Boosted Ensembles of Qubit and Continuous Variable Quantum Support Vector Machines for B Meson Flavour Tagging

Maxwell West[1], Martin Sevior[1] and Muhammad Usman[1,2]

[1]School of Physics, The University of Melbourne, Victoria, Australia

[2]Data61, CSIRO, Victoria, Australia

May 8, 2023

CP violation and the Standard Model

- CP symmetry violation is necessary to explain the preponderance of matter over anti-matter, however its known sources within the standard model (SM) are insufficient to explain the magnitude of the observed asymmetry.
- While there are not many known examples of CP violation, it can be seen in certain decays of B mesons.
- Belle-II and LHCb experimentally test this CP violation, and whether the SM describes it completely, or if New Physics is required...

CP violation in the $B^0 - \overline{B}^0$ system

- CP symmetry implies equality of the decay rates Γ of B^0 and $\overline{B}^0$ to a common CP eigenstate, however the SM predicts

$$\frac{\Gamma(\overline{B}^0 \to J/\psi K_S^0) - \Gamma(B^0 \to J/\psi K_S^0)}{\Gamma(\overline{B}^0 \to J/\psi K_S^0) + \Gamma(B^0 \to J/\psi K_S^0)} \sim \sin \delta m t \sin 2\beta$$

  where β is a phase from the CKM matrix and δm is the mass difference between the two mass eigenstates.

- Need to be able to reliably distinguish B^0 from $\overline{B}^0$ in order to do this analysis.

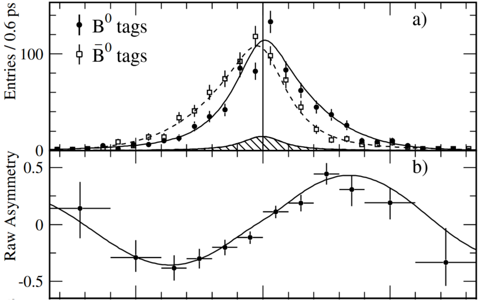


Figure taken from Ref.[1]

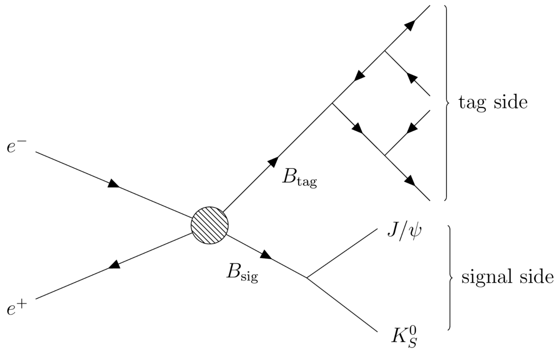
[1] B. Aubert *et al.* *Phys. Rev. Lett.* **89** 201802 (2002)

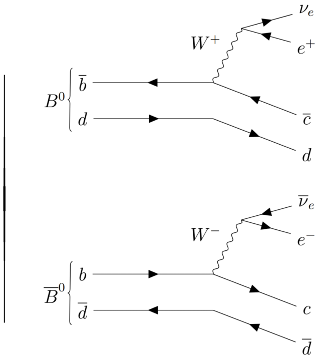
B^0 flavour tagging

- At Belle-II, $B^0 - \overline{B}^0$ pairs are created in entangled states by $e^- e^+$ collisions:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(\left| B^0 \overline{B}^0 \right\rangle - \left| \overline{B}^0 B^0 \right\rangle \right)$$

- *Flavour tagging* is the process of determining the quark content of the "tag-side" meson B_{tag} (i.e. $\bar{b}d$ or $b\bar{d}$)





B^0 flavour tagging at Belle-II

- Currently the state-of-the-art results at Belle-II are achieved via machine learning approaches on 130 input variables.
- The performance of the classifiers are characterised by the effective tagging efficiency Q, defined as

$$Q = \sum_{i=1}^{n_{\text{bins}}} \epsilon_i (1 - 2w_i)^2$$

  where ϵ_i is the fraction of events in the ith bin, and w_i is the fraction incorrectly tagged.
- Recent results[2] using fast boosted decision trees and deep neural networks give

$$Q_{\text{FBDT}} = 30.0\%$$

$$Q_{\text{DNN}} = 28.8\%$$

[2] Abudinen, F., et al. B-flavour tagging at Belle II. *The European Physical Journal C*, 82(4). (2022)

Qubit and Continuous Variable Quantum Computers

- The fundamental unit of a conventional quantum computer is the qubit, a state of a two level quantum system:

$$|\psi\rangle_{\text{qubit}} = \alpha |0\rangle + \beta |1\rangle$$

- We will also consider continuous variable (CV) quantum computers, the fundamental units of which are *qumodes*, quantum systems with continuous degrees of freedom:

$$|\psi\rangle_{\text{qumode}} = \int dx \; \psi(x) |x\rangle$$

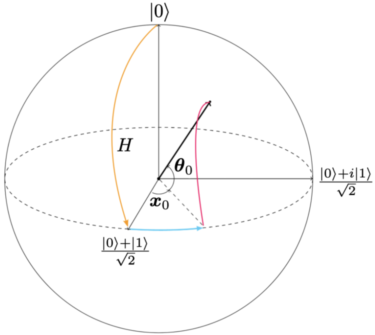
e.g. a set of bosonic modes (i.e. harmonic oscillators).

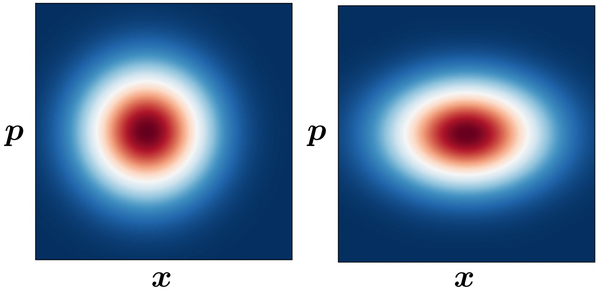
Visualising Quantum Operations

- Each qumode has a representation in terms of a pair $\hat{x}$, $\hat{p}$ of canonically conjugate operators formed from the creation and annihilation operators of the mode:

$$\hat{x} = \hat{a} + \hat{a}^\dagger, \quad \hat{p} = i\left(\hat{a}^\dagger - \hat{a}\right)$$

- Each qubit has a representation as a point on the surface of the *Bloch sphere*.





Quantum Support Vector Machines

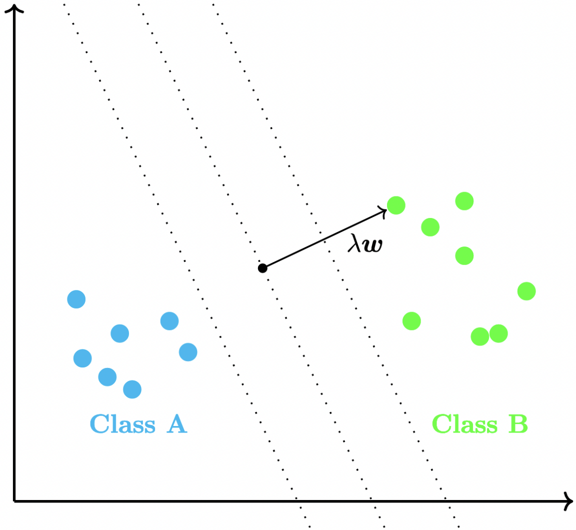
- SVMs are linear classifiers on data which has typically been mapped to a feature space,

$$\boldsymbol{x} \mapsto \Phi(\boldsymbol{x})$$

- In a QSVM this mapping is into a quantum Hilbert space,

$$\boldsymbol{x} \mapsto |\psi(\boldsymbol{x})\rangle = \mathcal{U}(\boldsymbol{x})\,|0\rangle$$

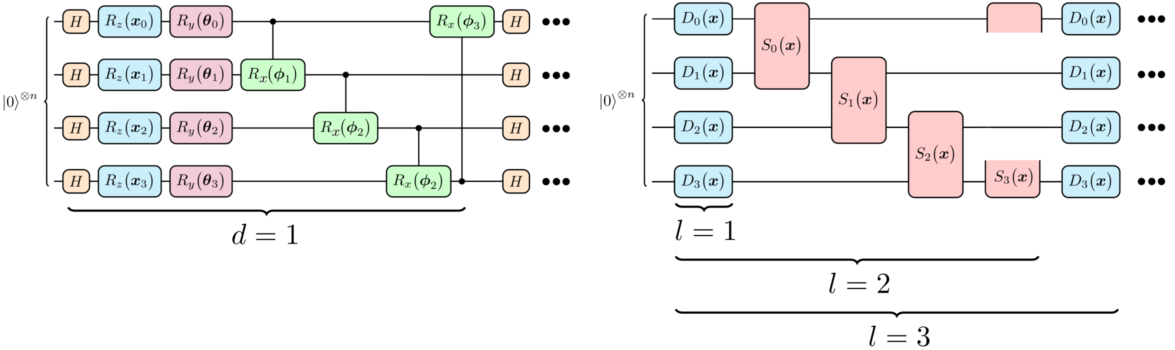
- QSVMs can be implemented on near-term (non error corrected) quantum computers.



Quantum Support Vector Machine Architectures

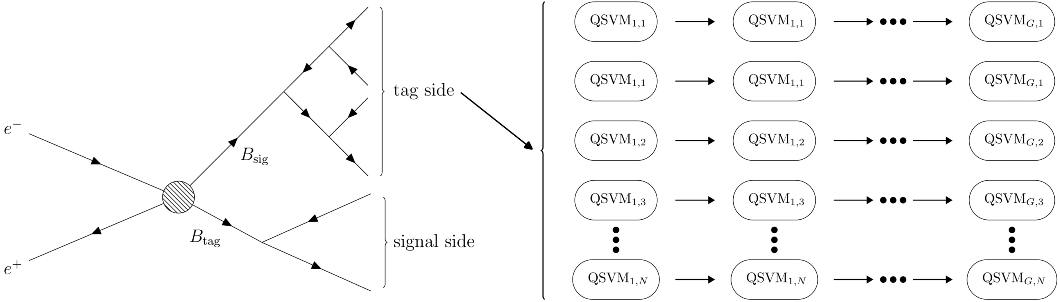
- The key component of a QSVM is the data encoding map $\boldsymbol{x} \mapsto |\psi(\boldsymbol{x})\rangle$.
- We consider mappings implementable on quantum computers consisting of
  - Qubit rotations
  - Qumode displacement and squeezing operations

on conventional and CV quantum computers respectively.



Boosted Ensembles of QSVMs

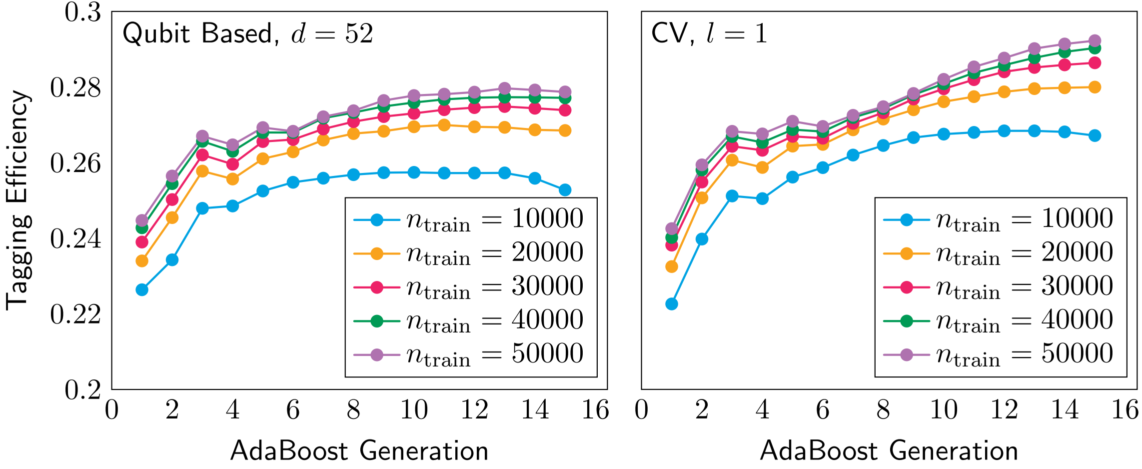
- We consider ensembles of $N = 200$ QSVMs. Each QSVM outputs a qr value, with the final prediction taken to be the average.
- Furthermore, each QSVM is boosted with the AdaBoost algorithm[3].



[3] Freund, Y. and Schapire, R.E. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1). (1997)

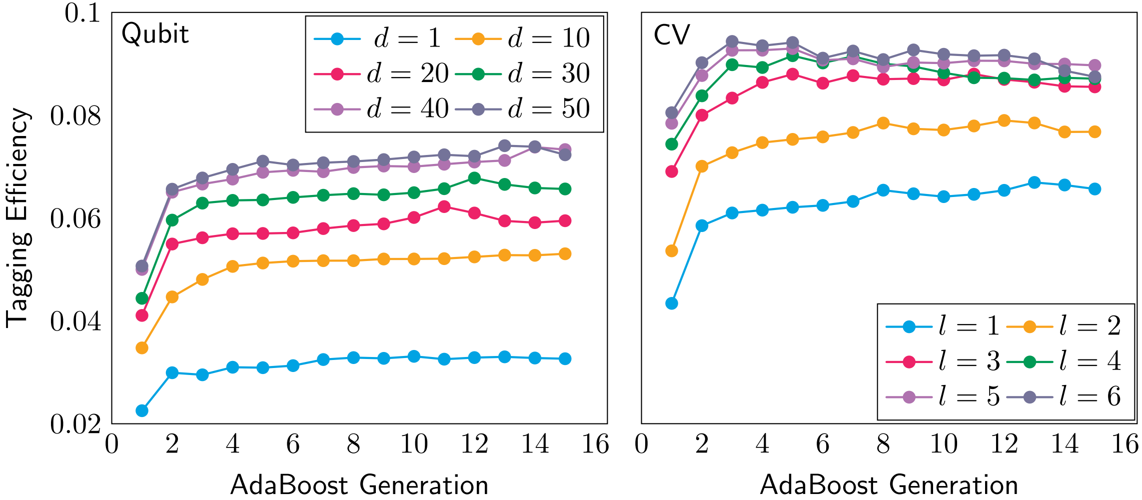
Quantum Support Vector Machines: Results

- We construct boosted ensembles of 200 QSVMs.
- Due to simulation constraints, the CV models only include the inital displacement.



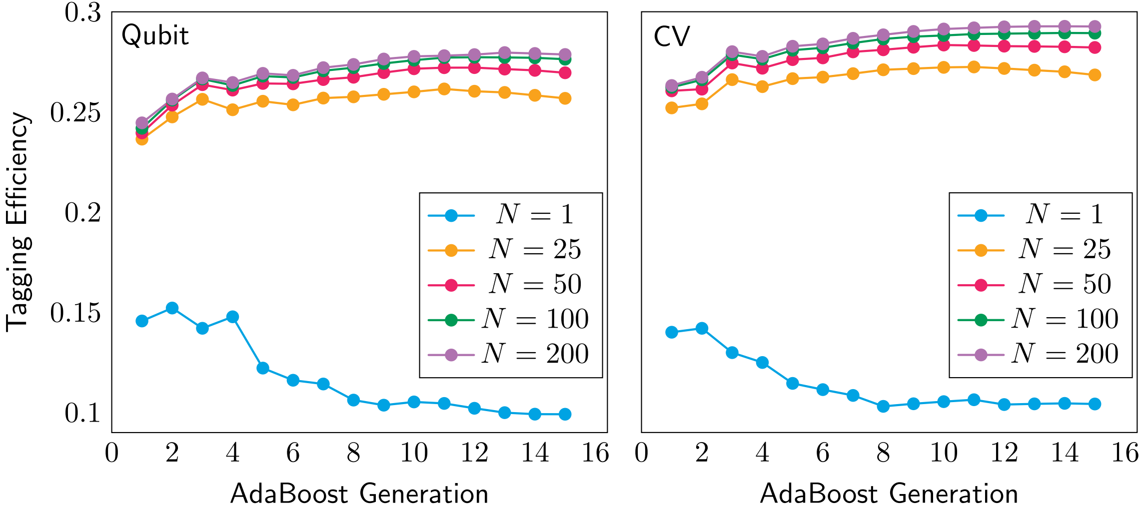
Top 5 PCA Components

- By doing a PCA transformation we can reduce the dimensionality of the data and simulate more powerful CV-QSVMs.



Effect of Ensemble Size

- We find drastic benefits in moving from single QSVMs to ensembles with $N \gtrsim 100$



Summary

- B meson flavour tagging is an important component of experiments which probe CP violation and heavy quark mixing.
- By using boosted ensembles of QSVMs we can achieve flavour tagging at level commensurate with state of the art classical algorithms.
- This is achieved despite imposing heavy restrictions on the CV-QSVMs in order to make them classically simulable.
- There is a tantalising prospect for outperforming classical methods as quantum computer hardware matures and it becomes possible to perform large-scale entangled kernels[4].

[4] West, M., Sevior, M. and Usman, M. Boosted Ensembles of Qubit and Continuous Variable Quantum Support Vector Machines for B Meson Flavour Tagging. arXiv:2305.02729 (2023)