# On Quantum Perceptron Learning via Quantum Search

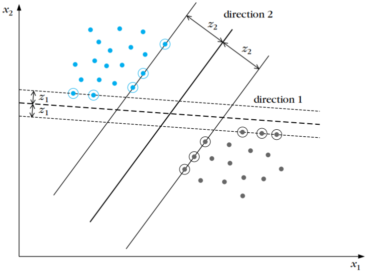
**Xiaoyu Sun**, Mathieu Roget, Giuseppe Di Molfetta, and Hachem Kadri
LIS, CNRS, Aix-Marseille Université, 13288, Marseille, France.

CANA, LIS, AMU

# Machine Learning Task: Binary Linear Classification

## Training dataset

- Training dataset $\mathcal{S}$: $N$ pairs of $\{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_N, y_N)\}$, where training examples $\boldsymbol{x}_{i\in[N]} \in \mathbb{R}^D$, labels $y_{i\in[N]} \in \{+1, -1\}$.




- **Perceptron learning** aims to find a hyperplane $\boldsymbol{w} \in \mathbb{R}^D$, such that $\{\boldsymbol{w} \mid \boldsymbol{w}^T\boldsymbol{x}_i y_i > 0, \forall i \in [N]\}$, and this **feasible set** is called **version space**.

# Dataset Assumption

## Geometric margin

The **geometric margin** $z_h := \min_{i \in [N]} \frac{|\boldsymbol{w}^T \boldsymbol{x}_i + b|}{\|\boldsymbol{w}\|_2}$ of a hyperplane $h = (\boldsymbol{w}, b)$.

Assumption: The maximal margin $\gamma = \max_h z_h \iff \exists\ \boldsymbol{u}^T \boldsymbol{x}_i y_i \geq \gamma$ for all $i \in [N]$.

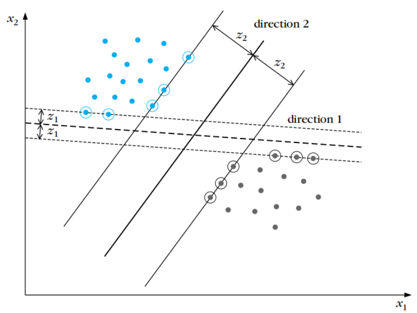



Figure 1: An example of two possible linear classifiers on a linearly separable dataset. Picture from [Theodoridis, S., & Koutroumbas, K. (2006)].

# Query Models

## Classical

A classical *data query* on a training data $(\boldsymbol{x}_i', y_i')$ refers to an evaluation of a Boolean function $f_{\boldsymbol{w}} = \{0, 1\}$, where

$$f_{\boldsymbol{w}}(\boldsymbol{x}_i' y_i') = 1 \iff \boldsymbol{w}^T \boldsymbol{x}_i' y_i' \leq 0 \text{ (misclassified)}. \tag{1}$$

## Quantum

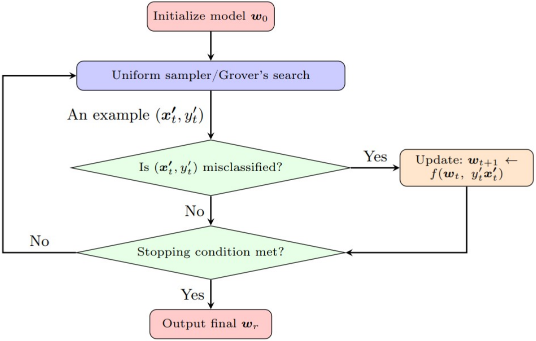
A quantum *data query* on a training data $|\boldsymbol{x}_i' y_i'\rangle$ refers to an evaluation of a quantum oracle $F_{\boldsymbol{w}}$, acting coherently on a superposition of data vectors,

$$F_{\boldsymbol{w}} |\boldsymbol{x}_i' y_i'\rangle = (-1)^{f_{\boldsymbol{w}}(\boldsymbol{x}_i' y_i')} |\boldsymbol{x}_i' y_i'\rangle , \tag{2}$$

or equivalently,

$$F |\boldsymbol{w}\rangle |\boldsymbol{x}_i' y_i'\rangle |0\rangle = |\boldsymbol{w}\rangle |\boldsymbol{x}_i' y_i'\rangle |0 \oplus f_{\boldsymbol{w}}(\boldsymbol{x}_i' y_i')\rangle . \tag{3}$$

- Online perceptron algorithm [Novikoff, 1962]
  Upon misclassification: $\boldsymbol{w} \leftarrow \boldsymbol{w} + \mathbf{x}'_i y'_i$. $\Rightarrow O^*(\frac{N}{\gamma^2})$ [1] classical *data queries*.
- Online quantum perceptron (OQP) [Kapoor et al., 2016]
  Using *Grover's search* [Grover, 1996] for speeding up searching for misclassified examples.
  $\Rightarrow O^*(\frac{\sqrt{N}}{\gamma^2})$ quantum *data queries*.




[1] The asterisk notation $O^*(\cdot)$ to suppress the terms with the error parameter, $\log(D)$, $\log\log(1/\gamma)$ and $\log\log(N)$.

# Previous Work - 2

- Quantum version space perceptron (QVSP) [Kapoor et al., 2016]
  An query to the hyperplane $\boldsymbol{w}_j$ :

$$G_{\mathcal{S}} \left| \boldsymbol{w}_j \right\rangle = (-1)^{1+(f_{\boldsymbol{w}_j}(\boldsymbol{x}_1, y_1) \vee f_{\boldsymbol{w}_j}(\boldsymbol{x}_2, y_2) \vee \cdots \vee f_{\boldsymbol{w}_j}(\boldsymbol{x}_N, y_N))} \left| \boldsymbol{w}_j \right\rangle , \tag{4}$$

  can be built using $O(N)$ quantum *data queries* as Eqn. (2). $\Rightarrow O^*(\frac{N}{\sqrt{\gamma}})$ quantum queries.
- Improvement [Liao et al., 2024]: quantum counting $\Rightarrow O^*(\sqrt{\frac{N}{\gamma}})$ quantum *data queries*.

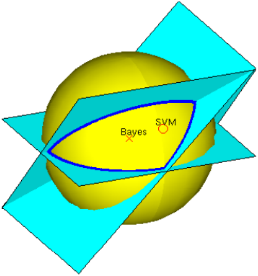



Figure 2: An illustration of version space. Picture from [Minka, T. P. (2001)].

# Result 1 - Classical

## Curse of Dimensionality in QVSP

We examine the QVSP algorithm proposed in [Kapoor et al., 2016], and show that instead of $\Theta(\gamma)$, sampling $\Pr[\boldsymbol{w} \mid \boldsymbol{w} \in \mathcal{VS}] = \Omega(\gamma^D)$ when $\gamma \to 0$.

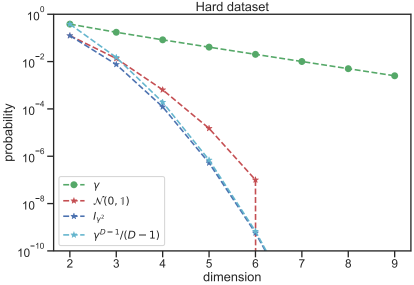



Figure 3: The perfect classification probability-dimension relationship (red) on an artificially designed dataset (Given from [Mohri, M. et al. (2018)].) with its margin $\gamma$ (blue) upper bounded by $\sqrt{1/2^{D-1}}$ (cyan).

- Applying *Grover's search* directly for the classifiers, in the worst case, QVSP uses $O^*(N/\sqrt{\gamma^D})$ instead of $O^*(N/\sqrt{\gamma})$ quantum *data queries*.

# Linear Programming Approaches

Feasibility algorithms: ellipsoid method [Khachiyan, 1979](first polynomial-time solvability), various cutting-plane methods [Vaidya, 1989, Bertsimas and Vempala, 2004].

## Cutting-Plane Random Walk (CP-RW)

Preserving an approximate centroid $\boldsymbol{z}_t$ of the current space $\mathcal{P}_t$.
The half space $H_{t+1} = \{\boldsymbol{w} \mid (\boldsymbol{w} - \boldsymbol{z}_t)^T \boldsymbol{x}'_{t+1} y'_{t+1} > 0\}$ cut by a misclassified example $(\boldsymbol{x}'_{t+1}, y'_{t+1})$.
Update: $\boldsymbol{z}_{t+1} \leftarrow \frac{1}{M} \sum_{j=1}^{M} \boldsymbol{w}'_j$, where $\boldsymbol{w}'_j \sim \mathcal{U}_{H_{t+1}}$. $(\mathcal{P}_{t+1} = H_{t+1} \cap \mathcal{P}_t = H_{t+1} \cap H_t \cap \cdots \cap \mathcal{P}_0)$

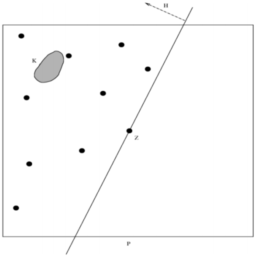



Figure 4: An illustration of the algorithm from [Bertsimas and Vempala, 2004].

- Optimal complexity of $O^*(D \log(1/\gamma))$ rounds (separation oracles).

## *Hit-and-Run* [Smith, 1984]

(1) Choose a line $\ell$ through the current point $\boldsymbol{x}$ uniformly at random.
(2) Move to a point $\boldsymbol{y}$ chosen **uniformly** from $\mathcal{P}_t \cap \ell$.

- The best-known mixing bound: $O^*(D^3)$ steps [Lovász, 1999].

? Quantum speed-ups

## Membership oracles

- Classically, with $t$ queries, $\mathcal{O}_{\mathcal{P}_t}(\boldsymbol{w}) = 0$ iff $\boldsymbol{w} \in \mathcal{P}_t$. where $f_{\boldsymbol{w}}(\boldsymbol{x}_j' y_j', \boldsymbol{z}_{j-1}) = 0$ iff $(\boldsymbol{w} - \boldsymbol{z}_{j-1})^T \boldsymbol{x}_j' y_j' > 0$, $\forall j \in [t]$.
- Quantumly, $\mathcal{Q}_{\mathcal{P}_t}$ can be implemented using $O(\sqrt{t})$ quantum *data queries* [Liao et al., 2024], with

$$P_t \left|\boldsymbol{w}\right\rangle \left|\boldsymbol{x}_j' y_j', \boldsymbol{z}_{j-1}\right\rangle \left|0\right\rangle = \left|\boldsymbol{w}\right\rangle \left|\boldsymbol{x}_j' y_j', \boldsymbol{z}_{j-1}\right\rangle \left|0 \oplus f_{\boldsymbol{w}}(\boldsymbol{x}_j' y_j', \boldsymbol{z}_{j-1})\right\rangle, \forall j \in [t+1]. \tag{5}$$

■ $t \leq O^*(D \log(1/\gamma)) \Rightarrow$ Naive Hybrid cutting-plane random walk (HCP-RW) algorithm.

? Further quantum speed-ups- sampling? [Wocjan and Abeyesinghe, 2008]

## Unique eigenvector

It was shown by [Szegedy, 2004, Magniez et al., 2007] that for an *irreducible* and time-reversible Markov chain $P$ with stationary distribution $\boldsymbol{\pi}$. The quantum walk $W(P) := R_{\mathcal{B}} R_{\mathcal{A}}$, where $R_{\mathcal{I}} = 2\,\Pi_{\mathcal{I}} - \mathbb{1}$, and $\mathcal{A} = \textbf{span}(|x\rangle\,|0\rangle : x \in \Omega)$, $\mathcal{B} = \textbf{span}(|0\rangle\,|x\rangle : x \in \Omega)$, has a unique eigenvalue-1 eigenvector

$$|\pi\rangle = |\pi\rangle'\,|0\rangle = |\pi\rangle' = \sum_{x \in \Omega} \sqrt{\pi_x}\,|x\rangle\,|0\rangle\,.$$

## Quantum walk operator

Implementation $W_t := W(P_t) = U(P_t)^\dagger S U(P_t) R_{\mathcal{A}} U(P_t)^\dagger S U(P_t) R_{\mathcal{A}}$. With the swap $S$, and quantum walk update $U(P_t)$ :

$$U(P_t)\,|x\rangle\,|0\rangle = |x\rangle\,|p_x\rangle = |x\rangle \sum_{y \in \Omega} \sqrt{p_{xy}}\,|y\rangle\,, \ \forall \boldsymbol{x} \in \Omega,$$

where $p_{xy}$ denotes the transition probability from $x$ to $y$.

# Result 2 - Quantum

## Fully-Quantum Cutting-Plane Quantum Walk Algorithm (QCP-QW)

- **High-level:**
  Outer loop: CP method, performing $O^*(D\log(1/\gamma))$ updates to prepare final solution $|\pi_r\rangle$.
  Inner loop: *Grover's search* to amplify the probability of finding a separating hyperplane.

- **Middle-level:**
  Applying $O^*(D)$ non-destructive quantum mean estimation [a] $U_t^{\text{mean}}$ on $|\tilde{\pi}_t\rangle$ to acquire the multivariate centroid $\boldsymbol{z}_{t+1}$ of $\tilde{\boldsymbol{\pi}}_{t+1}$.
  Similarly, applying $O^*(D)$ non-destructive affine estimation $U_t^{\text{aff}}$ on $|\tilde{\pi}_t\rangle$ to obtain the affine transformation matrix $S_{t+1}$ of $\tilde{\boldsymbol{\pi}}_{t+1}$.

- **Low-level:**
  Applying the affine transformation $g_t$ to the states $|\tilde{\pi}_t\rangle$ to ensure the convex body $\mathcal{P}_t$ is in near-isotropic position.
  Applying the quantum walk evolution $\tilde{U}_{t,m}$ to generate the distribution $|\tilde{\pi}_{t+1}\rangle$ via

$$|g_t\tilde{\pi}_{t+1}\rangle = \tilde{U}_{t,m}\,|g_t\tilde{\pi}_t\rangle\,.$$

[a] For non-destructive multidimensional quantum amplitude estimation circuits, we point out the possibility to do so from [Harrow and Wei, 2020, Cornelissen et al., 2022, Cornelissen and Hamoudi, 2023, van Apeldoorn et al., 2023].

# Quantum Walk Evolution Operator $U_{t,m}$

## Grover's $\frac{\pi}{3}$ fix point search [Grover, 2005]

Two arbitrary quantum states with overlap $|\langle\pi_t|\pi_{t+1}\rangle|^2 \geq p$ $(0 < p \leq 1)$.
Given $|\pi_t\rangle$, $|\tilde{\pi}_{t+1}\rangle = U_{t,m}|\pi_t\rangle$, such that $\||\tilde{\pi}_{t+1}\rangle - |\pi_{t+1}\rangle\| \leq \epsilon_1 > 0$.
Choosing $m = O(\log(\log 1/\epsilon_1))$, unitaries $\{R_t, R_t^\dagger, R_{t+1}, R_{t+1}^\dagger\}$ are invoked at most $O(\log 1/\epsilon_1)$ times.

## Recursive quantum walk evolutions

$U_{t,0} = \mathbb{1}$, and $U_{t,i+1} = U_{t,i} \cdot R_t \cdot U_{t,i}^\dagger \cdot R_{t+1} \cdot U_{t,i}$.

The unitaries $R_t$, $R_{t+1}$ operators are selective phase shifts that $R_t = e^{\frac{i\pi}{3}}\Pi_t + \Pi_t^\perp$, where $\Pi_t$ is the orthogonal projector onto **span** $(|\pi_t\rangle)$.

## Example

$$U_{t,1} = U_{t,0} \cdot R_t \cdot U_{t,0}^\dagger \cdot R_{t+1} \cdot U_{t,0} = R_t \cdot R_{t+1}$$
$$U_{t,2} = U_{t,1} \cdot R_t \cdot U_{t,1}^\dagger \cdot R_{t+1} \cdot U_{t,1} = (R_t \cdot R_{t+1}) \cdot R_t \cdot (R_{t+1}^\dagger R_t^\dagger) \cdot R_{t+1} \cdot (R_t \cdot R_{t+1})$$
$$\dots$$
$$|\tilde{\pi}_{t+1}\rangle = U_{t,m}|\pi_t\rangle = U_{t,m-1} R_t U_{t,m-1}^\dagger R_{t+1} U_{t,m-1}|\pi_t\rangle = R_t \cdot R_{t+1} \cdots R_t \cdot R_{t+1}|\pi_t\rangle$$

## Example

The state $|\pi_t\rangle$ is a uniform superposition of states on the convex body $\mathcal{P}_t$. It can be decomposed in the subspace $\mathcal{H}_{\pi_t} = \textbf{span}\,(|\pi_{t+1}\rangle, |\pi_{t+1}^{\perp}\rangle)$ as $|\pi_t\rangle = \sin\theta_{t+1}|\pi_{t+1}\rangle + \cos\theta_{t+1}|\pi_{t+1}^{\perp}\rangle$, where

$$|\pi_{t+1}\rangle = \frac{\Pi_{\mathcal{P}_{t+1}}|\pi_t\rangle}{\|\Pi_{\mathcal{P}_{t+1}}|\pi_t\rangle\|} = \frac{\sum_{x\in\mathcal{P}_{t+1}}|x\rangle\langle x|\sum_{x'\in\mathcal{P}_t}\sqrt{\pi_{t,x'}}|x'\rangle|0\rangle}{\sqrt{\langle\pi_t|\Pi_{\mathcal{P}_{t+1}}|\pi_t\rangle}} = \frac{\sum_{x\in\mathcal{P}_{t+1}}\sqrt{\pi_{t,x}}|x\rangle|0\rangle}{\sqrt{\sum_{x\in\mathcal{P}_{t+1}}\pi_{t,x}}},$$

From the actions of quantum oracle $P_{t+1}$ in Eqn. (5):

$$P_{t+1}|\boldsymbol{w}\rangle|\boldsymbol{x}'_{t+1}y'_{t+1},\boldsymbol{z}_t\rangle|0\rangle = |\boldsymbol{w}\rangle|\boldsymbol{x}'_{t+1}y'_{t+1},\boldsymbol{z}_t\rangle|0\oplus f_{\boldsymbol{w}}(\boldsymbol{x}'_{t+1}y'_{t+1},\boldsymbol{z}_t)\rangle,$$

it can be easily verified that $P_{t+1}|\pi_{t+1}\rangle|0\rangle = |\pi_{t+1}\rangle|0\rangle$ and $P_{t+1}|\pi_{t+1}^{\perp}\rangle|0\rangle = |\pi_{t+1}^{\perp}\rangle|1\rangle$. Herewith, $R_{t+1} = e^{\frac{i\pi}{3}}\Pi_{t+1} + \Pi_{t+1}^{\perp}$ can be constructed by letting

$$R_{t+1} = P_{t+1}^{\dagger}(\mathbb{1}\otimes(e^{\frac{i\pi}{3}}|0\rangle\langle 0| + |1\rangle\langle 1|))P_{t+1}. \tag{6}$$

As has been shown in [Magniez et al., 2007] and [Wocjan and Abeyesinghe, 2008, Corollary 2], the phase estimation [Cleve et al., 1998] circuit $PE(W_t)$ can be leveraged to realise $R_t$.
For some $\epsilon_2 > 0$, there is a quantum circuit $V_t$ of $\mathbb{C}^d \otimes \mathbb{C}^d \otimes (\mathbb{C}^2)^{\otimes ac}$, using $PE(W_t)$ circuit $c = O(\log \frac{1}{\sqrt{\epsilon_2}}))$ times with $a = O(\log \frac{1}{\Delta})$ ancillary qubits each.

$$\tilde{R}_t = V_t^\dagger (\mathbb{1} \otimes (e^{\frac{i\pi}{3}} |0\rangle \langle 0|^{\otimes ac} + (\mathbb{1} - |0\rangle \langle 0|^{\otimes ac})))V_t, \tag{7}$$

that $\| \tilde{R}_t |\psi_k^\pm\rangle |0\rangle^{\otimes ac} - R_t |\psi_k^\pm\rangle |0\rangle^{\otimes ac} \| \leq 2\sqrt{\epsilon_2}$.

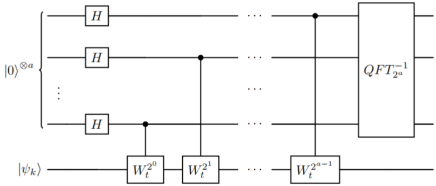



Figure 5: Phase estimation circuit $PE(W_t)$, where $H$ is the Hadamard gate, $QFT^{-1}$ denotes the inverse quantum Fourier transform circuit, and $W_t = U(P_t)^\dagger SU(P_t)R_{\mathcal{A}}U(P_t)^\dagger SU(P_t)R_{\mathcal{A}}$

# The Main Theorem

## Theorem (Hit-and-Run Quantum Walk)

*Let $P_t$ be an ergodic symmetric Markov chain defined by the* **Hit-and-Run** *algorithm with stationary uniform distribution $\boldsymbol{\pi}_t$ on the support $\mathcal{P}_t$ with an initial distribution $|\pi_t\rangle = |\pi'_t\rangle\,|0\rangle$. If the following properties hold:*

- **Mixing time:** $d_{\mathrm{TV}}(P_t^{\tau} \cdot \boldsymbol{\pi}_{t+1}, \boldsymbol{\pi}_t) \leq \epsilon$.
- **Warmness:** $\|\boldsymbol{\pi}_{t+1}/\boldsymbol{\pi}_t\| = O(1)$, *where* $\|\boldsymbol{\pi}/\boldsymbol{\sigma}\| := \int_{\mathbb{R}^D} \frac{\pi(\boldsymbol{x})}{\sigma(\boldsymbol{x})}\pi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$.
- **Overlap:** $|\langle \pi_t|\pi_{t+1}\rangle| = \int_{\mathbb{R}^D} \sqrt{\pi_t(\boldsymbol{x})\pi_{t+1}(\boldsymbol{x})}\,d\boldsymbol{x} = \Omega(1)$.

*Then, we can obtain $|\tilde{\pi}_{t+1}\rangle = \tilde{U}_{t,m}\,|\pi_t\rangle$ with $\|\,|\tilde{\pi}_{t+1}\rangle - |\pi_{t+1}\rangle\,\| \leq \epsilon$, using $O\big(\sqrt{\tau}\log(1/\epsilon)\big)$ calls to the quantum walk operators $W_t$.*

- Using *Hit-and-Run*, start from $\boldsymbol{\pi}_{t+1}$, takes $\tau = O^*(D^3)$ steps to generate a uniform distribution $\boldsymbol{\pi}_t$. The quantum walk evolution $\Rightarrow$ $U_{t,m}$ can be achieved by using $O^*\left(D^{1.5}\right)$ calls to the quantum walk operators $W_t$.

# Lemmas

For uniform $\boldsymbol{\pi}_t$ over $\mathcal{P}_t$, its quantum sample in the continuous space
$|\pi_t\rangle' = \int_{\mathcal{P}_t} \sqrt{\pi_{t,x}}\,|x\rangle\; dx = \int_{\mathcal{P}_t} \sqrt{\frac{1}{V(\mathcal{P}_t)}}\,|x\rangle\; dx.$

## Overlap lemma

From [Bertsimas and Vempala, 2004], the volume left in each round is at least 1/3 with high probability.

$$|\langle \pi_t | \pi_{t+1}\rangle| = \int_{\mathbb{R}^D} \sqrt{\frac{1}{V(\mathcal{P}_t)V(\mathcal{P}_{t+1})}}\mathrm{d}\boldsymbol{x} = \frac{V(\mathcal{P}_{t+1})}{\sqrt{V(\mathcal{P}_t)V(\mathcal{P}_{t+1})}} \geq \sqrt{\frac{1}{3}} = \Omega(1).$$

Hence, in QCP-QW algorithm, the overlap condition $|\langle \pi_t | \pi_{t+1}\rangle| = \Omega(1)$ is met.

## Warmness lemma

$$\|\boldsymbol{\pi}_{t+1}/\boldsymbol{\pi}_t\| = \int_{\mathbb{R}^D} \frac{\pi_{t+1}(\boldsymbol{x})}{\pi_t(\boldsymbol{x})}\pi_{t+1}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_{\mathcal{P}_{t+1}} \frac{\frac{1}{V(\mathcal{P}_{t+1})}}{\frac{1}{V(\mathcal{P}_t)}}\frac{1}{V(\mathcal{P}_{t+1})}\mathrm{d}\boldsymbol{x} = \frac{V(\mathcal{P}_t)}{V(\mathcal{P}_{t+1})} \leq 3 = O(1).$$

Hence, in QCP-QW algorithm, the warmness conditions $\|\boldsymbol{\pi}_{t+1}/\boldsymbol{\pi}_t\| = O(1)$ is met.

## A summary of algorithms and *data query* complexities

| Algorithm | Complexity |
|---|---|
| Version Space Perceptron | $O^*(D^{\frac{1}{2}}\frac{N}{\gamma^D})$ |
| QVSP | $O^*(D^{\frac{1}{4}}\sqrt{\frac{N}{\gamma^D}})^{\mathrm{a}}$ |
| Online Perceptron | $O^*(\frac{1}{\gamma^2}\log(1/\gamma^2)N)$ |
| OQP | $O^*(\frac{1}{\gamma^2}\log(1/\gamma^2)\sqrt{N})$ |
| CP-RW (Hit-and-Run) | $O^*(D\log(\frac{1}{\gamma})(N+D^5\log(\frac{1}{\gamma})))^{\mathrm{b}}$ |
| HCP-RW (1) | $O^*(D\log(\frac{1}{\gamma})(\sqrt{N}+D^{4.5}\sqrt{\log(\frac{1}{\gamma})}))$ $^{\mathrm{b}}$ |
| QCP-QW (2) | $O^*(D\log(\frac{1}{\gamma})(\sqrt{N}+D^3\sqrt{\log(\frac{1}{\gamma})}))$ |

[a] Here we only present the corrected result of the improved-QVSP of Liao et al. (2024).

[b] For Hit-and-Run CP-RW and HCP-RW introduced in Section 4.2, we simplify the result using $O(\log^2 N)\le O(D)$. Otherwise, tighter bounds can be achieved by replacing an $O^*(D)$ factor in $O^*(D^5)$ and $O^*(D^{4.5})$ with $\log^2 N$.

Bertsimas, D. and Vempala, S. (2004).
Solving convex programs by random walks.
*J. ACM*, 51(4):540–556.

Cleve, R., Ekert, A., Macchiavello, C., and Mosca, M. (1998).
Quantum algorithms revisited.
*Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):339–354.

Cornelissen, A. and Hamoudi, Y. (2023).
A sublinear-time quantum algorithm for approximating partition functions.
In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1245–1264. SIAM.

Cornelissen, A., Hamoudi, Y., and Jerbi, S. (2022).
Near-optimal quantum algorithms for multivariate mean estimation.
In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, page 33–43, New York, NY, USA. Association for Computing Machinery.

Grover, L. K. (1996).
A fast quantum mechanical algorithm for database search.
In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96, page 212–219, New York, NY, USA. Association for Computing Machinery.

Grover, L. K. (2005).
A different kind of quantum search.
*arXiv preprint quant-ph/0503205*.

Harrow, A. W. and Wei, A. Y. (2020).

Adaptive quantum simulated annealing for bayesian inference and estimating partition functions.
In *Proceedings of the Thirty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '20, page 193–212, USA. Society for Industrial and Applied Mathematics.

Kapoor, A., Wiebe, N., and Svore, K. (2016).
Quantum perceptron models.
In Lee, D., Sugiyama, M., Luxburg, U., Guyon, I., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc.

Khachiyan, L. G. (1979).
A polynomial algorithm in linear programming.
In *Doklady Akademii Nauk*, volume 244, pages 1093–1096. Russian Academy of Sciences.

Liao, P., Sanders, B. C., and Byrnes, T. (2024).
Quadratic quantum speedup for perceptron training.
*Physical Review A*, 110(6):062412.

Lovász, L. (1999).
Hit-and-run mixes fast.
*Mathematical Programming*, 86:443–461.

Magniez, F., Nayak, A., Roland, J., and Santha, M. (2007).
Search via quantum walk.
In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07, page 575–584, New York, NY, USA. Association for Computing Machinery.

Novikoff, A. B. (1962).
On convergence proofs on perceptrons.

In *Proceedings of the Symposium on the Mathematical Theory of Automata*, volume 12, pages 615–622. New York, NY.

Smith, R. L. (1984).
Efficient monte carlo procedures for generating points uniformly distributed over bounded regions.
*Operations Research*, 32(6):1296–1308.

Szegedy, M. (2004).
Quantum speed-up of Markov chain based algorithms.
In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 32–41. IEEE.

Vaidya, P. (1989).
A new algorithm for minimizing convex functions over convex sets .
In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 338–343, Los Alamitos, CA, USA. IEEE Computer Society.

van Apeldoorn, J., Cornelissen, A., Gilyén, A., and Nannicini, G. (2023).
Quantum tomography using state-preparation unitaries.
In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1265–1318. SIAM.

Wocjan, P. and Abeyesinghe, A. (2008).
Speedup via quantum sampling.
*Phys. Rev. A*, 78:042336.

Thanks for your attention.

Figure 6: https://arxiv.org/abs/2503.17308