Saint Petersburg State University

Solving linear equations on an adiabatic quantum computer (AQC)

Vladimir S. Shalgin
Scientific advisor
prof. Sergey B. Tikhomirov

Sochi, 2021

Adiabatic quantum algorithm

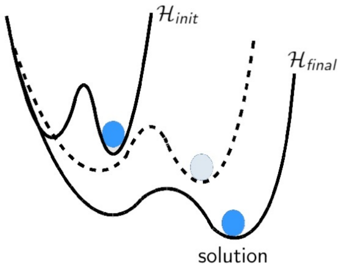
The Hamiltonian of the system is equal to the energy of the system.

- $\mathcal{H}_{init}(\sigma)$ is an initial Hamiltonian, σ is a state of the system.
- $\mathcal{H}_{final}(\sigma)$ is a final Hamiltonian encoding the solution of some problem.

Evolution of $\mathcal{H}_{init}$ into $\mathcal{H}_{final}$ is an evolution of the system with the Hamiltonian

$$\mathcal{H}(t) = (1 - t/T)\mathcal{H}_{init} + (t/T)\mathcal{H}_{final},$$

where T is the time of system evolution.



Types of the final Hamiltonian

Ising problem

$$H_{ising}(\mathbf{s}) = \sum_{i=1}^{N} h_i s_i + \sum_{i=1}^{N} \sum_{j=i+1}^{N} J_{ij} s_i s_j,$$

$s_i \in \{+1, -1\},\ J_{ij}, h_i \in \mathbb{R}$

QUBO problem

$$H_{qubo}(\mathbf{q}) = \sum_{i=1}^{N} h_i q_i + \sum_{i=1}^{N} \sum_{j=i+1}^{N} J_{ij} q_i q_j = \sum_{i=1}^{N} \sum_{j=i}^{N} J_{ij} q_i q_j,$$

$q_i \in \{0, 1\},\ J_{ij}, h_i \in \mathbb{R},\ J_{ii} = h_i,\ (q_i^2 = q_i).$

The main idea

The desired solution of some problem is the point where H has its global minimum.

The problem

- We consider a one-dimensional equation $ax = b$.
- Construct a final Hamiltonian $H(x) = (ax - b)^2$ whose minimum point is b/a.
- Represent the variable as $x = c\chi - d$, where $\chi = \sum_{i=0}^{R-1} 2^{-i} q_i \in [0, 2)$ with $\forall q_i \in \{0, 1\}$ and $c, d > 0$.
- Hence $x \in [-d, 2c - d)$
- The Hamiltonian with variables q_i has the form

$$H(q_0, \ldots, q_{R-1}) = \sum_{i=0}^{R-1} h_i q_i + \sum_{i=0}^{R-1} \sum_{j=i+1}^{R-1} J_{ij} q_i q_j$$

$$h_i = 2^{-i} ac \left(2^{-i} ac - 2(ad + b)\right)$$
$$J_{ij} = 2a^2 c^2 2^{-i-j}$$

D-Wave's computers

D-Wave 2000Q (2041 qubits)

Advantage System (5627 qubits)

Probabilistic nature of adiabatic quantum computers

Boltzmann distribution (BD)

Let $H(x)$ be a Hamiltonian of the system,
$x_1, \ldots, x_N$ be all possible states of the system.
Then the probability that the system is in the state x_k

$$P(x_k) = \frac{1}{Q} e^{-\beta H(x_k)},$$

where $Q = \sum e^{-\beta H(x_j)}$, $\beta = \frac{1}{k_B T}$ is the parameter of the distribution, k_B is the Boltzmann constant, T is the thermodynamic temperature of the system. We will omit physical nature of the parameter β in further considerations.

The limit distribution

If the number of states of the system tends to infinity and the states become more dense then the distribution becomes normal.

Statement

1) Let X_n be a random variable such that

$$X_n \in \Omega_n = \left\{ \pm \sum_{i=-n}^{n-1} q_i \, 2^i \;\middle|\; q_i \in \{0,1\} \right\} \subset \mathbb{R}.$$

2) Let P_n be a probability measure on Ω_n such that

$$P_n(x) = P_n(X_n = x) \propto e^{-\beta H(x)},$$

where $H(x) = (ax-b)^2$, $\beta > 0$. Then

$$X_n \xrightarrow{\mathcal{D}} N\left(\frac{b}{a}, \frac{1}{a\sqrt{2\beta}}\right), \quad n \to \infty$$

Proof

- We let $\mu_n(A) := \sum_{x_k \in \Omega_n} P_n(x_k)\delta_x(A)$ be a measure on $\mathbb{R}$ where
- $P_n(x_k) = \frac{1}{Q_n} e^{-\beta(ax_k - b)^2}, \ Q_n = \sum_{x_i \in \Omega_n} e^{-\beta(ax_i - b)^2}$
- We then compute an asymptotic formula for Q_n

$$Q_n = 2^n \sum_{x_i \in \Omega_n} \frac{1}{2^n} e^{-\beta(ax_i - b)^2}$$

$$\sim 2^n \int_{-\infty}^{\infty} e^{-\beta(at - b)^2} dt = \frac{2^n \sqrt{\pi}}{a\sqrt{\beta}}$$

- Now we can calculate the limit of μ_n

$$\mu_n(A) = \frac{1}{Q_n} \sum_{x_k \in \Omega_n} e^{-\beta(ax_k - b)^2} \cdot \delta_{x_k}(A)$$

$$\xrightarrow{n \to \infty} \frac{a\sqrt{\beta}}{\sqrt{\pi}} \int_A e^{-\beta(at - b)^2} dt = \frac{1}{\sigma\sqrt{2\pi}} \int_A e^{-\frac{(t - b/a)^2}{2\sigma^2}} dt, \quad \sigma = \frac{1}{a\sqrt{2\beta}}$$

Parameters of the distribution of solutions

Normal distribution

- Let $F_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{(t-b/a)^2}{2\sigma^2}} dt$ be the CDF of the limit normal distribution, $\sigma = \frac{1}{a\sqrt{2\beta}}$
- Let $F_N(x)$ be the eCDF of a sample then

$$\hat{\sigma} = \arg\min_{\sigma} \|F_\sigma(x) - F_N(x)\|_{L^2}$$

Boltzmann distribution

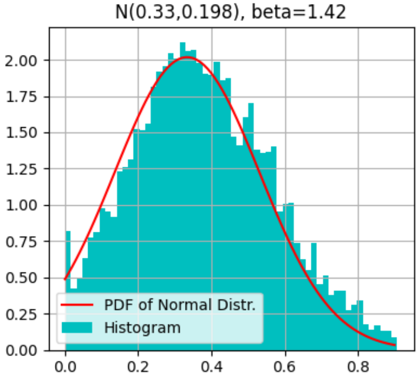
- Let Q be the BD, q_i be the probability of x_i,
- Let P be the empirical distribution, p_i be the frequency of x_i.

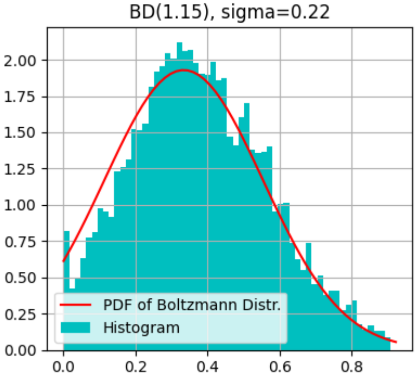
$$\hat{\beta} = \arg\min_{\beta} \mathrm{JSD}(P||Q)$$

$\mathrm{JSD}(P||Q) = \frac{1}{2}D_{KL}(P||M) + \frac{1}{2}D_{KL}(Q||M)$, $M = \frac{1}{2}(P+Q)$, e.g. $D_{KL}(P||Q) = \sum_i p_i \log \frac{p_i}{q_i}$

Examples

- $3x = 1$
- $x = \chi = \sum_{i=0}^{R-1} 2^{-i} q_i \Rightarrow x \in [0, 2)$, we use $\sigma = \frac{1}{a\sqrt{2\beta}}$, $R = 7$

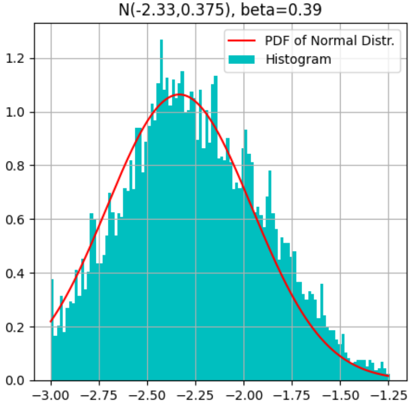


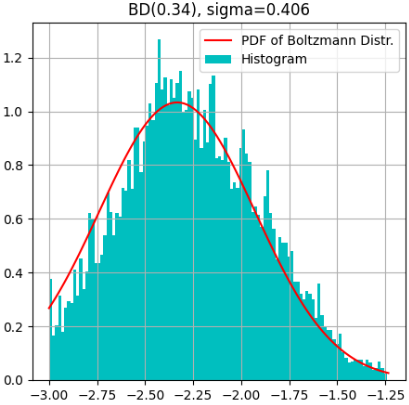


- The best solution is 0.328
- Sample mean is 0.387

Examples

- $3x = -5$
- $x = 2\chi - 3 \in [-3, 1), \;\; R = 8$





- The best solution is -2.328
- Sample mean is -2.246

Improving the solution

- Let x_k be an approximate solution of $ax = b$
- The exact solution $x = x_k + \Delta_k$
- The equation for Δ_k is $a\Delta_k = b - ax_k$,
- The next approximation to the solution is $x_{k+1} = x_k + \Delta_k$
- Assume $\Delta_k \sim N\left(\frac{b-ax_k}{a}, \frac{1}{a\sqrt{2\beta}}\right)$, according to the statement about the limit distribution
- Then x_k does not converge to $\frac{b}{a}$

Improving the solution

Another scheme

- Let $a \in \left(\frac{1}{2}, 1\right)$ and $\widetilde{\Delta}_k = 2^l \Delta_k$ such that $|\widetilde{\Delta}_k| \in \left(\frac{1}{2}, 1\right)$
- Equation for $\widetilde{\Delta}_k$ then is $a\widetilde{\Delta}_k = 2^l(b - ax_k)$
- The same assumption gives $\Delta_k \sim \frac{1}{2^l} N\left(\frac{2^l(b-ax_k)}{a}, \frac{1}{a\sqrt{2\beta}}\right)$
- If $a \in \left(\frac{1}{2}, 1\right)$ then $|b - ax_k| \in \left(\frac{1}{2^{l+2}}, \frac{1}{2^l}\right)$ but assume $|b - ax_k| \approx \frac{1}{2^l}$
- Thus we can write $\Delta_k \sim (b - ax_k)N\left(\frac{1}{a}, \frac{1}{a\sqrt{2\beta}}\right)$
- Then $x_{k+1} = x_k + \Delta_k$ does converge to $\frac{b}{a}$

Improving the solution

Theorem (convergence of the approximations)

Let $x_0 = 0$, $x_{k+1} = x_k + \Delta_k$ for $k \geqslant 0$, where Δ_k is a random variable distributed as $(b - ax_k)N\left(\frac{1}{a}, \sigma\right)$, where $a > 0, b \in \mathbb{R}$. Then $x_k \to \frac{b}{a}$ in probability as $k \to \infty$ if $a\sigma < \sqrt{2}\, e^{\frac{\gamma}{2}}$, where $\gamma \approx 0.577$ is the Euler-Mascheroni constant.

It can be reformulated if $\sigma = \frac{1}{a\sqrt{2\beta}}$ then the condition $a\sigma < \sqrt{2}\, e^{\frac{\gamma}{2}}$ becomes $\beta > \frac{1}{4}e^{-\gamma}$

Proof of the theorem

Consider $z_k = x_k - \frac{b}{a}$ and prove $z_k \xrightarrow{P} 0$. Denote $\xi_k \sim N\left(\frac{1}{a}, \sigma\right)$. We can simply derive the recurrence relation for z_k

$$z_k = -(a\xi_{k-1} - 1)z_{k-1} = -\zeta_{k-1}z_{k-1}$$

where $\zeta_k \sim N(0, a\sigma)$. Since $z_0 = -\frac{b}{a}$ then

$$z_k = (-1)^{k+1}\frac{b}{a}\zeta_0\zeta_1 \dots \zeta_{k-1}$$

Taking log of the above relation we get

$$\log|z_k| = \log\left|\frac{b}{a}\right| + \log|\zeta_0| + \log|\zeta_1| + \dots + \log|\zeta_{k-1}|$$

Next we find an expectation of $\log|\zeta_i|$

Proof of the theorem

Omitting details of calculations of an integral we get

$$\mathbb{E} \log |\zeta_i| = \log \frac{a\sigma}{\sqrt{2}} - \frac{\gamma}{2}$$

(we use here a nice equality $\int\limits_0^\infty e^{-x^2} \log x \, dx = -\frac{1}{4}\sqrt{\pi}\,(\gamma + 2\log 2)$)

Since $a\sigma < \sqrt{2}e^{\frac{\gamma}{2}}$ then $\mathbb{E} \log |\zeta_i| < 0$. So by Law of large numbers we have

$$\frac{1}{k}\left(\log|\zeta_0| + \log|\zeta_1| + \ldots + \log|\zeta_{k-1}|\right) \xrightarrow{P} \log \frac{a\sigma}{\sqrt{2}} - \frac{\gamma}{2} < 0$$

Hence $\log\left|\frac{b}{a}\right| + \log|\zeta_0| + \ldots + \log|\zeta_{k-1}| \xrightarrow{P} -\infty$ exponentiating we get $|z_k| \xrightarrow{P} 0$.

Speed of convergence

Statement

Let $x_0 = 0$, $x_{k+1} = x_k + \Delta_k$, *where* $\Delta_k \sim (b - a x_k) N\left(\frac{1}{a}, \sigma\right)$ *and* $a\sigma < \sqrt{2}\, e^{\frac{\gamma}{2}}$. *Let*

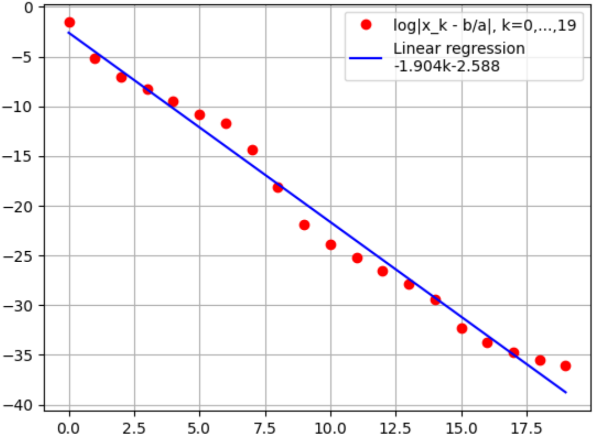
$$s_k = \left(\frac{\sqrt{2}\, e^{\gamma/2-\delta}}{a\sigma}\right)^k = \left(2\sqrt{\beta e^{\gamma-2\delta}}\right)^k$$

where $\delta > 0$. *Then* $\left(x_k - \frac{b}{a}\right) s_k \xrightarrow{P} 0$.

- Let $\delta = 0$ and assume $s_k \left|x_k - \frac{b}{a}\right| \approx 1$
- So $\left|x_k - \frac{b}{a}\right| \approx e^{\rho k}$ such that $\left(e^{\rho} \cdot 2\sqrt{\beta e^{\gamma}}\right)^k \approx 1$
- Hence $\beta \approx \frac{1}{4e^{\gamma+2\rho}}$

Speed of convergence

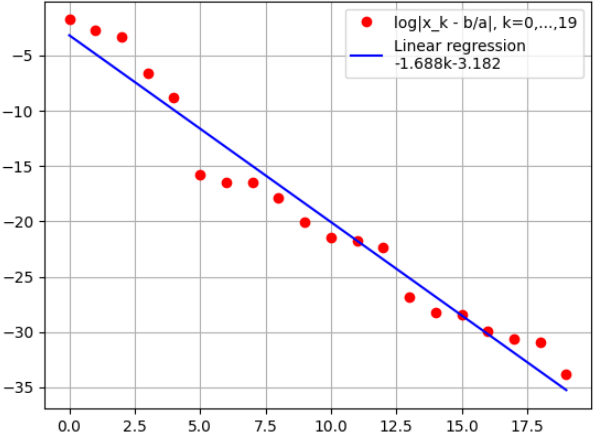
- $0.75x = 0.25$
- On each step we are in $[-1, 1)$ and $R = 10$



- $\beta = 6.32$

Speed of convergence

- $0.875x = -0.625$
- On each step we are in $[-1, 1)$ and $R = 10$



- $\beta = 4.1$

The case of a truncated normal distribution

Limit distribution

- It the statement about the limit distribution of solutions X_n was a random variable which took values from $\Omega_n = \left\{\pm \sum_{i=-n}^{n} q_i \, 2^i \,\middle|\, q_i \in \{0,1\}\right\} \subset \mathbb{R}$
- Now let $X_n \in \Omega_n = \left\{\pm \sum_{i=-n}^{n-1} q_i \, 2^i \in (r_1, r_2) \,\middle|\, q_i \in \{0,1\}\right\}$, $r_1 < r_2 \in \mathbb{R}$
- As before $P_n(X_n = x) \propto e^{-\beta(ax-b)^2}$ — Boltzmann distribution over Ω_n
- Then $X_n \xrightarrow{\mathcal{D}} N\left(\frac{b}{a}, \frac{1}{a\sqrt{2\beta}}, r_1, r_2\right)$ which has the density

$$p(t) = \frac{1}{\sigma\sqrt{2\pi}Q} e^{-\frac{(t-b/a)^2}{2\sigma^2}} \mathbf{1}_{(r_1,r_2)}(t),$$

where $Q = \Phi\left(\frac{r_2 - b/a}{\sigma}\right) - \Phi\left(\frac{r_1 - b/a}{\sigma}\right)$ and $\sigma = \frac{1}{a\sqrt{2\beta}}$

Case of a truncated normal distribution

Improving the solution

- Let x_k be an approximate solution of $ax = b$, then the exact solution is $x = x_k + \Delta_k$
- The equation for Δ_k is $a\Delta_k = b - ax_k$, the next approximation to the solution is $x_{k+1} = x_k + \Delta_k$
- Assume $\Delta_k \sim N\left(\frac{b-ax_k}{a}, \sigma, \frac{b-ax_k}{a}(1 \pm \varepsilon)\right)$, where $\sigma = \frac{1}{a\sqrt{2\beta}}$, $\varepsilon > 0$
- $\widetilde{\Delta}_k = 2^l \Delta_k$ such that $|\widetilde{\Delta}_k| \in \left(\frac{1}{2}, 1\right)$, then $a\widetilde{\Delta}_k = 2^l(b - ax_k)$
- Hence $\Delta_k \sim \frac{1}{2^l}\widetilde{N}\left(\frac{2^l(b-ax_k)}{a}, \sigma, \frac{2^l(b-ax_k)}{a}(1 \pm \varepsilon)\right)$
- If $a \in \left(\frac{1}{2}, 1\right)$ then $|b - ax_k| \in \left(\frac{1}{2^{l+2}}, \frac{1}{2^l}\right) \Rightarrow |b - ax_k| \approx \frac{1}{2^l}$
- Thus $\Delta_k \sim (b - ax_k)\widetilde{N}\left(\frac{1}{a}, \sigma, \frac{1}{a}(1 \pm \varepsilon)\right)$

Theorem

Let $x_0 = 0$, $x_{k+1} = x_k + \Delta_k$ where Δ_k is the random variable distributed as $(b - ax_k)\widetilde{N}\left(\frac{1}{a}, \sigma, \frac{1}{a}(1 \pm \varepsilon)\right)$ where $a, \varepsilon, \sigma > 0, b \in \mathbb{R}$. Then if $\varepsilon \leqslant e$ then $x_k \xrightarrow{P} \frac{b}{a}$. If $\varepsilon > e$ and $F(\varepsilon, a\sigma) < 0$ where

$$F(\varepsilon, a\sigma) = \log\left(\sqrt{2}a\sigma\right) + \left(\int_0^{\frac{\varepsilon}{\sqrt{2}a\sigma}} e^{-x^2}\, dx\right)^{-1} \left(\int_0^{\frac{\varepsilon}{\sqrt{2}a\sigma}} e^{-x^2} \log x\, dx\right),$$

then $x_k \xrightarrow{P} \frac{b}{a}$.

Remark

If $\varepsilon \to \infty$ then the first integral $\to \frac{\sqrt{\pi}}{2}$, the second one $\to -\frac{\sqrt{\pi}}{4}(\gamma + 2\log 2)$, so $F(\varepsilon, a\sigma) \to \log\frac{a\sigma}{\sqrt{2}} - \frac{\gamma}{2} < 0 \Rightarrow a\sigma < \sqrt{2}e^{\gamma/2}$

For an equation

$$ax + by = c$$

We have the distribution of solutions

$$P((x, y)) \propto e^{-\beta(ax+by-c)^2}$$

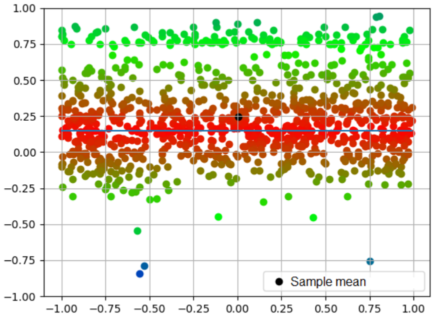
- If we want to find the limit distribution of solutions then we cannot impose no restrictions on the values of solutions since the integral $\iint_{\mathbb{R}^2} e^{-\beta(ax+by-c)^2} dxdy$ does not converge.
- If we let Ω be the set of values that solutions can take then the limit distribution has the density function

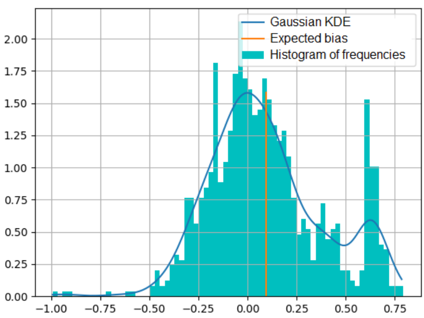
$$\frac{e^{-\beta(ax+by-c)^2}}{\iint\limits_{\Omega} e^{-\beta(ax+by-c)^2} dxdy} \cdot \mathbf{1}_{\Omega}(x, y)$$

An equation with two variables

Limit distribution of solutions

- We take $\Omega = [-1,1]^2$ and sample 1000 times the «solutions» of the equations $2y = 0.3$ with precision $R = 8$ bits.
- Also we project samples onto the line orthogonal to $2y = 0.3$.
- Expected bias is the mean distance between samples and the line $2y = 0.3$.

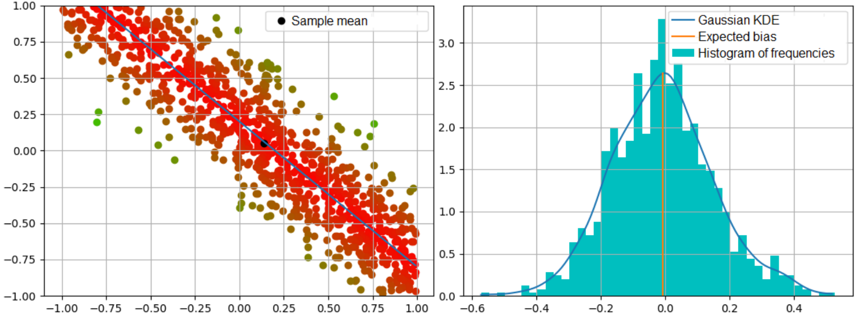




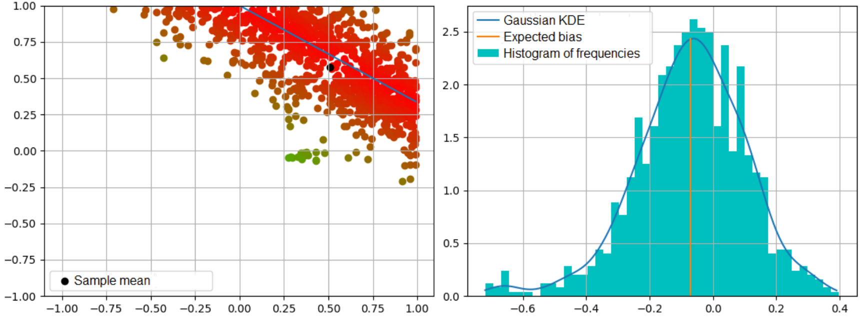
An equation with two variables

Limit distribution of solutions

Samples for $x + y = 0.2$.



Samples for $2x + 3y = 3$.



Two equations with one variable

$$\begin{cases} a_1 x = b_1, \\ a_2 x = b_2. \end{cases}$$

- Solving this system is equivalent to solving one equation

$$||a||x = \left(\frac{a}{||a||}, b \right),$$

  whose solution is pseudo-solution of minimal norm of the above system. Here $a = (a_1, a_2),\ b = (b_1, b_2)$.
- This can be generalized to any overdeterminated system, i.e. finding a normal pseudo-solution of the system $(n + k) \times n$ is equivalent to solving some system $n \times n$.
- But is there an efficient algorithm for finding this quadratic system?

$$Ax = b \quad \text{or} \quad \begin{cases} a_{11}x_1 + a_{12}x_2 = b_1, \\ a_{21}x_1 + a_{22}x_2 = b_2. \end{cases}$$

- Let the Boltzmann distribution of solutions have the Hamiltonian $||Ax - b||^2$ and parameter β.
- Let $\det A \neq 0$, A_1 and A_2 be the first and the second columns of A. If there is no restrictions on the range of values of x
- then the limit distribution of solutions is $\mathcal{N}(\mu, \mathcal{K})$, where

$$\mu = A^{-1}b, \quad \mathcal{K}^{-1} = 2\beta \begin{pmatrix} (A_1, A_1) & (A_1, A_2) \\ (A_1, A_2) & (A_2, A_2) \end{pmatrix}$$

Sampling from «almost» normal distribution using AQC

- For the one dimensional case if we want to sample from normal distribution $\mathcal{N}(\mu, \sigma)$ then we can construct a Hamiltonian for the Boltzmann distribution with parameter β.

$$H(x) = \left(\frac{x}{\sigma\sqrt{2\beta}} - \frac{\mu}{\sigma\sqrt{2\beta}} \right)^2$$

  where we used the statement about the limit distribution of solutions of an equation $ax = b$.
- The same we can do for multivariate normal distribution.
- But it is time-consuming to construct a Hamiltonian because in order to do this we (again) need to invert given covariance matrix.