Importance of Kernel Bandwidth in Quantum Machine Learning

Ruslan Shaydulin* Stefan M. Wild*

*Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, IL 60439, USA

Supplementary Information

1 Additional numerical experiments

Figure S1 illustrates the relationship between model performance and scaling factor for varying levels of the second hyperparameter, namely, the number of Trotter steps T. No significant variation in performance with T is observed. Figure S2 shows the scaling of the kernel for feature maps and datasets not included in the main text.

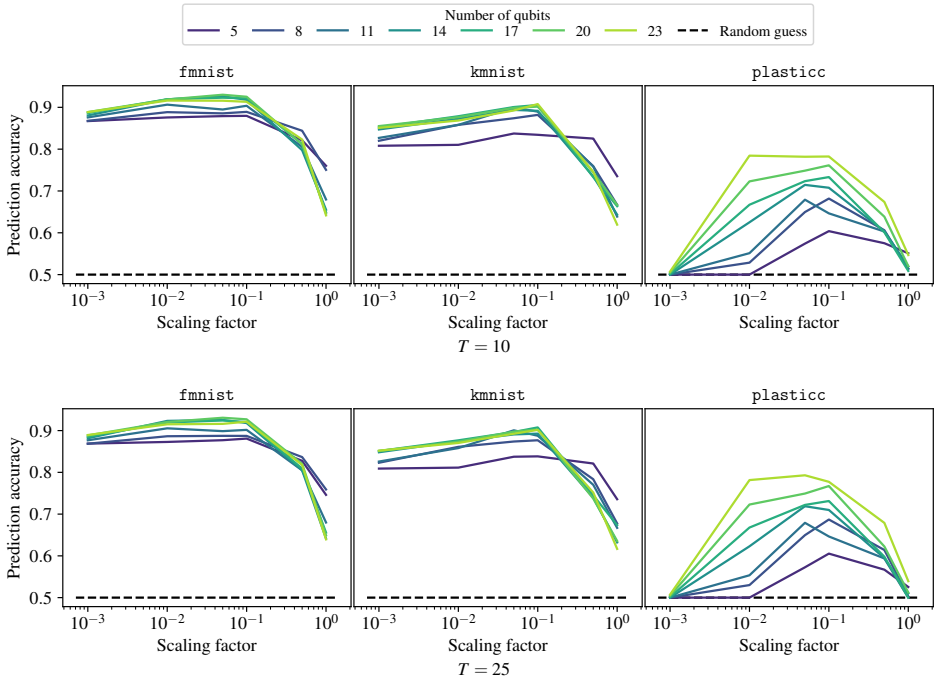
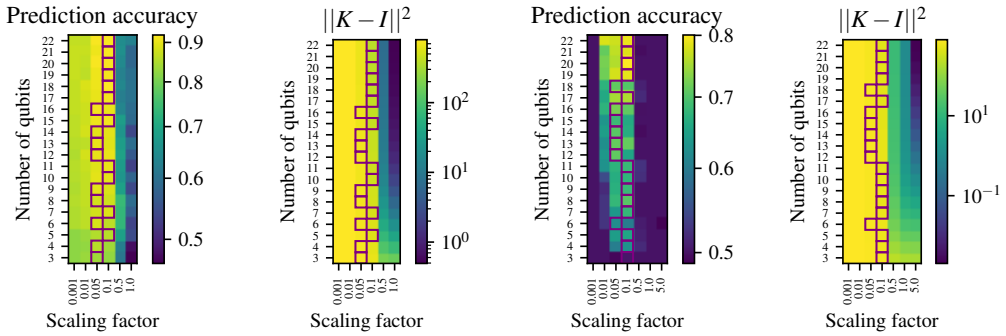


Figure S1: Prediction accuracy of the quantum kernel method with the Hamiltonian evolution feature map as a function of the scaling factor for the number of Trotter steps $T = 25$ and $T = 40$.

Figures S3 and S4 show results without cross-validation for SVM hyperparameter optimization. Concretely, the penalty term C is optimized to maximize the training score. Figure S3 shows the prediction accuracy change with scaling factor, and Figure S4 shows scaling of prediction accuracy with qubit count for optimized bandwidth.

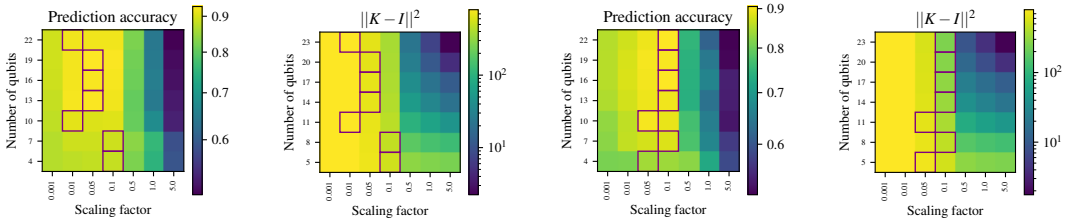


(a) Prediction accuracy for the IQP feature map applied to the `kmnist` dataset

(b) Kernel matrix entries for the IQP feature map applied to the `kmnist` dataset

(c) Prediction accuracy for the IQP feature map applied to the `kmnist` dataset

(d) Kernel matrix entries for the IQP feature map applied to the `plasticc` dataset

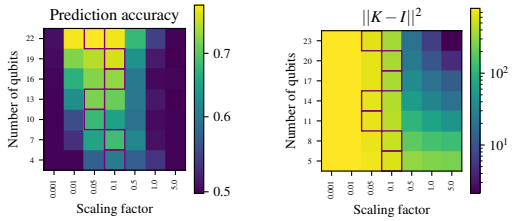


(e) Prediction accuracy for the Hamiltonian evolution feature map applied to the `fmnist` dataset

(f) Kernel matrix entries for the Hamiltonian evolution feature map applied to the `fmnist` dataset

(g) Prediction accuracy for the Hamiltonian evolution feature map applied to the `kmnist` dataset

(h) Kernel matrix entries for the Hamiltonian evolution feature map applied to the `kmnist` dataset



(i) Prediction accuracy for the Hamiltonian evolution feature map applied to the `plasticc` dataset

(j) Kernel matrix entries for the Hamiltonian evolution feature map applied to the `plasticc` dataset

Figure S2: Additional numerical evidence of kernel matrix entry scaling and model performance variation with the scaling factor. For each qubit count, the value of the scaling factor corresponding to the highest accuracy is highlighted with a purple square. For each model and each dataset considered, a clear "Goldilocks" regime is visible where the scaling factor is neither too small nor too large.

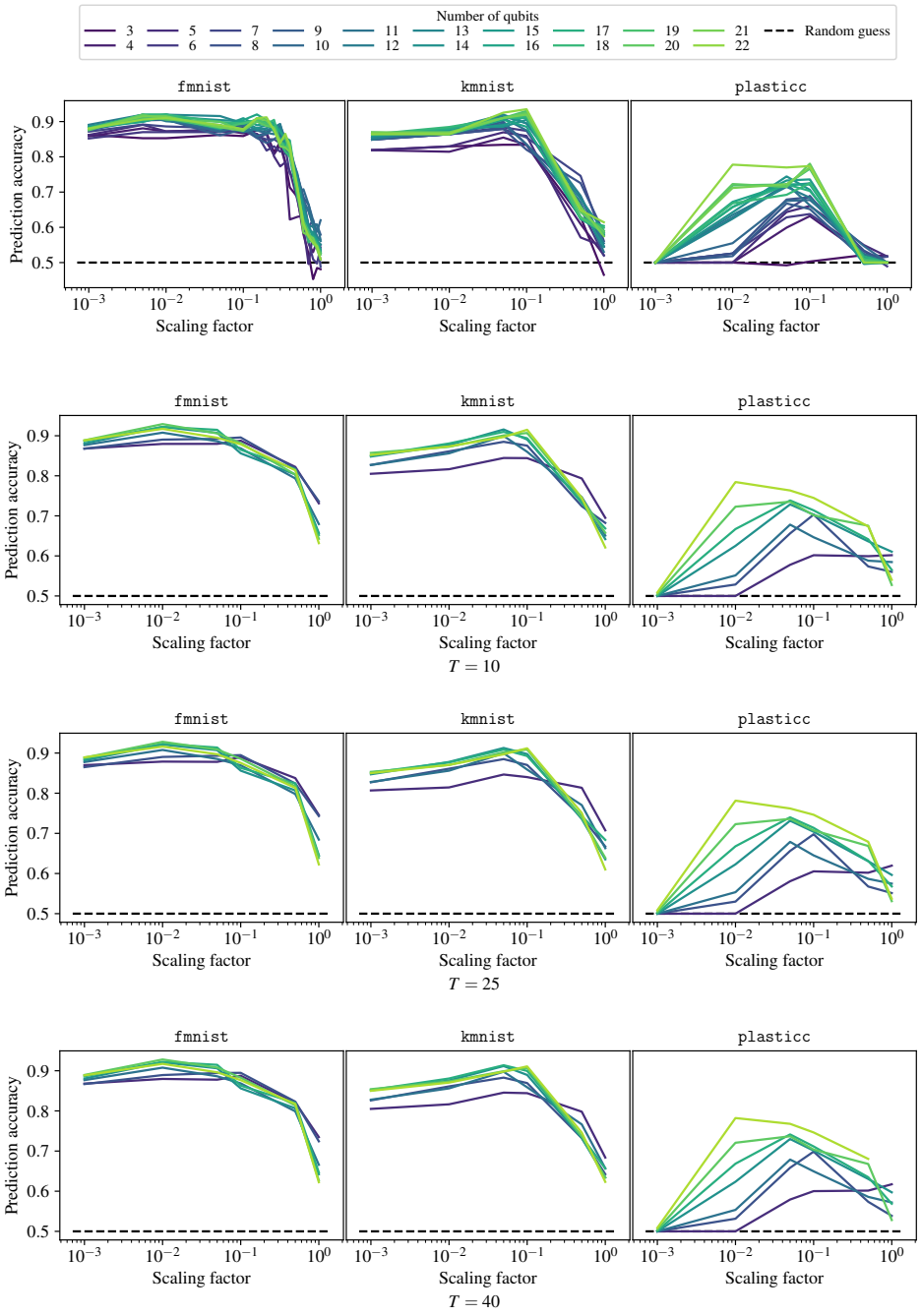


Figure S3: Prediction accuracy as a function of the scaling factor without cross-validation for SVM hyper-parameter optimization.

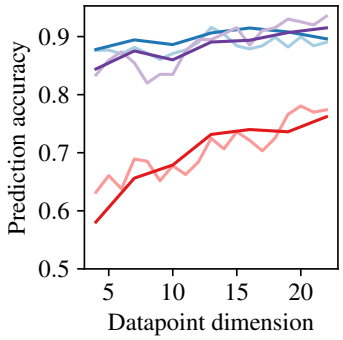


Figure S4: Prediction accuracy as a function of the datapoint dimension. Even without cross-validation, performance improves with qubit count.