La Thuile, March 6 2024

Machine Learning Applied to $b \to s\ell^+\ell^-$

Presented by Jason Aebischer

University of Zurich

Outline

1. Motivation
2. $B \to K^* \mu^+ \mu^-$
3. Neural network
4. Summary

In collaboration with Jernej Kamenik and Sandro Mächler

Outline

Motivation

Angular observables in $B \to K^* \mu^+ \mu^-$

Tension: $\sim 2\sigma$

Long distance (LD) effects

Hard to compute

Solution

Neural Network

Outline

Effective Hamiltonian

$$\mathcal{H}_{\text{eff}} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i C_i O_i \,,$$

$$O_9 = \frac{e^2}{16\pi^2} (\bar{s}\gamma_\nu P_L b)(\bar{\mu}\gamma^\nu \mu)\,, \quad O_{10} = \frac{e^2}{16\pi^2} (\bar{s}\gamma_\nu P_L b)(\bar{\mu}\gamma^\nu \gamma_5 \mu)\,.$$

$\bar{B}^0 \to \bar{K}^{*0}(\to K^-\pi^+)\mu^+\mu^-$

Differential decay rate

$$\frac{d^4\Gamma}{dq^2\, d\cos\theta_l\, d\cos\theta_{K^*}\, d\phi} = \sum_i a_i J_i$$

Angular coefficients

$$J_i = J_i(C_i, FFs, LD, ..)$$

Issue

Dependence on LD effects

Optimized observables

Descotes-Genon/Hurth/Matias/Virto: 1303.5794

P_i

Linear combinations of J_i

Example

$P_{5'} = (J_5 + \bar{J}_5)/2\mathcal{N}'$

Advantage

small dependence on FFs

Goal

Find optimized observables

Small dependence on LD effects

Form

Lin. comb. of J's

Parameter inference

Find C_9, C_{10}

Neural network: Classifier

Input

q^2-bin, FFs, J_i

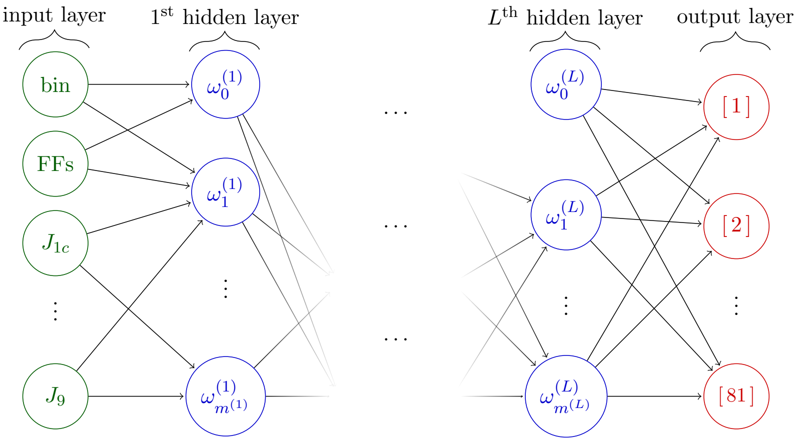
Output

$C_9 - C_{10}$ class

Class

range for C_9 and C_{10}

Neural Network



Outline

1. Motivation
2. $B \to K^* \mu^+ \mu^-$
3. Neural network
4. Summary

Procedure

Create training data

Vary J_i via parameters

Training

Labelled data

Comparison

with global fit

Tools

flavio

Straub: 1810.08132

Training data generation

pytorch

Paszke et.al: 1912.01703

training, validation

smelli

JA/Kumar/Stangl/Straub: 1810.07698

Comparison of performance

Training data

WCs

$C_9^{NP} \in [-C_9^{SM}, C_9^{SM}], \quad C_{10}^{NP} \in [-C_{10}^{SM}, C_{10}^{SM}]$

FFs

Bharucha/Straub/Zwicky: 1503.05534

SSE parameters

LD

$C_9 \to C_9 + \Delta C_9^{LD}$

LD

Helicity dependence

$(\Delta C_9^{LD})_\lambda = a_\lambda + i \cdot b_\lambda$

$\lambda = 0, +, -$

LD: here

charm loop contributions

Other approach

z-expansion

Gubernari/Reboud/van Dyk/Virto: 2206.03797

Data sets

q^2-bins [GeV^2]

low: [0.1, 0.98], [1.1, 2.5], [2.5, 4], [4, 6]

high: [15, 17], [17, 19]

FFs

small, moderate

LD

small, moderate, large

Neural Network

Hidden layers, nodes

8 , 200

Activation function

ReLU

Loss function

Cranmer/Pavez/Louppe: 1506.02169

Cross-entropy

Validation data

J_i

normally distributed

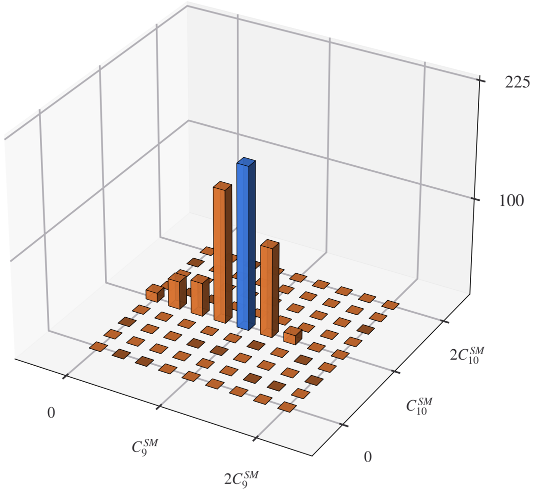
FFs

small

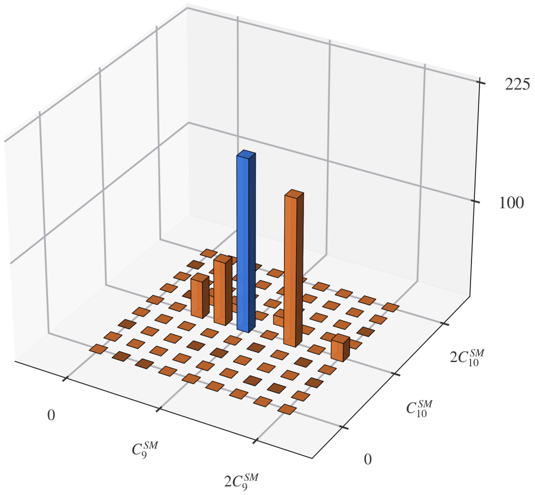
LD

small, moderate, large

Result: Small LD



Result: large LD



Outline

Summary

Angular observables for $B \to K^* \mu^+ \mu^-$

Depend on LD contributions

ML techniques

NN that minimizes dependence on LD

Parameter inference

comparable results