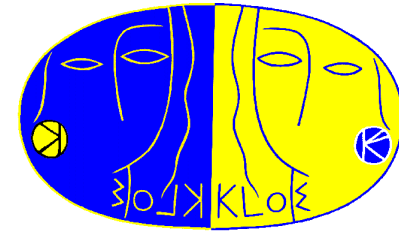


Large scale MC production II



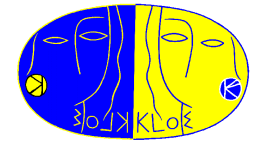
M. Moulson

Offline Discussion, 30 October 2002

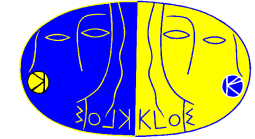
Outline

- *Status reports*
 - *Logistics: a proposal*
 - *Combine or separate channels?*
 - *Timescale and checklist*
-

Status reports



M. Antonelli	ISR in the MC
A. Antonelli, S. Dell'Agnello	DC geometry review
S. Miscetti, M. Palutan	EmC geometry review
M. Moulson	Bank-reduction code for DST's
C. Bloise, S. Miscetti	QCAL banks for MC DST's
M. Moulson, S. Miscetti, V. Patera	Background insertion
S. Giovannella	GEANFI on IBM
I. Sfiligoi	DB modifications for DST's



MC DST's: A/C path

If an identified K_L tag, recalculate t_0 and re-run NVR code:

- talk kloe_drop DROP VNVO 0 DROP KNVO 0 DROP INVO 0
- talk eclsfilt sel kltag yes ret ret
- use/path=1 kbkmd d emcdbini dcd bini eclsfilt kloe_drop
t0_find/par=2 clufixtim/par=2 eclmod/par=2
- filter/path=1 eclsfilt on

Compress banks (all events, including unclassified events and K_L tags)

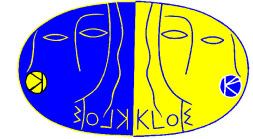
- use/path=2 **mc_dst** **clu2trig** **track_dst** **trig_dst**
- output select event/path=(1,2)

Zip output (level 4)

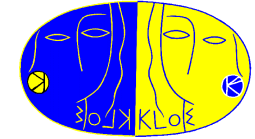
Specify bank list

- output select drop
 - output select kept
-

MC DST's: banks currently present



Headers, etc.	LRID HEAD EVCL BRIN
MC truth	KINE VERT
t_0 -related	T0MC T0GL
EmC recon.	CLPS CLLS CSPS
EmC truth	CFHI
Event class.	ECLS ECLO VNVO INVO KNVO
DC recon.	DCNH DTFS DVFS TCLO
DC truth	MKIN MDTF
Trigger	TCOE TDST CTRG
QCAL	QCAE QWRK QIHI QCKA



MC DST summary banks for tracking

MKIN: MC details for KINE tracks

One bank per charged KINE track, 20 words:

- Number of DHIT hits and layers
- **x, p** at first and last DHIT hits
- Path length and TOF

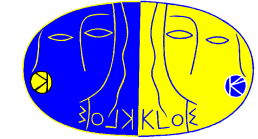
MDTF: MC truth for DTFS tracks

One bank per DTFS track, 28 words:

- Indices of 3 main KINE contributors
- Number of hits and number consecutive hits contributed by each
- Index of KINE at first DTFS hit; layer, **x, p** for first hit
- Index of KINE at last DTFS hit; layer, **x, p** for last hit

PROD2NTU routines to be modified:

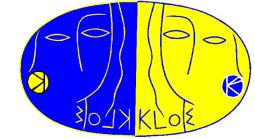
getdhitval, tfmcttrue, tfmchit, mainkineintrk, getfirstlastdhit



MC DST's: banks currently dropped

Trigger	TMON TELE TPIE PZZA TCAF TPAS
EmC recon.	MCEL CWRK CPPS CPLS CELE
DC recon.	MDTC DTCE DHRE DPRS DHSP DCHD
MC run init.	PART MATE TMED
MC truth	DHIT CHIT CEKA DTKA DTHA

Additionally, all lower-case MC banks compressed with **sqzlib**



MC DST's: status and size estimate

Output size estimate:

1000 $K_S \rightarrow$ all, $K_L \rightarrow$ all events

Generated and reconstructed on AIX w/ standard path

.mcr 23.9 MB (i.e., KB/evt)

.dst 4.2 MB

Target size:

6 KB/evt $\rightarrow$ 500 pb⁻¹ = 500 M evts = 3 TB

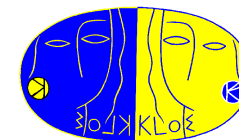
Work to be done:

Summary banks for EmC, trigger, QCAL

Need assistance with definition of these banks!

Time to complete development:

1 week full-time; 2 weeks realistically



Background insertion: status

S. Miscetti:

Selection criteria for clean $\gamma\gamma$ events

Nearly done

Isolation of $\gamma\gamma$ clusters/fragments

In progress

M. Moulson:

A/C module for event insertion

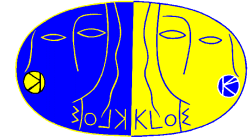
1st version working

V. Patera (A. Ventura):

Installation of MBCKADD routines
to merge DC hits; testing

In progress

A/C module for background insertion



New A/C module: **insert**

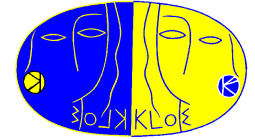
- Opens background file
- Reads events from the file into secondary YBOS array, reusing events appropriately
- Decompresses (unzips) events from the secondary file
- Gracefully handles EOF of the secondary file

Proof of principle: no analysis at present

Only copies LRID from 'background' file into BRID in the 'MC' file

Next step: open the secondary file using KID

Straightforward, just drop-in subroutine replacements.



Proposal for production logistics

Goal: Best possible reproduction of time-variable conditions

Mainly background, but also $\sqrt{s}$, $\mathbf{p}_\phi$, dead channels, etc.

These quantities available by run

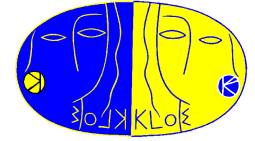
Proposal: generate an MC file for each run

Need to simulate chunks of runs for which $\int L dt$ available

$\sigma(K_S K_L) = 750 \text{ nb}$: Max MC file size 25000 evts $\approx$ **33 nb⁻¹**

Run size in 2001-2002 data **20 $\rightarrow$ 200 nb⁻¹**

Have $\int L dt$ for each raw file!



Proposal for production logistics

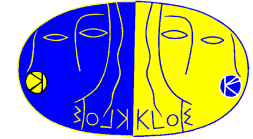
Generate MC files with number of events in proportion to luminosity in each raw file:

- Tailor generation to each run: $\sqrt{s}$, $\mathbf{p}_\phi$, $\mathbf{x}_\phi$ (others?)
- Tailor reconstruction to each run: dead channels, trigger (others?)
- Insert background from each raw file
- All necessary information should already exist

Relatively easy to implement; database driven

No need to define data set a priori

Individual analyses can use different MC data sets specific to their data



Proposal for production logistics

Number of reconstructed raw files 2001-2002: **84237**

Raw file size: **2-6 nb⁻¹ in 2001, 6-15 nb⁻¹ in 2002**; tails extend to zero

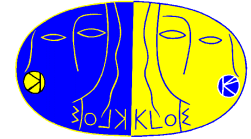
Problems can arise when individual raw files too small ($N = \int L \sigma dt \rightarrow 0$)

- Absurdly small number of events per MC file, fluctuations to zero
Not a problem for $K_S K_L$: $\sigma = 750$ nb
Rare channels have to be generated with something non-rare
- Fluctuations to zero in number of background events to insert

Can solve all of above problems by:

- Not generating rare channels alone ($K_S K_L$ slightly preferable here)
 - Summing f05-f07 for 2001 (reduce number of MC files)
 - Don't generate very small files: $\langle N_{\gamma\gamma} \rangle < 10 \rightarrow \int L dt < \sim 0.3 \text{ nb}^{-1}$
Cut assumes $\gamma\gamma$ sampling limiting factor: $\sigma(\text{MC}) \gg \sigma(\gamma\gamma \text{ background})$
Error on total luminosity = 0.04%
-

Miscellaneous upgrades to MC



Generators:

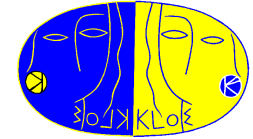
Any that need to be checked?

Branching ratios?

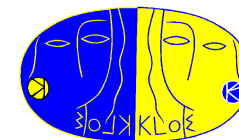
Is it time to review and update these?

Other issues?

Combine or separate neutral kaon runs?



Combined production $K_S \rightarrow \text{all}, K_L \rightarrow \text{all}$	Separate production $K_S \rightarrow \pi^0\pi^0, K_L \rightarrow \text{all}$ $K_S \rightarrow \pi^+\pi^-, K_L \rightarrow \text{all}$ Also differentiated by K_L decay in DC?
Simpler to produce	Simpler to analyze
Fewer files	Smaller files
Less disk turnover? (if people cooperate)	Less disk turnover? (if one K_S channel dominates interest)
No need to prioritize	Possible to prioritize
Naturally treats channels with small BR	
Well-suited for background studies for rare K_S decays	



Task list

ISR in the MC

DC geometry review

EmC geometry review

Bank-reduction code for DST's

Assistance w/ EmC DST structures

Assistance w/ trigger DST structures

Assistance w/ QCAL structures

Background insertion

DC dead channel simulation

EmC dead channels, thresholds, etc.

Trigger simulation by run

GEANFI on IBM

Other MC tuning (BR's, etc)

DB modifications for DST's

M. Antonelli

A. Antonelli, S. Dell'Agnello

S. Miscetti, M. Palutan

M. Moulson

S. Miscetti, M. Moulson, V. Patera

F. Bossi

S. Giovannella

I. Sfiligoi
