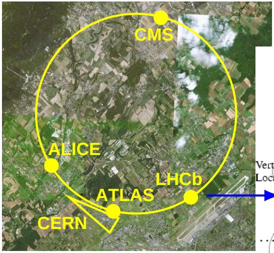
Lifetime Measurements at LHCb

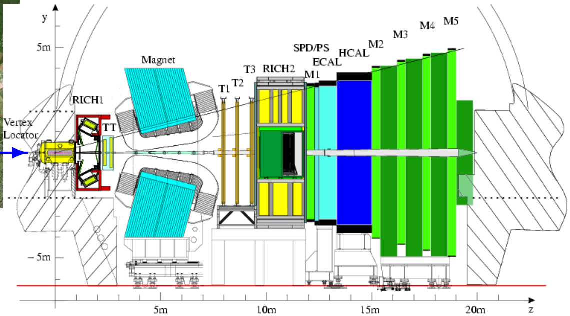
Marco Gersabeck

University of Glasgow

IOP HEPP Annual Meeting, Oxford, 06 April 2009

Flavour physics at its finest





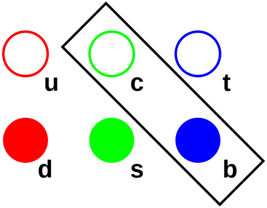
- LHCb: single arm spectrometer with world's most precise vertex detector and excellent tracking and particle identification system

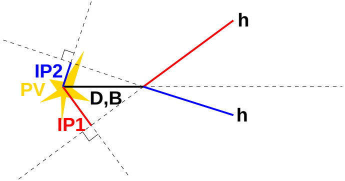
University of Glasgow

Lifetime Measurements with. . .

. . . heavy flavour charmless, charged two body decays

- Heavy flavour:
- D, B_d, B_s, Λ_b, ...





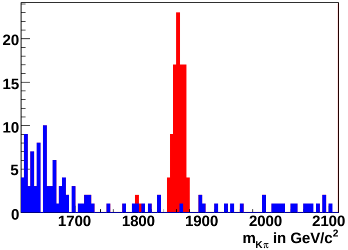
- Charmless, charged two body decays:
- $h = \pi^{\pm}, K^{\pm}, p^{\pm}$

- A broad range of decays
- Physics from SM tests to NP discoveries

University of Glasgow

1. $\tau(D)$ from $D \to K^- \pi^+$

- Abundance of data
- *D*s produced more frequently than *B*s
- $D \to K^- \pi^+$ has a large branching ratio (4%)
- Trigger rates of several Hz possible
- Can apply very tight selection
  $\Rightarrow$ almost background free
- First lifetime measurements after only days of data taking
- Current precision: 0.4%
  $\Rightarrow$ calibration measurement



University of Glasgow

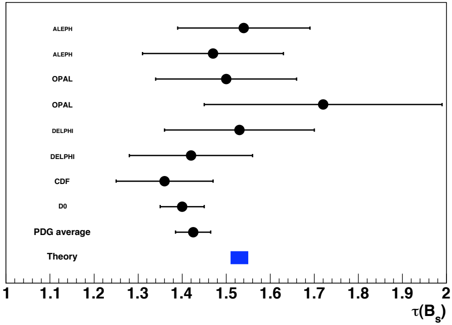
2. $\tau(D \to K\pi)/\tau(D \to hh)$ from $D \to hh'$

- New physics in the second course?
- Also $D \to hh$ are produced in abundance
- Lifetime ratios:
  $y_{\mathcal{CP}} = \frac{\tau_{K\pi}}{<\tau_{hh}>} - 1; \quad \Delta y = \frac{\tau_{K\pi}}{<\tau_{hh}>} A_\tau; \quad A_\tau = \frac{\tau_{hh}^{+} - \tau_{hh}^{-}}{\tau_{hh}^{+} + \tau_{hh}^{-}}$
- Compare $y_{\mathcal{CP}}$ to mixing parameter $y \equiv \Delta\Gamma/2\Gamma$

3 scenarios:

- No mixing: $y_{\mathcal{CP}} = \Delta y = 0$
- No $\mathcal{CP}$ violation: $y_{\mathcal{CP}} = y$, $\Delta y = 0$
- $\mathcal{CP}$ violation: $y_{\mathcal{CP}} \neq y$, $\Delta y \neq 0$
  $\Rightarrow$ any sizeable $\mathcal{CP}$ violation would be sign of NP
- Current HFAG average: $y_{\mathcal{CP}} = (1.07 \pm 0.26)\%$

3. $\tau(B_s)$ from $B_s \to KK$



$$\text{Theory} = [\tau(B_s)/\tau(B_d)]_{theory} \times \tau(B_d)_{PDG}$$

- Tension between theory and experiment
- Tension caused only by hadron collider results

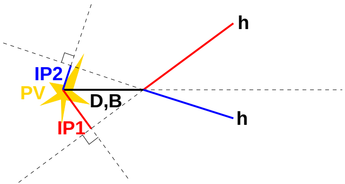
University of Glasgow

3. $\tau(B_s)$ from $B_s \to KK$

- K^+K^- is an (almost) $\mathcal{CP}$ even final state
  $\Rightarrow$ measure $\Gamma_s - \mathcal{A}_{\Delta\Gamma}\frac{1}{2}\Delta\Gamma_s$
- SM: $-1 \lesssim \mathcal{A}_{\Delta\Gamma} \lesssim -0.95$; $\Delta\Gamma_s/\Gamma_s \approx 0.1$
- NP: possible new mixing phase ϕ^{NP}
  or $\mathcal{CP}$ violating phase σ^{NP}
  $\Rightarrow$ measure $\Gamma_s - \mathcal{A}_{\Delta\Gamma}^{SM}\cos(\phi^{NP} + 2\sigma^{NP})\frac{1}{2}\Delta\Gamma_s^{SM}\cos(\phi^{NP})$
  $\Rightarrow$ compare to tree dominated $B_s \to (J/\psi\phi)_{\mathcal{CP}+}$

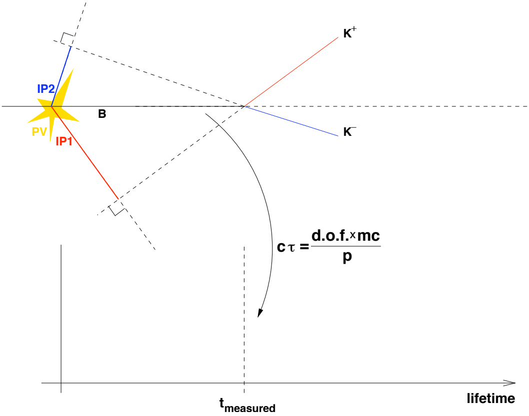
Measuring lifetimes

- Hadronic final states are lifetime biased by the trigger
  $\Rightarrow$ IP cuts are applied to enhance long-lived B sample
- Have to correct for this bias when measuring lifetimes



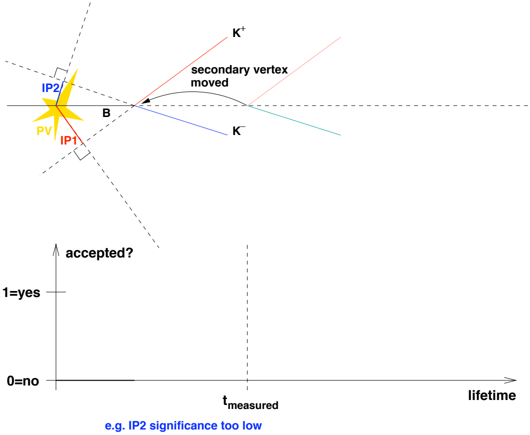
- Monte Carlo independent measurement desirable
- How to evaluate trigger acceptance without MC?
  $\Rightarrow$ Swim!

Event-by-event lifetime acceptance



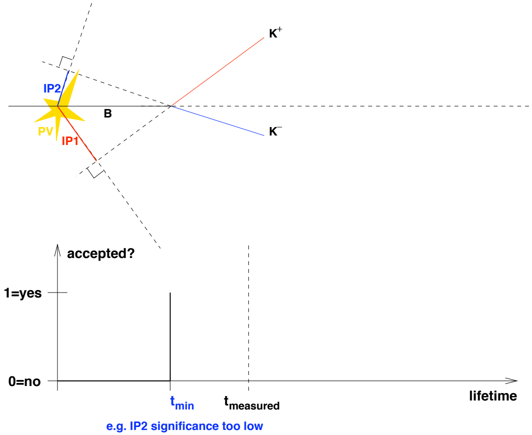
- Measure lifetime and determine event-by-event acceptance as function of lifetime

The ’swimming’



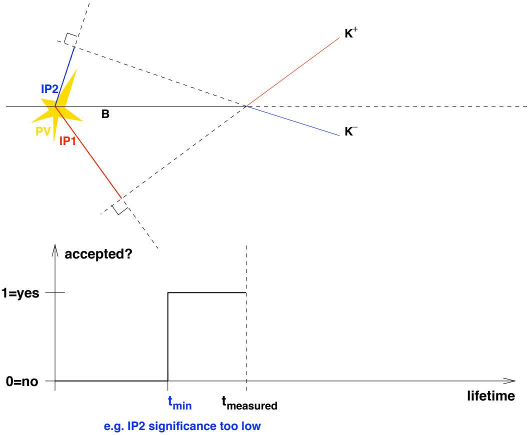
- ’Swim’ secondary vertex and re-evaluate trigger decision
- *IP* significance too small ⇒ *not accepted*

The ’swimming’



- *IP* significance at threshold ⇒ *just accepted*

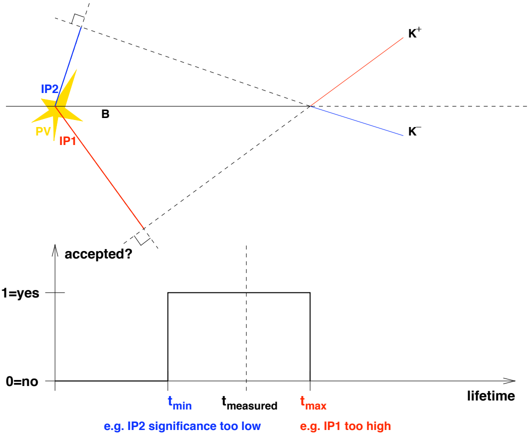
The 'swimming'



- Initially measured value ⇒ *accepted*

University of Glasgow

The 'swimming'



- *IP* at upper threshold ⇒ *just accepted*

In short

- 3 stages of the fit
  - Data selection: basically in line with standard $B \rightarrow hh$ selection
  - Swimming: determine trigger decision as function of lifetime according to decay kinematics as saved in NTuple
  - Fit: Produce likelihood function and perform an unbinned, combined mass-lifetime fit of signal and background

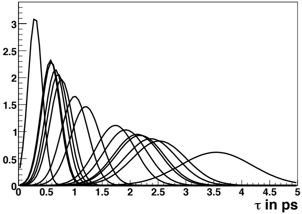
Fitting the signal:

$$P_\tau(t|sig) = \frac{\frac{1}{\tau}e^{-t/\tau}}{\int_{t_{min}}^{t_{max}} \frac{1}{\tau}e^{-t'/\tau}dt'} = \frac{\frac{1}{\tau}e^{-t/\tau}}{e^{-t_{min}/\tau} - e^{-t_{max}/\tau}}$$

How to model the non-specific background time PDF?

- Best assumption for a background PDF is no assumption
- Obtain smooth background PDF with per-event kernel functions

$$P(t|bg) = \frac{\sum_{bg} Gauss(t, t_i, \sigma)}{N_{bg}}$$



- Unparametrised PDFs used in LEP analyses for neural net output (K. Cranmer et al.),
  Introduced to LHCb for $B \rightarrow K^*\mu\mu$ by F. Marinho

University of Glasgow

Sensitivities

- Only statistical uncertainties given
- B/S should be much smaller than 1
- Systematic limits expected at per mille level

Year*	$\mathcal{L}_{int}$ (fb^{-1})	$D \to K\pi$ yield	$D \to K\pi$ stat. error	$B_s \to KK$ yield	$B_s \to KK$ stat. error
2009	0.005	$\mathcal{O}(10^5)$	0.3%	185	7.4%
2010	0.2	$\mathcal{O}(10^6)$	0.1%	7400	1.2%
2011	2	$\mathcal{O}(10^7)$	0.03%	74000	0.4%
2015	10	$\mathcal{O}(10^8)$	0.01%	370000	0.2%

* indicative!

- Recall: B_s lifetime too low w.r.t. theory by almost 10%

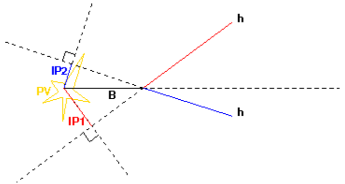
Conclusion

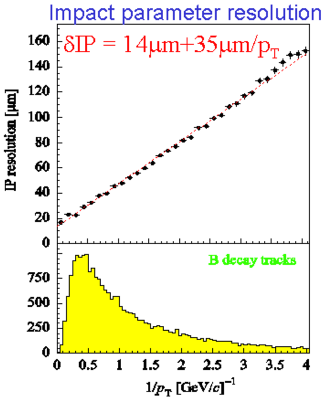
- Lifetime measurements offer many opportunities for interesting physics
- Complete MC-free fitting technique developed
- D lifetime and lifetime ratio are early measurements
- $B_s \to KK$: lifetime measurements with possible NP
- Many opportunities of $B/D \to hh'$ decays not covered: $\Rightarrow$ B lifetime ratios, Baryonic lifetimes, $\mathcal{CP}$ violation, CKM angles α and γ, rare decays, …

- SPARES

The VErtex LOcator

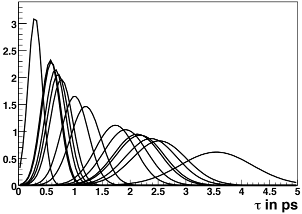
- Vertex resolution:
  $\sigma_{x,y} \approx 10\ \mu\text{m},\ \sigma_z \approx 40\ \mu\text{m}$
- Proper time resolution:
  $\sigma_\tau \approx 40$ fs
- Impact parameter resolution:
  $\sigma_{IP} \approx 40\ \mu\text{m}$





The 'kernel' method

- Use kernel functions to obtain a smooth distribution from a finite number of events
- Adjust the width of the kernel functions to reflect the density of events
  $\Rightarrow$ balance wiggles from statistical fluctuations to oversmoothing
- In reality: Sample all events an subtract known signal distribution to obtain pure background PDF
- Can be applied to sample virtually any distribution



University of Glasgow

What can we expect?

- Try to estimate numbers for $5\ \mathrm{pb}^{-1}$
- Inclusive $c\bar{c}$ cross-section: $3.6\ \mathrm{mb}$
  $\Rightarrow 18 \times 10^9 c\bar{c}$
- Probability for at least one c to hadronise to a D^* and to produce a D^0: ??? (assume 20%)
  (B analogon would yield 34%, $BR(D^* \to D\pi) = 68\%$)
- $BR(D \to h^+ h'^-) \approx 4\%$
- Trigger efficiency: ??? (assume 10%)
- Geometrical acceptance, reconstruction & selection efficiency: ??? (assume 1%)

$\Rightarrow$ $D \to h^+ h'^-$ yield: 144×10^3

Physics opportunities

- $\mathcal{CP}$ violation in decays of $D \to KK/\pi\pi$
  Current publications with 30k/50k $D \to \pi\pi$ (5k)
- D-Mixing using DCS $D \to K^+\pi^-$
  Current publications with 4k/4k/13k $D \to K^+\pi^-$ (0.5k)
- ($\mathcal{CP}$ violation in) D-mixing via lifetime ratio
  Current publications with 70k/110k $D \to KK$ (13k)
- D lifetime itself
  Currently measured to 0.4% (125k $\Rightarrow$ sub-% possible)

At Glasgow focus on lifetime measurements
Complemetary to Oxford activities

(LHCb-2009 expectations in brackets)

University of Glasgow

Lifetimes with $D \to h^+ h'^-$?

Why?

- We might have a sizeable amount of data in 2009
- This amount might even be compatible with existing data sets
- At the very least good calibration for (not only lifetime) tools as similar to $B_{(s)} \to h^+ h'^-$

What?

- Lifetime itself: starting point, calibrate lifetime fitting method
  $D \to K^- \pi^+$, 125k
- Lifetime ratio
  $y_{\mathcal{CP}} = \frac{\tau_{K\pi}}{<\tau_{hh}>} - 1$; $\Delta y = \frac{\tau_{K\pi}}{<\tau_{hh}>} A_\tau$; $A_\tau = \frac{\tau^+_{hh} - \tau^-_{hh}}{\tau^+_{hh} + \tau^-_{hh}}$
  $D \to K^- K^+$, 13k; $D \to \pi^- \pi^+$, 5k

University of Glasgow

The total probability

The total probability for each event to observe a decay at time t with mass m is

$$P_{\tau,sf}(t,m) = P_{\tau}(t|m) \times P_{sf}(m)$$

with the lifetime probability given the mass and the probability of observing this mass. The mass can be replaced by a set of discriminating variables, e.g. to include PID information.

Perform separate log-likelihood fits for signal fraction(s), sf, and lifetime(s), τ.

The mass probability

$P_{\tau,sf}(t,m) = P_\tau(t|m) \times P_{sf}(m)$

For a set of event classes (signal(s), background) the probability of measuring a mass m is given by

$$P_{sf}(m) = \sum_{class} P(m|class) \times \underbrace{P(class)}_{sf_{class}}$$

where $P(m|class)$ is the PDF of m for a given class and $P(class)$ is the probability of an event being of that class, i.e. $P(class) = N_{class}/N_{tot} = sf_{class}$ which will be a fit parameter.

Need $P(class|m)$ for time fit, which is given by

$$P(class|m) = \frac{P(m|class) \times P(class)}{P(m)}$$

The time probability

$P_{\tau,sf}(t,m) = P_\tau(t|m) \times P_{sf}(m)$

The lifetime part can be split up in signal and background components:

$$P_\tau(t|m) = \sum_{class} P_\tau(t|class) \times P(class|m)$$

where $P_\tau(t|class)$ are the different lifetime models for signal(s) and background, while $P(class|m)$ determines the probability of this event belonging to a certain class.

The background time PDF

$P_{\tau,sf}(t,m) = P_\tau(t|m) \times P_{sf}(m)$; $\quad P_\tau(t|m) = \sum_{class} P_\tau(t|class) \times P(class|m)$

Only in an ideal world can we sum exclusively over a pure background sample

$$P(t|bg) = \frac{\sum_{bg} Gauss(t, t_i, \sigma) \times P(bg|m_i)}{\sum_{bg} P(bg|m_i)}$$

BUT the sum here goes over all events:

$$P(t|bg) = \frac{\sum_{all} Gauss(t, t_i, \sigma) \times P(bg|m_i)}{\underbrace{\sum_{all} P(bg|m_i)}_{N_{bg}}}$$

$$= \frac{\sum_{bg} Gauss(t, t_i, \sigma) \times P(bg|m_i) + \sum_{sig} Gauss(t, t_i, \sigma) \times P(bg|m_i)}{\sum_{all} P(bg|m_i)}$$

An improved background time PDF

$P_{\tau,sf}(t,m) = P_\tau(t|m) \times P_{sf}(m)$; $P_\tau(t|m) = \sum_{class} P_\tau(t|class) \times P(class|m)$

Try to subtract signal contribution in numerator

$$\sum_{all} Gauss(t, t_i, \sigma) = \sum_{bg} Gauss(t, t_i, \sigma) + \sum_{sig} Gauss(t, t_i, \sigma),$$

where

$\sum_{bg} Gauss(t, t_i, \sigma) = \sum_{bg} Gauss(t, t_i, \sigma) \times P(bg|m_i) = N_{bg} \times P(t|bg)$,

and

$\sum_{sig} Gauss(t, t_i, \sigma) = N_{sig} \times P(t|sig) = \sum_{all} P(t|sig) \times P(sig|m_i)$.

Hence

$$P(t|bg) = \frac{\sum_{all} (Gauss(t, t_i, \sigma) - P(t|sig) \times P(sig|m_i))}{\sum_{all} P(bg|m_i)}$$

Speeding things up

$P_{\tau,sf}(t,m) = P_\tau(t|m) \times P_{sf}(m); \qquad P_\tau(t|m) = \sum_{class} P_\tau(t|class) \times P(class|m)$

The standard implementation involves a loop over all events for each event to evaluate the Gaussians:

$$P(t|bg) = \frac{\sum_{all}\left(Gauss(t, t_i, \sigma) - P(t|sig) \times P(sig|m_i)\right)}{\sum_{all} P(bg|m_i)}$$

Abbreviate where possible:

$$P(t|bg) = \frac{\overbrace{\sum_{all} Gauss(t, t_i, \sigma)}^{f(t)} - P(t|sig) \times \overbrace{\sum_{all} P(sig|m_i)}^{const}}{\underbrace{\sum_{all} P(bg|m_i)}_{const}}$$

$f(t)$ is independent of τ and hence can be calculated before the fit. Use a finely binned histogram to determine $f(t)$ prior to fit.

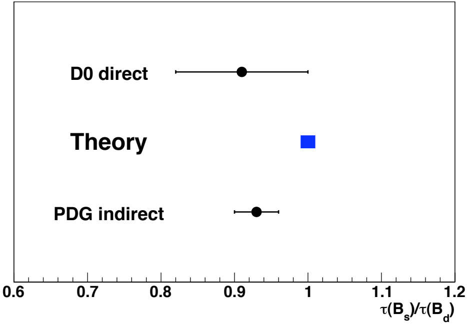
Lifetime measurements

Served in five courses:

1. $\tau(D)$ from $D \to K^- \pi^+$
2. $\tau(D \to K\pi)/\tau(D \to hh)$ from $D \to hh'$
3. $\tau(B_s)$ from $B_s \to KK$
4. $\tau(B_s)/\tau(B_d)$ from the same
5. $\tau(\Lambda_b)/\tau(B_d)$ from $\Lambda_b \to ph$ and $B_d \to \pi h$

- Similar final states
- Coherence and synergy effects in analyses

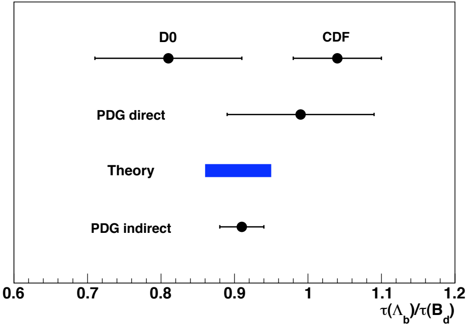
4. $\tau(B_s)/\tau(B_d)$ from $B_{(s)} \to h^+ h'^-$



PDG indirect = $\tau(B_s)_{PDG}/\tau(B_d)_{PDG}$

- Same tension as mentioned above
- No precise direct measurement yet
- Can test HQE to the % level

5. $\tau(\Lambda_b)/\tau(B_d)$ from $\Lambda_b \to ph$ and $B_d \to \pi h$



PDG indirect = $\tau(\Lambda_b)_{PDG}/\tau(B_d)_{PDG}$

- Two direct measurements in disagreement
- Potential for better theory prediction

Systematic effects: misalignments

- Effects on resolutions from both VELO & IT/OT misalignments for $B_{(s)} \rightarrow h^+ h'^-$ events.

Resolution	Affected by VELO misalignments	Affected by T misalignments
π momentum	NO	YES
B mass	NO	YES
B vertex	YES	NO
B IP	YES	NO
B $c\tau$	YES	NO

NO = very small/ negligible effects

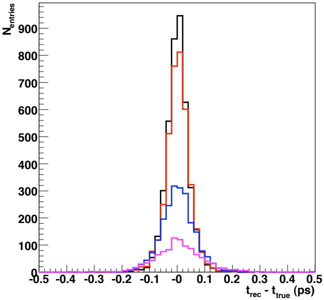
YES = significant effects

Effect on proper time resolution - I

Random misalignments

- Proper time resolution after standard $B_{(s)} \rightarrow h^+ h'^-$ selection with VELO & T misalignments.

	proper time resolution (in fs, sgl. Gaussian)
0σ	37.7
1σ	40.9
3σ	58.0
5σ	78.6



0σ
1σ
3σ
5σ

- Full study including many more variables and effects on physics parameters ($\mathcal{CP}$ asymmetries, lifetimes) is ongoing and will be presented in due time.

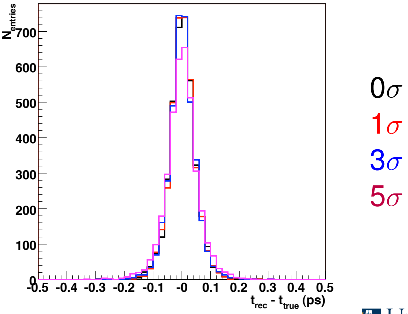
University of Glasgow

Effect on proper time resolution - II

z-scaling

- Proper time resolution after standard $B_{(s)} \to h^+ h'^-$ selection.

z-scale	proper time resolution (in fs)	bias (in fs)
1.00000	37.7	2.3
1.00003	37.7	2.3
1.00010	37.7	2.6
1.00033	38.5	2.6
1.00100	46.8	1.9



In practical terms

- Create smooth distribution with kernels
- Sample in variable steps
- Save sampling values to text file
- Read back and linearly interpolate

$\Rightarrow$ Fitting and toy generation