Beyond-the-standard-model contributions to rare B decays analyzed with variational-Bayes enhanced adaptive importance sampling

Stephan Jahn

March 16, 2015

Overview

Bayes' formula:

$$P(\boldsymbol{\theta}|\mathcal{D},\mathrm{M}) = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{P(\mathcal{D}|\mathrm{M})} = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{\int P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})d\boldsymbol{\theta}}$$

Bayes' formula:

$$P(\boldsymbol{\theta}|\mathcal{D},\mathrm{M}) = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{P(\mathcal{D}|\mathrm{M})} = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{\int P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})d\boldsymbol{\theta}}$$

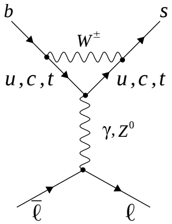
model independent search for new physics (effective theory):

Overview

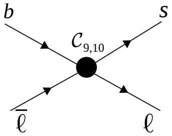
Bayes' formula:

$$P(\boldsymbol{\theta}|\mathcal{D},\mathrm{M}) = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{P(\mathcal{D}|\mathrm{M})} = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{\int P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})d\boldsymbol{\theta}}$$

model independent search for new physics (effective theory):



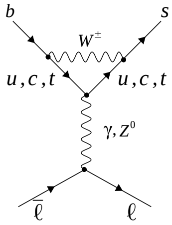
$$\boldsymbol{\theta} = \text{effective couplings } \mathcal{C}_i, \ldots$$



Bayes' formula:

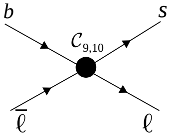
$$P(\boldsymbol{\theta}|\mathcal{D},\mathrm{M}) = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{P(\mathcal{D}|\mathrm{M})} = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{\int P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})d\boldsymbol{\theta}}$$

model independent search for new physics (effective theory):



$$\boldsymbol{\theta} = \text{effective couplings } \mathcal{C}_i, \ldots$$

$$\mathcal{D} = \text{detector events}$$

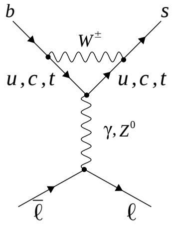


Overview

Bayes' formula:

$$P(\boldsymbol{\theta}|\mathcal{D},\mathrm{M}) = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{P(\mathcal{D}|\mathrm{M})} = \frac{P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})}{\int P(\mathcal{D}|\boldsymbol{\theta},\mathrm{M})P(\boldsymbol{\theta}|\mathrm{M})d\boldsymbol{\theta}}$$

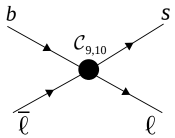
model independent search for new physics (effective theory):



$$\boldsymbol{\theta} = \text{effective couplings } \mathcal{C}_i, \ldots$$

$$\mathcal{D} = \text{detector events}$$

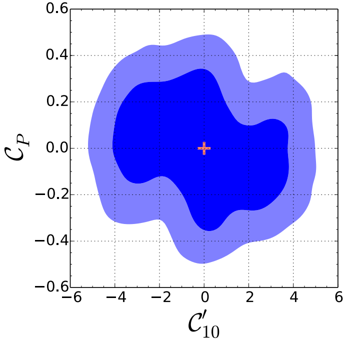
$$\mathrm{M} = \mathrm{EFT}, \mathrm{SM}, \ldots$$



Goals

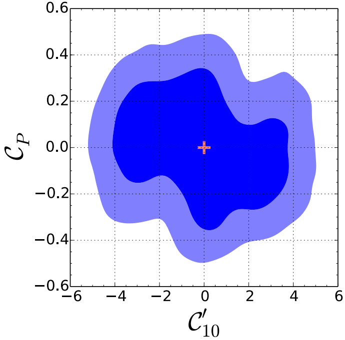
Goals

- draw marginal plots of the posterior



Goals

- draw marginal plots of the posterior



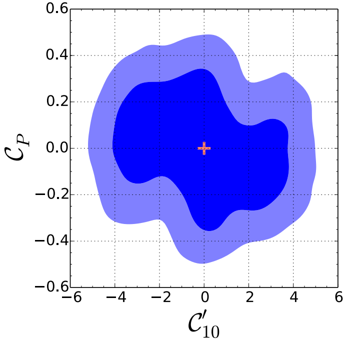
- compare models
  $\mathrm{NP} \leftrightarrow \mathrm{SM}$

$$\frac{P(\mathrm{NP}|\mathcal{D})}{P(\mathrm{SM}|\mathcal{D})} = \frac{P(\mathcal{D}|\mathrm{NP})}{P(\mathcal{D}|\mathrm{SM})} \cdot \frac{P(\mathrm{NP})}{P(\mathrm{SM})}$$

$$P(\mathrm{M}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathrm{M})P(\mathrm{M})}{P(\mathcal{D})}$$

Goals

- draw marginal plots of the posterior



- compare models
  $\mathrm{NP} \leftrightarrow \mathrm{SM}$

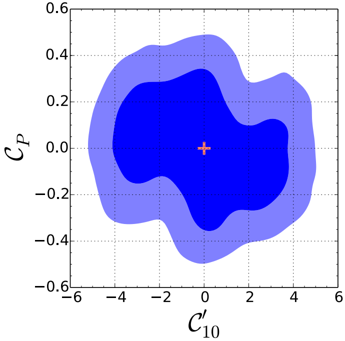
$$\frac{P(\mathrm{NP}|\mathcal{D})}{P(\mathrm{SM}|\mathcal{D})} = \frac{P(\mathcal{D}|\mathrm{NP})}{P(\mathcal{D}|\mathrm{SM})} \cdot \frac{P(\mathrm{NP})}{P(\mathrm{SM})}$$

$$P(\mathrm{M}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathrm{M})P(\mathrm{M})}{P(\mathcal{D})}$$

$\frac{P(\mathrm{NP}|\mathcal{D})}{P(\mathrm{SM}|\mathcal{D})} > 1$ new physics ☺

Goals

- draw marginal plots of the posterior



- compare models
  $\mathrm{NP} \leftrightarrow \mathrm{SM}$

$$\frac{P(\mathrm{NP}|\mathcal{D})}{P(\mathrm{SM}|\mathcal{D})} = \frac{P(\mathcal{D}|\mathrm{NP})}{P(\mathcal{D}|\mathrm{SM})} \cdot \frac{P(\mathrm{NP})}{P(\mathrm{SM})}$$

$$P(\mathrm{M}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathrm{M})P(\mathrm{M})}{P(\mathcal{D})}$$

$\frac{P(\mathrm{NP}|\mathcal{D})}{P(\mathrm{SM}|\mathcal{D})} > 1$ new physics ☺

$\frac{P(\mathrm{NP}|\mathcal{D})}{P(\mathrm{SM}|\mathcal{D})} < 1$ confirm SM ☹

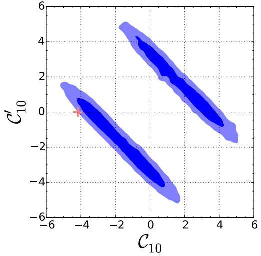
Difficulties

Difficulties

- curse of dimensionality

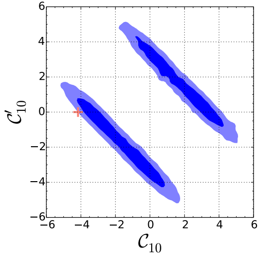
Difficulties

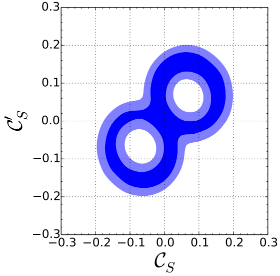
- curse of dimensionality
- multimodality



Difficulties

- curse of dimensionality
- multimodality
- degeneracies

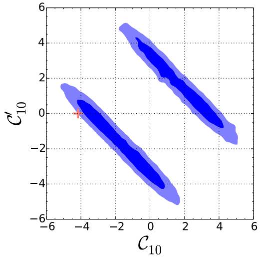


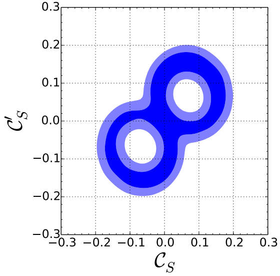


Difficulties

- curse of dimensionality
- multimodality
- degeneracies

no standard algorithm so far





Contents

Adaptive importance sampling with the variational-Bayes approach

Adaptive importance sampling

$$\int P(x)\mathrm{d}x = \int \frac{P(x)}{q(x)} q(x)\mathrm{d}x \approx \frac{1}{N}\sum_{n=1}^{N} \frac{P(x_n)}{q(x_n)} \equiv \hat{\mu}^N \;\; where \;\; x_n \sim q$$

Adaptive importance sampling

$$\int P(x)\mathrm{d}x = \int \frac{P(x)}{q(x)} q(x)\mathrm{d}x \approx \frac{1}{N}\sum_{n=1}^{N} \frac{P(x_n)}{q(x_n)} \equiv \hat{\mu}^N \quad \textit{where} \quad x_n \sim q$$

squared uncertainty (variance):

$$var(\hat{\mu}^N) = \frac{1}{N}\left[\int \frac{P(x)}{q(x)} P(x)\mathrm{d}x - \left(\int P(x)\mathrm{d}x\right)^2\right]$$

Adaptive importance sampling

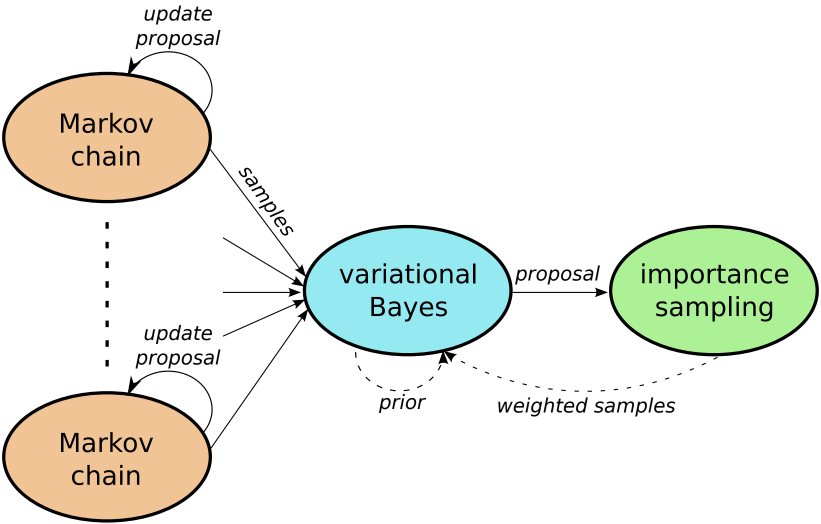
$$\int P(x)\mathrm{d}x = \int \frac{P(x)}{q(x)} q(x)\mathrm{d}x \approx \frac{1}{N}\sum_{n=1}^{N} \frac{P(x_n)}{q(x_n)} \equiv \hat{\mu}^N \quad where \quad x_n \sim q$$

squared uncertainty (variance):

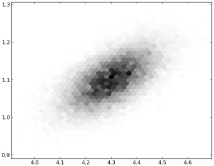
$$var(\hat{\mu}^N) = \frac{1}{N}\left[\int \frac{P(x)}{q(x)} P(x)\mathrm{d}x - \left(\int P(x)\mathrm{d}x\right)^2\right]$$

minimize $var(\hat{\mu}^N)$ with respect to the *proposal* q

Adaptive importance sampling with the variational-Bayes approach



```python
'''This example illustrates how to run a Markov Chain using pypmc'''

import numpy as np
import pypmc

# define a proposal
prop_dof   = 50.
prop_sigma = np.array([[0.1 , 0.  ]
                      ,[0.  , 0.02]])
prop = pypmc.density.student_t.LocalStud

# define the target; i.e., the function
# In this case, it is a Gaussian with me
# covariance "target_sigma".
#
# Note that the target function "log_tar
# unnormalized gaussian density.
target_sigma = np.array([[0.01 , 0.003 ]
                         ,[0.003, 0.0025]])
inv_target_sigma = np.linalg.inv(target_sigma)
target_mean  = np.array([4.3, 1.1])

def unnormalized_log_pdf_gauss(x, mu, inv_sigma):
    diff = x - mu
    return -0.5 * diff.dot(inv_sigma).dot(diff)

log_target = lambda x: unnormalized_log_pdf_gauss(x, target_mean, i

# choose a bad initialization
start = np.array([-2., 10.])

# define the markov chain object
mc = pypmc.sampler.markov_chain.AdaptiveMarkovChain(log_target, prc

# run burn-in
mc.run(10**4)

# delete burn-in from history
mc.history.clear()
```



(effective samples).

- **rel_tol** –
  Relative tolerance ϵ. If two consecutive values of the log likelihood bound, L_t, L_{t-1}, are close, declare convergence. More precisely, check that
  $$\left\| \frac{L_t - L_{t-1}}{L_t} \right\| < \epsilon.$$
- **abs_tol** –
  Absolute tolerance ϵ_a. If the current bound L_t is close to zero, ($L_t < \epsilon_a$), declare convergence if
  $$\| L_t - L_{t-1} \| < \epsilon_a.$$
- **verbose** – Output status information after each update.

`set_variational_parameters()`

Reset the parameters to the submitted values or default.

Use this function to set the prior value (indicated by the subscript 0 as in α_0) or the initial value (e.g., α) used in the iterative procedure to find the posterior value of the variational distribution.

Every parameter can be set in two ways:

1. It is specified for only one component, then it is copied to all other components.
2. It is specified separately for each component as a K vector.

The prior and posterior variational distributions of $\boldsymbol{\mu}$ and $\boldsymbol{\Lambda}$ for each component are given by

$$q(\boldsymbol{\mu}, \boldsymbol{\Lambda}) = q(\boldsymbol{\mu}|\boldsymbol{\Lambda})q(\boldsymbol{\Lambda}) = \prod_{k=1}^{K} \mathcal{N}(\boldsymbol{\mu}_k|\boldsymbol{m}_k, (\beta_k \boldsymbol{\Lambda}_k)^{-1})\mathcal{W}(\boldsymbol{\Lambda}_k|\boldsymbol{W}_k, \nu_k),$$

where $\mathcal{N}$ denotes a Gaussian and $\mathcal{W}$ a Wishart distribution. The weights $\boldsymbol{\pi}$ follow a Dirichlet distribution

$$q(\boldsymbol{\pi}) = Dir(\boldsymbol{\pi}|\boldsymbol{\alpha}).$$

Warning

This function may delete results obtained by `update()`.

https://pypi.python.org/pypi/pypmc

Motivation

The standard model (SM) of particle physics cannot explain:

- dark matter
- neutrino masses
- hierarchy problem
- strong CP problem
- ...

Motivation

The standard model (SM) of particle physics cannot explain:

- dark matter
- neutrino masses
- hierarchy problem
- strong CP problem
- ...

new physics (NP) required

Motivation

The standard model (SM) of particle physics cannot explain:

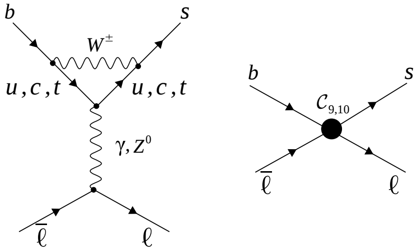
- dark matter
- neutrino masses
- hierarchy problem
- strong CP problem
- ...

new physics (NP) required

exact structure unknown $\Rightarrow$ model independent analysis effective Lagrangian for $b \to s\ell^+\ell^-$ (SM):

$$\mathcal{L}_{int} = \frac{4G_F}{\sqrt{2}} \frac{\alpha_e}{4\pi} V_{tb} V_{ts}^* \sum_i \mathcal{C}_i \mathcal{O}_i + \ldots + \text{h.c.}$$

$$\mathcal{O}_9 = \left[\bar{s}\gamma_\mu P_L \quad b\right]\left[\bar{\ell}\gamma^\mu \ell\right] \qquad \mathcal{O}_{10} = \left[\bar{s}\gamma_\mu P_L \quad b\right]\left[\bar{\ell}\gamma^\mu \gamma_5 \ell\right]$$



Effective theory

effective Lagrangian for $b \to s\ell^+\ell^-$ (beyond-SM):

$$\mathcal{L}_{int} = \frac{4G_F}{\sqrt{2}} \frac{\alpha_e}{4\pi} V_{tb} V_{ts}^* \sum_i \mathcal{C}_i \mathcal{O}_i + ... + \text{h.c.}$$

$$\mathcal{O}_9^{(\prime)} = \left[\bar{s}\gamma_\mu P_{L(R)} b\right]\left[\bar{\ell}\gamma^\mu \ell\right] \qquad \mathcal{O}_{10}^{(\prime)} = \left[\bar{s}\gamma_\mu P_{L(R)} b\right]\left[\bar{\ell}\gamma^\mu \gamma_5 \ell\right]$$

$$\mathcal{O}_S^{(\prime)} = \left[\bar{s} P_{R(L)} b\right]\left[\bar{\ell}\ell\right] \qquad \mathcal{O}_P^{(\prime)} = \left[\bar{s} P_{R(L)} b\right]\left[\bar{\ell}\gamma_5 \ell\right]$$

$$\mathcal{O}_T = \left[\bar{s}\sigma_{\mu\nu} b\right]\left[\bar{\ell}\sigma^{\mu\nu} \ell\right] \qquad \mathcal{O}_{T5} = \left[\bar{s}\sigma_{\mu\nu} b\right]\left[\bar{\ell}\sigma^{\mu\nu}\gamma_5 \ell\right]$$

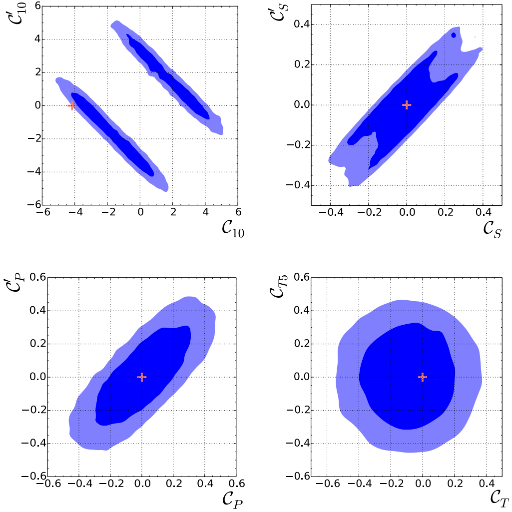
Experimental constraints

- $B \to K\mu^+\mu^-$: $\mathcal{B}, A_{FB}, F_H$
  - LHCb 2014 (arXiv:1403.8044 , arXiv:1403.8045)
  - CDF 2012

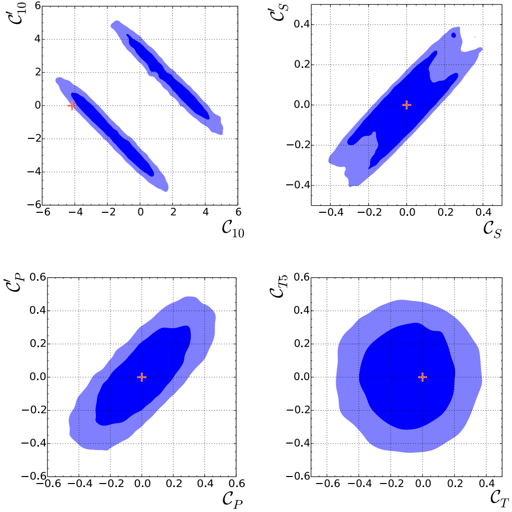
    (http://www-cdf.fnal.gov/physics/new/bottom/120628.blessed-b2smumu_96)
- $B_s \to \mu^+\mu^-$: $\mathcal{B}$
  - LHCb+CMS 2014 (arXiv:1411.4413)
- $B \to K^*\mu^+\mu^-$: $\mathcal{B}$
  - LHCb 2013 (arXiv:1304.6325)
  - CMS 2013 (arXiv:1308.3409)
  - CDF 2012

    (http://www-cdf.fnal.gov/physics/new/bottom/120628.blessed-b2smumu_96)

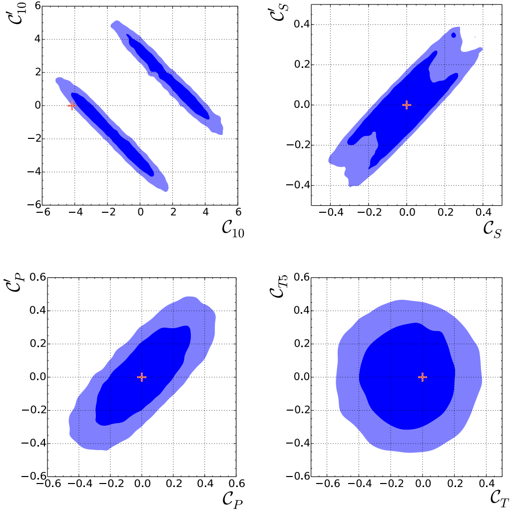
Joint fit of $\mathcal{C}_{10}^{(\prime)}, \mathcal{C}_S^{(\prime)}, \mathcal{C}_P^{(\prime)}, \mathcal{C}_T, \mathcal{C}_{T5}$, and 29 nuisance parameters



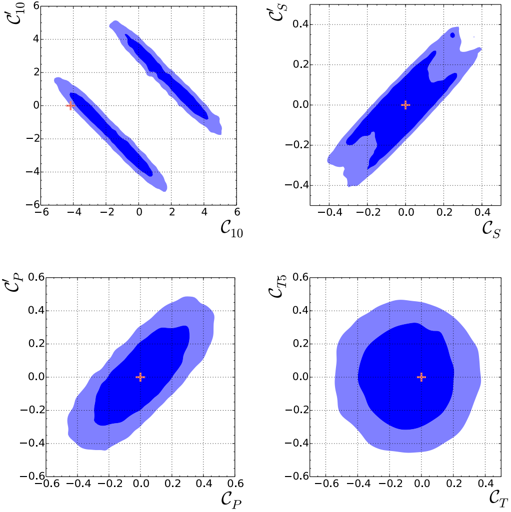
Joint fit of $\mathcal{C}_{10}^{(\prime)}, \mathcal{C}_S^{(\prime)}, \mathcal{C}_P^{(\prime)}, \mathcal{C}_T, \mathcal{C}_{T5}$, and 29 nuisance parameters



- first *simultaneous* fit



- first *simultaneous* fit
- interference $\mathcal{C}_{10}^{(\prime)} \leftrightarrow \mathcal{C}_{S,P}^{(\prime)}$ in $\mathcal{B}(B_s \to \mu^+\mu^-)$

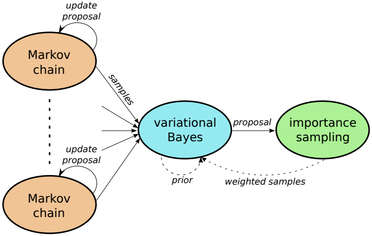


- first *simultaneous* fit
- interference $\mathcal{C}_{10}^{(\prime)} \leftrightarrow \mathcal{C}_{S,P}^{(\prime)}$ in $\mathcal{B}(B_s \to \mu^+\mu^-)$
  $\Rightarrow$ larger uncertainty than obtained for fixed $\mathcal{C}_{10}^{(\prime)} = \mathcal{C}_{10}^{(\prime)\mathrm{SM}}$
  arXiv:1205.5811,
  arXiv:1206.0273,
  arXiv:1407.7044

Summary

Summary

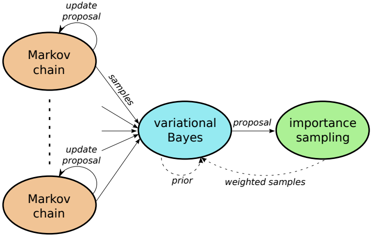
algorithm to sample and integrate in dim = $\mathcal{O}(40)$



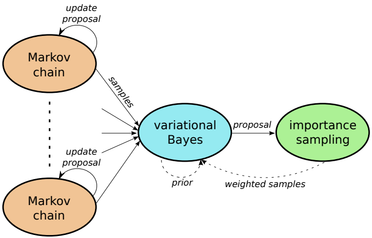
Summary

algorithm to sample and integrate in dim $= \mathcal{O}(40)$

model-independent search for new physics



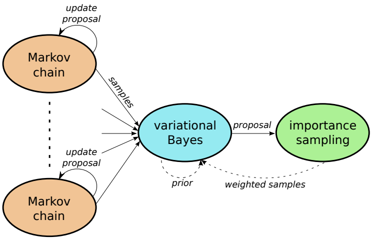
algorithm to sample and integrate in dim = $\mathcal{O}(40)$



model-independent search for new physics

- simultaneous fit of $\mathcal{C}_{10}^{(\prime)}, \mathcal{C}_S^{(\prime)}, \mathcal{C}_P^{(\prime)}, \mathcal{C}_T$, and $\mathcal{C}_{T5}$
  $\Rightarrow$ updated constraints

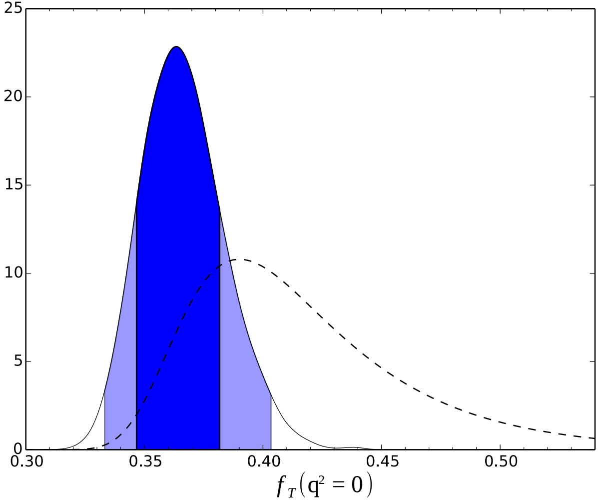
algorithm to sample and integrate in dim = $\mathcal{O}(40)$



model-independent search for new physics

- simultaneous fit of $\mathcal{C}_{10}^{(\prime)}, \mathcal{C}_S^{(\prime)}, \mathcal{C}_P^{(\prime)}, \mathcal{C}_T$, and $\mathcal{C}_{T5}$
  $\Rightarrow$ updated constraints
- no significant deviation from the SM

Nuisance parameters



Nuisance parameters