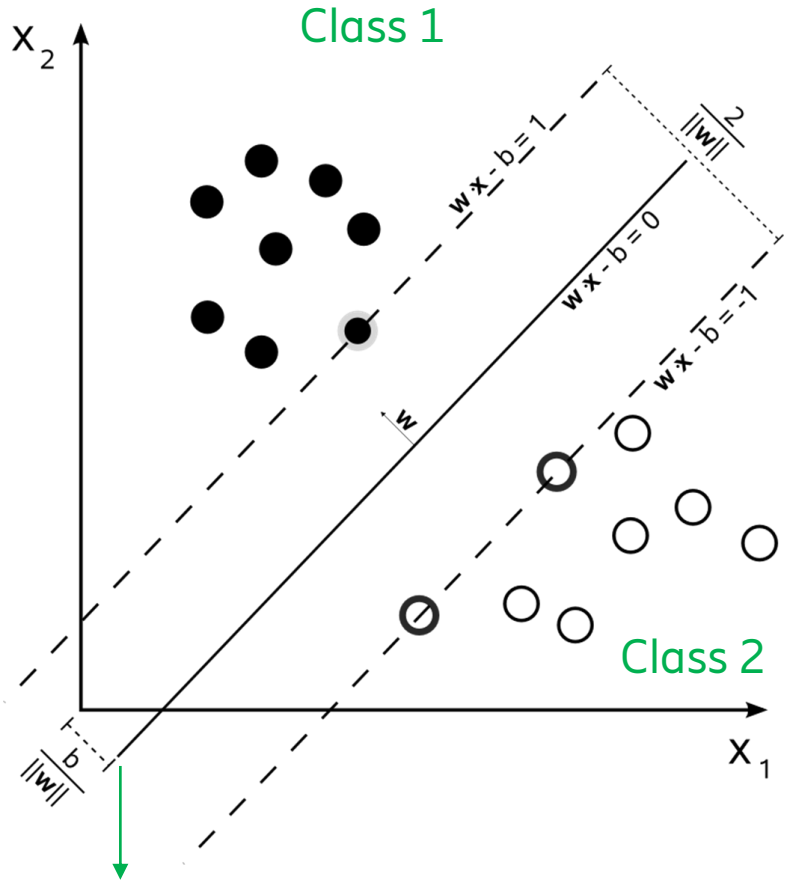
Quantum Support Vector Machine

Claire Yang 2019-03-14

Classical SVM



Hyperplane
(dimention: N-1)

M training data points: $\{(\vec{x_j}, y_j) : \vec{x_j} \in \mathbb{R}^N, y_j = \pm 1\}, j = 1 \dots M$

Goal → find the maximum-margin hyperplane → maximize $\frac{2}{|\vec{\omega}|}$ → minimize $\frac{|\vec{\omega}|^2}{2}$

The constraint:

$$\begin{cases} \vec{w} \cdot \vec{x_j} + b \geq 1 & \text{if } y_j = +1 \ (y_j \ belongs \ to \ class \ 1) \\ \vec{w} \cdot \vec{x_j} + b \leq -1 & \text{if } y_j = -1 \ (y_j \ belongs \ to \ class \ 2) \end{cases} \Leftrightarrow y_i(\vec{w} \cdot \vec{x_j} + b) \geq 1$$

Computational Complexity: $O(log(\epsilon^{-1})poly(N, M))$

Dimension of feature space (input data)

Number of training data points

Image: https://zh.wikipedia.org/wiki/File:Svm_max_sep_hyperplane_with_margin.png

SVM structure

Goal → find the maximum-margin hyperplane → maximize $\frac{2}{|\vec{\omega}|}$ s.t. $y_j(\vec{w}\cdot\vec{x_j}+b)\geq 1$ → minimize $\frac{|\vec{\omega}|^2}{2}$ s.t. $y_j(\vec{w}\cdot\vec{x_j}+b)\geq 1$ → Fulfils KKT condition

$$\max_{\vec{\alpha}}\min_{\vec{\omega},b} L(\vec{\omega},b,\vec{\alpha})=\frac{|\vec{\omega}|^2}{2}-\sum_{j=1}^{M}\alpha_j y_j(y_j(\vec{\omega}\cdot\vec{x_j}+b)-1), \text{s.t.} \sum_{j=1}^{M}\alpha_j=0 \text{ and } \alpha_j y_j\geq 0$$

Partial derivative:

$$\frac{\partial L}{\partial \vec{\omega}}=0\Rightarrow\vec{\omega}=\sum_{j=1}^{M}\alpha_j y_j\vec{x_j}$$

$$\frac{\partial L}{\partial b}=0\Rightarrow\sum_{j=1}^{M}\alpha_j y_j^2=0$$

Kernel, Inner product ←Problem 1

$$\max_{\vec{\alpha}} L(\vec{\alpha})=\sum_{j=1}^{M}\alpha_j y_j-\frac{1}{2}\sum_{j,k=1}^{M}\alpha_j\vec{x_j}\cdot\vec{x_k}\alpha_k$$

Find optimal $\vec{\alpha}^*$ ←Problem 2

Optimal hyperplane with parameters:

$$\vec{\omega}=\sum_{j,k=1}^{M}\alpha_j^*\vec{x_j}$$

$$b=y_i-\vec{\omega}$$

Quantum SVM

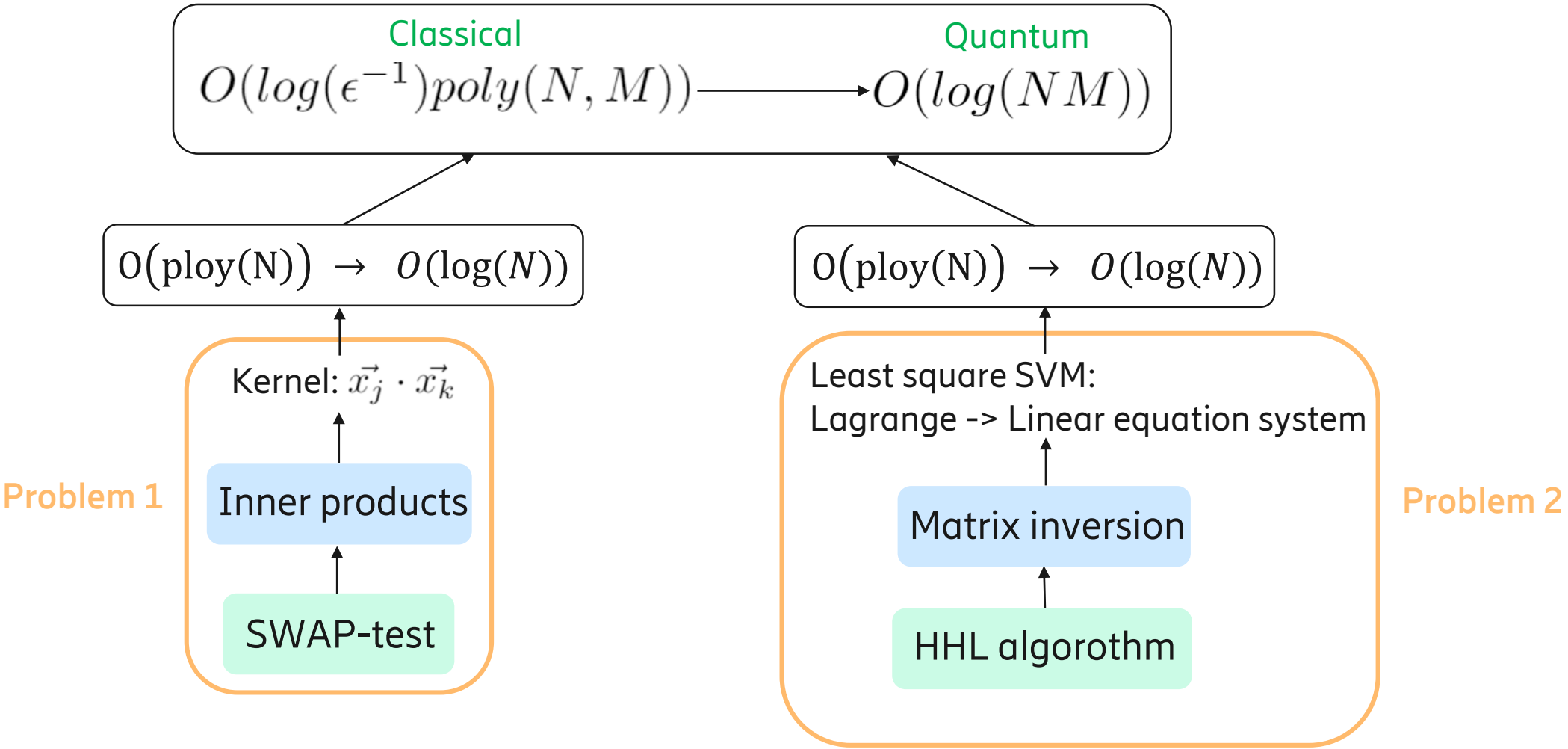
Computational Complexity:

Classical $O(log(\epsilon^{-1})poly(N, M))$ → Quantum $O(log(NM))$

training

testing

HHL based qSVM



Problem 1 – Inner product & SWAP-test

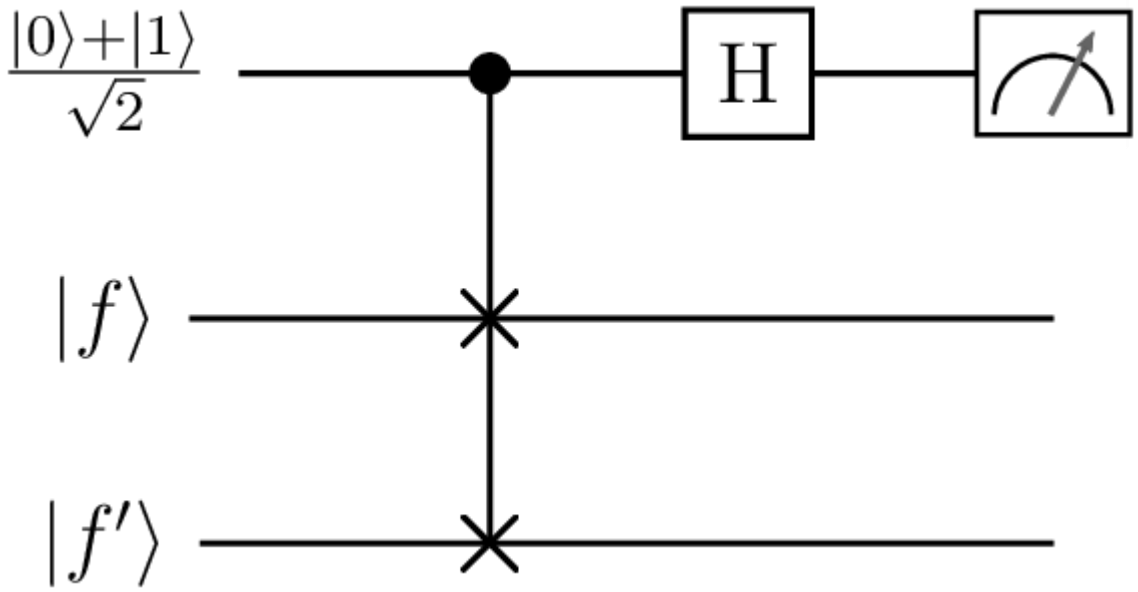


Image: https://peterwittek.com/understanding-quantum-svms.html

Problem 2 – Least Square SVM & HHL

Slack variable

Constraint: $y_j(\vec{\omega}\cdot\vec{x_j}+b)\geqslant 1 \overset{y_j^2=1}{\rightarrow} \vec{\omega}\cdot\vec{x_j}+b=y_j-y_j e_j$

New Lagrange function: $L(\vec{\omega},b,\vec{e},\vec{\alpha})=\frac{|\vec{\omega}|^2}{2}+\frac{\gamma}{2}\sum_{j=1}^{M}e_j^2-\sum_{j=1}^{M}\alpha_j y_j(\vec{\omega}\cdot\vec{x_j}+b-y_j+y_j e_j)$

Penalty term

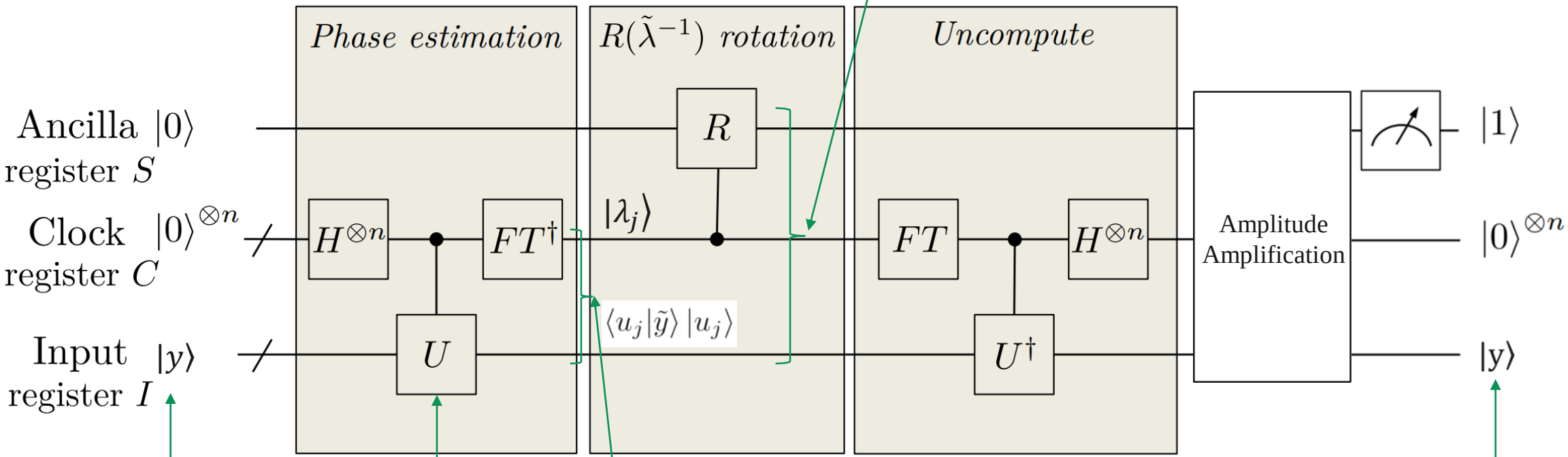
Linear equation system: $F\begin{pmatrix}b\\ \vec{\alpha}\end{pmatrix}\equiv\begin{pmatrix}0 & -\vec{1}^T\\ \vec{1} & K+\gamma^{-1}I\end{pmatrix}\begin{pmatrix}b\\ \vec{\alpha}\end{pmatrix}=\begin{pmatrix}0\\ \vec{y}\end{pmatrix}$

HHL

Problem 2 – Least Square SVM && HHL

Solve $F\begin{pmatrix} b \\ \vec{\alpha} \end{pmatrix} = \begin{pmatrix} 0 \\ \vec{y} \end{pmatrix}$

$\sum_{j=1}^{M+1}\left(\sqrt{1-\frac{C^2}{\lambda_j^2}}|0\rangle + \frac{C}{\lambda_j}|1\rangle\right)\langle u_j|\tilde{y}\rangle\,|u_j\rangle|\lambda_j\rangle$



$|\tilde{y}\rangle = \sum_{j=1}^{M+1}\langle u_j|\tilde{y}\rangle\,|u_j\rangle$

$F \rightarrow \mathrm{U} = e^{i\hat{F}t}$

$\sum_{j=1}^{M+1}\langle u_j|\tilde{y}\rangle\,|u_j\rangle|\lambda_j\rangle$

$\sum_{j=1}^{M+1}\frac{\langle u_j|\tilde{y}\rangle}{\lambda_j}|u_j\rangle$

Image: https://arxiv.org/pdf/1802.08227.pdf

References

— Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. Physical review letters, 113(13):130503, 2014.

— Wechat article (link)