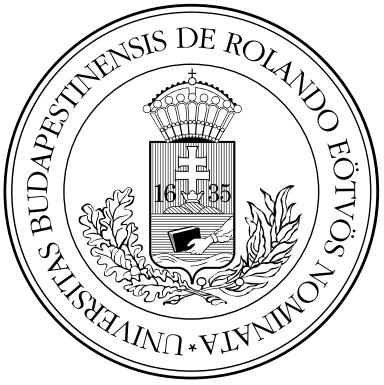


Eötvös Loránd Tudományegyetem
Természettudományi Kar

MSc Szakdolgozat
Alkalmazott matematikus

Kvantumos algoritmusok a várhatóérték becslésére

Készítette:
Csatári Jakab

Témavezető:
Dr. Gilyén András

2024

NYILATKOZAT

Név: Csatári Jakab

ELTE Természettudományi Kar, szak: Alkalmazott matematikus MSc

NEPTUN azonosító: D8REY9

Szakdolgozat címe:
Kvantumos algoritmusok a várhatóérték becslésére

A **szakdolgozat** szerzőjeként fegyelmi felelősségem tudatában kijelentem, hogy a dolgozatom önálló szellemi alkotásom, abban a hivatkozások és idézések standard szabályait következetesen alkalmaztam, mások által írt részeket a megfelelő idézés nélkül nem használtam fel.

Budapest, 2024.06.04.

Csatári Jakab

a hallgató aláírása

Tartalomjegyzék

Köszönetnyilvánítás

MSc-s tanulmányaim alatt a matematika sok számomra érdekes részével találkoztam, ezek egyike volt a kvantumszámítás, melynek alapjait volt szerencsém Gilyén András előadásai alapján elsajátítanom. Hálás köszönetem szeretném kifejezni neki, hogy ezt követően több féléven keresztül vállalta az önálló projektek és a szakdolgozatom témavezetését. Nagyon köszönöm a témaajánlásokat, a sok türelmet, tanácsot és hogy a konzultációk során olyan jól rávezetett, hogy mely részeket érdemes még jobban kibontanom és megértenem.

Illetve nem utolsó sorban köszönöm feleségemnek is a sok türelmet és megértést, amit a tanulmányaim során kaptam tőle.

Bevezetés

A kvantumalgoritmusok kitalálásának és fejlesztésének célja, hogy a klasszikus algoritmusokhoz képesti gyorsításokat érjünk el. A legközismertebb példa Shor exponenciális gyorsítást elérő algoritmusa a prímfaktorizációra. Sok esetben (mint például Grover keresésénél) négyzetes gyorsítás is igen jelentős lehet, különösen ha a módszer széleskörben alkalmazható. A szakdolgozatomban a Monte Carlo módszerek hátterében álló várhatóérték becslési problémáról írok, és arról, hogy hogyan lehet kvantumos gyorsítást elérni. A probléma arról szól, hogy minták alapján szeretnénk megbecsülni a várható értékét egy X diszkrét valószínűségi változónak nagy valószínűséggel minél pontosabban. Azt hasonlítom össze, hogy klasszikus és kvantumos esetben mekkora különbség van aközött, hogy mennyi mintát kell venni ugyanolyan pontosság eléréséhez.

Kothari és O'Donnell megmutatta, hogy egydimenziós valószínűségi változó esetén kvantumosan négyzetesen kevesebb mintát elég venni. Ez az optimális gyorsítás magában foglalja a probléma egy döntési problémává redukálását, valamint egy speciális kapu fázisbecslését, amely nagy valószínűséggel eldönti a két eset közötti különbséget. Magasabb dimenzióban a legjobb ismert módszer logaritmikus faktoroktól eltekintve optimális. Különbséget teszünk azon esetek között, amikor a dimenzió d legalább annyi, mint a megengedett minták száma n - ekkor nincs jobb algoritmus, mint a klasszikus - valamint aközött, amikor $n > d$, ahol $\widetilde{O}(\sqrt{d/n})$-szeres gyorsítás érhető el. Cornelissen, Hamoudi és Jerbi ügyes technikáit tárgyalom, a fő gondolat az, hogy a várható érték bizonyos vektorok szerinti skalárszorzatai amplitúdókká kódolhatóak és az inverz kvantum Fourier-transzformáció alkalmazásával az $\mathbb{E}[X]$ függvényét kapjuk.

A szakdolgozat utolsó fejezetében egy pénzügyi, gyakorlati alkalmazást is tárgyalok, az amerikai opció problémát. Ez egy optimális megállási probléma, amire klasszikus esetben alkalmazzák az LSM (least squares Monte Carlo) módszert, ami mintavételekre és átlagokkal való számolásra támaszkodik. A kvantumos megközelítés jelentős gyorsítást kínál azáltal, hogy a várható értékeket kvadratikusan kevesebb mintából számítja ki. Míg a klasszikus módszernél az összes mintát elég kezdetben venni, a kvantumos algoritmus többszöri mintavételt igényel, de az összesített időkomplexitás így is számottevően jobb - ha egy mintavételt egységnyi idejűnek tekintünk, akár négyzetes gyorsítás érhető el.

1. Kvantumszámítás és előkészületek

Az első fejezetben a kvantumszámítás alapjairól szeretnék írni. Alapvető fogalmakat és koncepciókat mutatok be, illetve előkészítek néhány bevezető kvantumos algoritmust, amiket az elkövetkező fejezetekben tárgyalt várhatóérték becslési problémához használunk. Ezek leírását részben [dW19] alapján szedtem össze.

1.1. A kvantumszámításról

A klasszikus számítási modell alapegysége a bit, ami két lehetséges állapotban lehet (0 vagy 1 az értéke). Ezzel szemben a kvantumszámítás az úgynevezett qubitekre épül, ami algebrai szempontból a bit általánosítása.

Jelölés 1.1.1.: Legyenek $|0\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ és $|1\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

Egy qubit állapota lehet $|0\rangle$, $|1\rangle$ vagy ezeknek egy szuperpozíciója.

Definíció 1.1.2.: Szuperpozíciónak hívjuk az $\alpha_0|0\rangle + \alpha_1|1\rangle$ kvantumállapotot, ahol $\alpha_0, \alpha_1 \in \mathbb{C}$ és $|\alpha_0|^2 + |\alpha_1|^2 = 1$.

Tehát egy szuperpozícióban levő qubit állapota felírható egy $\mathbb{C}^2$ vektorral (ahogy az 1.1.1. jelölés is sugallta). Az α_0, α_1 értékeket amplitúdóknak hívjuk.

Definíció 1.1.3.: (Mérés) Egy szuperpozícióban lévő qubit állapotát nem tudjuk vizsgálni, ahhoz hogy tudjunk róla mondani valamit meg kell mérni. Mérés alatt pedig azt értjük, hogy az állapota "összeomlik", azaz innentől $|\alpha_0|^2$ valószínűséggel $|0\rangle$ lesz, $|\alpha_1|^2$ valószínűséggel pedig $|1\rangle$. Illetve előbbi esetben 0-t, utóbbiban 1 értéket mérünk.

Több qubit összekapcsolt állapotát a tenzorszorzat segítségével tudjuk felírni, pl. két qubit esetén $|00\rangle := |0\rangle \otimes |0\rangle$. Egy n qubites rendszer kvantumállapotát, pedig a $|00...00\rangle, |00...01\rangle, ..., |11...11\rangle$ klasszikus állapotok lineáris kombinációjaként írhatjuk fel - ahol továbbra is teljesülni fog hogy az

amplitúdók abszolútértékének négyzetösszege 1, illetve megmérve értelemszerűen egy n hosszú bitsorozatot kapunk.

Megjegyzés 1.1.4.: Több qubit lehet egymással összefonódott állapotban, amikor nem létezik olyan állapota a qubiteknek külön-külön, amik szorzata adná az együttes állapotot. Ilyen például az EPR-párnak nevezett két qubites $\frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)$ állapot.

Tekintsünk egy $\mathcal{U} \in \mathbb{C}^{2^n \times 2^n}$ unitér mátrixot. Ekkor ha $|\psi\rangle \in \mathbb{C}^{2^n}$ egy n qubites kvantumállapot, akkor $|\phi\rangle := \mathcal{U}|\psi\rangle$ szintén egy kvantumállapot. Egy kvantumalgoritmus unitér mátrix operátorok egymásutánjával és mérésekkel leírható. Egy és két qubitre ható mátrixokat tipikusan kapuknak szoktunk hívni, néhány elemi kaput definiálva előállíthatóak a komplexebb operátorok is - az elemi kapuk célja hogy egy kvantumalgoritmus kapu komplexitását mérhessük.

Néhány fontosabb kapu:

- $\mathcal{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $\mathcal{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\mathcal{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
- Forgatás: $\mathcal{R}_\phi = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}$
- Hadamard kapu: $\mathcal{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

Ha van egy $\mathcal{U} \in \mathbb{C}^{2^n \times 2^n}$ operátorunk, akkor controlled-$\mathcal{U}$ alatt $\begin{pmatrix} \mathcal{I}_n & 0 \\ 0 & \mathcal{U} \end{pmatrix}$ -t értjük, ahol $\mathcal{I}_n$ is $2^n \times 2^n$-es és a 0-k is az ugyanekkora csupanulla mátrixok.

1.2. Definíciók és jelölések

Néhány jelölést és definíciót vezetek be, ami a szakdolgozat során több helyen előfordul.

Használni fogom a braket-jelölést (ú.n. Dirac-notation) a kvantumállapotok leírására. Ket jelölés: $|\psi\rangle$, bra jelölés: $\langle\psi|$.

Jelölés 1.2.1.: A $\langle\psi|$ jelöléssel $|\psi\rangle$ adjungáltját fogjuk jelölni, az adjungáltat pedig kvantumos irodalomban †-dal jelöljük. Tehát $\langle\psi| = |\psi\rangle^\dagger$.

Megkülönböztetem a $|0\rangle \in \mathbb{C}^2$ vektort a $|\vec{0}\rangle$ vektortól, ami tetszőlegesen vagy megfelelően sok csupanulla állapot lesz, vagyis $|\vec{0}\rangle = |0\rangle \otimes ... \otimes |0\rangle$. (Bizonyos helyeken, ahol hangsúlyozni szeretném a dimenzióját kiírom, hogy $|0\rangle^n$.) A mátrixokat kalligrafikus betűtípussal, egyéb változókat tipikusan dőlt betűtípussal jelölöm.

A kvantumalgoritmusokat néhol illusztrálom kvantumáramkörrel is, ezeket az ábrákat úgy kell értelmezni, hogy egy vonal felel meg egy qubitnek és balról jobbra haladva alkalmazzuk rá a rárajzolt operátort. Illetve mind kvantumáramkörökön, mind algoritmusokban $meas(|\psi\rangle)$-vel jelölöm a méréseket.

A szakdolgozat során X diszkrét valószínűségi változó várhatóértékének becsléséről lesz szó. Valószínűségi mező alatt a szokásos $(\Omega, 2^\Omega, \mathbb{P})$ jelöléssel azt értem, hogy Ω a diszkrét eseménytér, 2^Ω a hatványhalmaza és minden $\omega \in \Omega$ eseményhez $\mathbb{P}[\omega]$ valószínűség tartozik.

Definíció 1.2.2.: (Kvantilis) Legyen X diszkrét valószínűségi változó. Egy valós $p \in [0, 1]$ esetén $Q_X(p)$ jelöli az X p-rendű kvantilisét, és a következőt értjük alatta:

$$Q_X(p) := sup\{x \in \mathbb{R} : \mathbb{P}[X \geq x] \geq p\}$$

Ha a szövegkörnyezetből X egyértelmű, $Q(p)$-vel jelölöm a p-rendű kvantilisét.

Kvantliseket mind a 2. mind a 3. fejezetben fogunk használni és mindkét helyen használni fogjuk Yassine Hamoudi eredményét ([Ham21] Theorem 4.3.4.), miszerint a kvantilis kvantumosan hatékonyan becsülhető.

Tétel 1.2.3. (Hamoudi): Legyen X diszkrét valószínűségi változó. Létezik kvantumalgoritmus, mely két valós értéket $p, \delta \in (0, 1)$ vár inputként és $Q(p)$ becslését, $\widetilde{Q}$-t adja $O\left(\frac{log(1/\delta)}{\sqrt{p}}\right)$ kvantum mintavétellel. Ahol $c > 1$ univerzális konstansra $1 - \delta$ valószínűséggel teljesül, hogy:

$$Q(p) \leq \widetilde{Q} \leq Q(cp)$$

1.3. Grover algoritmusa

A kvantum számítás egyik legismertebb algoritmusa a Grover algoritmus. Adott egy 2^n hosszú $0-1$ string, célunk hogy visszaadjuk egy indexét, melyen 1 áll:

Feladat: Adott $f : \{0,1,...,N-1\} \to \{0,1\}$, ahol $N = 2^n$ és $n \in \mathbb{Z}^+$. Adjunk meg olyan $i \in \{0,1,...,N-1\}$ -t, melyre $f(i) = 1$.

Azt mondjuk, hogy $f(i) = 1$ esetén i jelölt, különben jelöletlen. Klasszikus esetben nyílván $\Theta(N)$ lekérdezésre van szükségünk jelölt i találásához, viszont kvantumosan elég négyzetesen kevesebb $O(n) = O(\sqrt{N})$ is.

Legyen $\mathcal{R} := 2|0\rangle^n\langle 0|^n - \mathcal{I}_n$ és $\mathcal{O}_{\pm,f}$ ami $i \in \{0,1\}^n$ esetén a $|i\rangle\mathcal{H}|0\rangle \to |i\rangle\mathcal{H}|0\rangle$ és $|i\rangle\mathcal{H}|1\rangle \to (-1)^f|i\rangle\mathcal{H}|1\rangle$ leképezést csinálja. Továbbá legyen

$$\mathcal{G} := \mathcal{H}^{\otimes n}\mathcal{R}\mathcal{H}^{\otimes n}\mathcal{O}_{\pm,f}$$

ezt a $\mathcal{G}$-t Grover iterációnak fogjuk hívni.

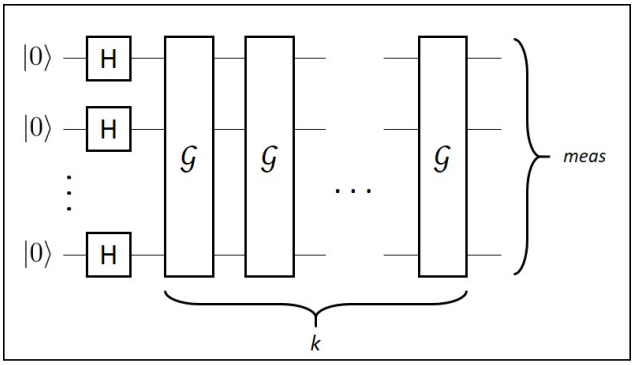
Algoritmus 1.3.1. (Grover): Az algoritmus során n qubittal fogunk dolgozni, kezdetben minden amplitúdó egyenlő lesz és néhány Grover iterációval folyamatosan növeljük a jelölt i-k amplitúdóját, végül a megmért számot válaszoljuk. Legyen t a jelöltek száma, ekkor:

1. Tegyünk n qubitet szuperpozícióba: $|U\rangle := \mathcal{H}^{\otimes n}|0\rangle^n$
2. $k = O\Big(\sqrt{\frac{N}{t}}\Big)$-szer alkalmazzunk Grover iterációt: $\big|\widetilde{G}\big\rangle := \mathcal{G}^k|U\rangle$
3. Mérjük meg a kapott kvantumállapotot a számítási bázisban: $i := meas\big(\big|\widetilde{G}\big\rangle\big)$

Az algoritmus elemzése

Tegyük fel hogy $0 < t < N$. Továbbá jelölje $|U\rangle$ a kezdőállapotot az algoritmus 1. pontja szerint és jelöljük $|G\rangle$, $|B\rangle$-vel a következő "jó" és "rossz" állapotokat:

$$|U\rangle = \sum_{i=0}^{N-1} \frac{1}{\sqrt{N}}|i\rangle \qquad |G\rangle := \sum_{i:f(i)=1} \frac{1}{\sqrt{t}}|i\rangle \qquad |B\rangle := \sum_{i:f(i)=0} \frac{1}{\sqrt{N-t}}|i\rangle$$



A célunk, hogy a jó állapotba jussunk, hiszen $|G\rangle$-t megmérve 1-valószínűséggel olyan i-t kapunk, melyre $f(i) = 1$.

Tekintsük azt a 2-dimenziós alteret, amit $|G\rangle$ és $|B\rangle$ feszít. Legyen θ a $|U\rangle$ és $|B\rangle$ által bezárt szög, ekkor $\theta = arcsin\Big(\sqrt{\frac{t}{N}}\Big)$, ugyanis:

$$sin(\theta) = cos\Big(\frac{\pi}{2} - \theta\Big) = \langle G|U\rangle = \sum_{i:f(i)=1} \frac{1}{\sqrt{t}} \cdot \frac{1}{\sqrt{N}} = \sqrt{\frac{t}{N}}$$

Állítás 1.3.2.: Ezen síkon $\mathcal{G}$ két tükrözés egymásutánja, mely egy 2θ szöggel való elforgatást eredményez. Továbbá $\mathcal{G}^k|U\rangle$ és $|B\rangle$ által bezárt szög $(2k+1)\theta$.

Biz.: Az $\mathcal{O}_{f,\pm}$ fázisorákulumról tudjuk, hogy azon $|i\rangle$-ket, melyekre $f(i) = 1$ negálja, míg $f(i) = 0$ esetén helybenhagyja, tehát $\mathcal{O}_{f,\pm}$ egy tükrözés $|B\rangle$-n keresztül. Mivel

$$\begin{aligned}\mathcal{H}^{\otimes n}\mathcal{R}\mathcal{H}^{\otimes n} = \mathcal{H}^{\otimes n}(2|0\rangle^n\langle 0|^n - \mathcal{I}_n)\mathcal{H}^{\otimes n} = \mathcal{H}^{\otimes n}2|0\rangle^n\langle 0|^n\mathcal{H}^{\otimes n} - \mathcal{H}^{\otimes n}\mathcal{I}_n\mathcal{H}^{\otimes n} = \\ = 2(\mathcal{H}^{\otimes n}|0\rangle^n)(\langle 0|^n\mathcal{H}^{\otimes n}) - (\mathcal{H}^{\otimes n})^2 = 2|U\rangle\langle U| - \mathcal{I}_n\end{aligned}$$

ezért $\mathcal{H}^{\otimes n}\mathcal{R}\mathcal{H}^{\otimes n}$ tükrözés $|U\rangle$-n keresztül.

Tehát, ha a $|B\rangle$-vel bezárt szög kezdetben α, akkor $|B\rangle$-re vett tükrözés után (ugyanolyan körüljárás szerint) $-\alpha$, majd $|U\rangle$-ra vett tükrözés után $2\theta + \alpha$. Így valóban egy Grover iteráció alkalmazása egy 2θ-s elforgatás. Mivel kezdetben $|U\rangle$ és $|B\rangle$ által bezárt szög θ, így k Grover iteráció után a $\mathcal{G}^k|U\rangle$ és $|B\rangle$ által bezárt szög valóban $(2k+1)\theta$.

Hogyan válasszuk meg k értékét?

Azt szeretnénk, hogy k Grover iteráció után $|G\rangle$-ben legyünk. Szeretnénk tehát, hogy fennálljon:

$$(2k+1)\theta = \frac{\pi}{2}$$

Tegyük fel, hogy t-t ismerjük és $t << N$, ekkor

$$(2k+1)\theta = (2k+1)\cdot arcsin\Big(\sqrt{\frac{t}{N}}\Big) \approx (2k+1)\sqrt{\frac{t}{N}}$$

Tehát ha $k \approx \frac{\pi}{4}\sqrt{\frac{N}{t}} - \frac{1}{2}$, akkor

$$(2k+1)\theta \approx \left(2\Big(\frac{\pi}{4}\sqrt{\frac{N}{t}} - \frac{1}{2}\Big) + 1\right)\sqrt{\frac{t}{N}} = \frac{\pi}{2}$$

Megjegyzés 1.3.3.: Tetszőleges ismert t-re módosítható az algoritmus, hogy pontosan $\frac{\pi}{2}$-be érkezzünk $k \in \mathbb{Z}$ iterációval. (A kezdeti θ-t lehet megfelelően módosítani és a k értéket meghatározni.)

Megjegyzés 1.3.4.: Ha t nem ismert, akkor is meg lehet csinálni úgy, hogy a lekérdezések számának várható értéke $O\Big(\sqrt{\frac{N}{t}}\Big)$ legyen (exponenciálisan növekvő k tippekkel).

Amplitúdó amplifikáció

A Grover algoritmus ötlete általánosabban is alkalmazható. Tekintsük most a következő állapotot $\sqrt{1-p}|\psi_0\rangle + \sqrt{p}|\psi_1\rangle$, ahol $|\psi_0\rangle$ és $|\psi_1\rangle$ normalizált, egymással ortogonális helyzetben lévő álllapotok (hasonlóan a Grovernél vett $|G\rangle, |B\rangle$ -hez). Ha most $|\psi_1\rangle$-et tekintjük "jó" állapotnak és $|\psi_0\rangle$-t rossznak, akkor növelni szeretnénk $|\psi_1\rangle$ amplitúdóját. Legyen tehát adott most egy $\mathcal{U}_p$ kapu, melyre

$$\mathcal{U}_p|0\rangle^n \longrightarrow \sqrt{1-p}|\psi_0\rangle + \sqrt{p}|\psi_1\rangle$$

Ekkor egy amplitúdó amplifikáció a következőképp nézhet ki: kiindulva $|U\rangle := \mathcal{U}_p|0\rangle^n$-ből ismételgessük, hogy tükrözünk $|\psi_0\rangle$-on, majd $|U\rangle$-n keresztül.

1.4. Kvantum Fourier transzformáció és fázisbecslés

A Fourier transzformáció egy széles körben használt eszköz, alkalmazzák például jelfeldolgozásban, tömörítéshez (jpeg) vagy polinomszorzáshoz. Ennek egyik fő oka, hogy klasszikus Fourier transzformáció esetén alkalmazható a gyors Fourier transzformáció, ami a triviális $O(N^2)$ helyett $O(NlogN)$ műveletet igényel. Ha $N = 2^n$, a kvantum Foruier transzformáció esetén ez n qubiten $O(n^2)$ kapuval végrehajtható, azaz exponenciális a gyorsítás. Viszont fontos megjegyezni, hogy a kvantum Foruier transzformáció esetén a kvantumállapotunk amplitúdóiként kapjuk meg az eredményt.

Legyen

$$\mathcal{F}_N = \frac{1}{\sqrt{N}} \begin{pmatrix} & \vdots & \\ \cdots & \omega_N^{jk} & \cdots \\ & \vdots & \end{pmatrix} \qquad \text{ahol } \omega_N = e^{\frac{2\pi i}{N}} \text{ primitív egységgyök és } j,k = 0,1,...,N-1.$$

Például:

$$\mathcal{F}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \mathcal{H} \qquad \mathcal{F}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & e^{\frac{2\pi i}{3}} & e^{\frac{4\pi i}{3}} \\ 1 & e^{\frac{4\pi i}{3}} & e^{\frac{2\pi i}{3}} \end{pmatrix}$$

Egy $v \in \mathbb{R}^N$ vektor Fourier transzformáltja $\mathcal{F}_N v$. Kvantumos esetben arról van szó, hogy van egy kvantumállapotunk, vagyis vektorok szuperpozíciója valamilyen amplitúdókkal. A kvantum Fourier transzformáció tehát ezen amplitúdókra $\mathcal{F}_N$ alkalmazása, így egy másik kvantumállapotot eredményezve.

Állítás 1.4.1.: $\mathcal{F}_N$ unitér mátrix (azaz $\mathcal{F}_N$ tekinthető egy kvantumkapunak).

Biz.: Tekintsük az $\mathcal{F}_N$ $i.$ és $j.$ oszlopát, ekkor

$$\big(\langle i|\mathcal{F}_N^\dagger\big)\big((\mathcal{F}_N)|j\rangle\big) = \sum_{k=0}^{N-1} \Big(\frac{1}{\sqrt{N}}(\omega_N^{ki})^{-1}\Big)\Big(\frac{1}{\sqrt{N}}\omega_N^{kj}\Big) = \frac{1}{N}\sum_{k=0}^{N-1} \omega_N^{(j-i)k} = \begin{cases} 1 & \text{ha } i = j \\ 0 & \text{különben} \end{cases}$$

Tehát valóban, $\mathcal{F}_N$ oszlopai ortonormáltak.

Hatékony kvantumimplementáció:
Szükségünk van tehát a $|k\rangle \to \mathcal{F}_N|j\rangle = \sum_{j=0}^{N-1} \omega_N^{jk}|j\rangle$ kapura. Ehhez kihasználjuk, hogy ha j bináris alakja $j_1 j_2 ... j_n$, akkor $j = \sum_{\ell=1}^{n} j_\ell \cdot 2^{n-\ell}$, és így:

$$\mathcal{F}_N|k\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{\frac{2\pi i j k}{N}} |j_1 j_2 ... j_n\rangle = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{N-1} \prod_{\ell=1}^{n} e^{2\pi i \cdot j_\ell \cdot \frac{k}{2^\ell}} |j_1 j_2 ... j_n\rangle = \bigotimes_{\ell=1}^{n} \frac{1}{\sqrt{2}} \Big(|0\rangle + e^{2\pi i \frac{k}{2^\ell}} |1\rangle\Big)$$

Fontos megfigyelés, hogy ha tekintjük $e^{2\pi i \frac{k}{2^\ell}}$ -et, akkor $\frac{k}{2^\ell}$ egészrészétől eltekinthetünk (hisz a kitevőben $2\pi i$ egészszerese 1). Ez a tulajdonság adja alapját, a hatékony kvantumimplementációnak.

A kvantum áramkört $\mathcal{H}$ és controlled-$\mathcal{R}_s$ kapukkal tudjuk létrehozni, ahol $\mathcal{R}_s := \begin{pmatrix} 1 & 0 \\ 0 & e^{\frac{2\pi i}{2^s}} \end{pmatrix}$. Nézzük meg például $n = 2$ esetben hogy néz ki $\mathcal{F}_4$:

$$\mathcal{F}_4|k\rangle = \mathcal{F}_4|k_1 k_2\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i \cdot 0.k_2}|1\rangle\big) \otimes \frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i \cdot 0.k_1 k_2}|1\rangle\big)$$

Ekkor egyrészt

$$\frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i \cdot 0.k_2}|1\rangle\big) = \mathcal{H}|k_2\rangle$$

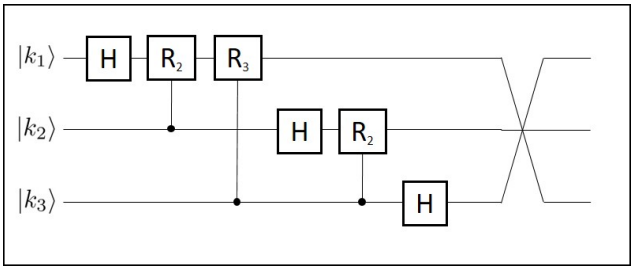
másrészt

$$\frac{1}{\sqrt{2}}\big(|0\rangle e^{2\pi i \cdot 0.k_1 k_2}|1\rangle\big) = \frac{1}{\sqrt{2}}\big(|0\rangle + e^{2\pi i \cdot 0.k_1} e^{2\pi i \cdot 0.0 k_2}|1\rangle\big)$$

hasonló az előzőhöz, itt $\mathcal{H}|k_1\rangle$ után még k_2-től függően vagy be kell szorozzuk $|1\rangle$-et $e^{\frac{\pi i}{2}}$-vel, vagy sem. Ezt controlled-$\mathcal{R}_2$ kapuval tudjuk megtenni.

Általánosan is hasonlóan járhatunk el, n qubit esetén $|k_1\rangle$-re $\mathcal{H}$ és controlled $\mathcal{R}_2$, $\mathcal{R}_3$, ..., $\mathcal{R}_n$ kapukat kell alkalmazni, majd $|k_2\rangle$-re $\mathcal{H}$ és $\mathcal{R}_2$, ..., $\mathcal{R}_{n-1}$ és így tovább (lásd ábra $n = 3$ esetben). A kapuk száma pedig $n + (n-1) + ... + 1 = \frac{n(n+1)}{2} = O(n^2)$ lesz.

Alkalmazás fázisbecslésre:
A kvantum Fourier transzformáció több kvantumalgoritmusban játszik fontos szerepet, mint például

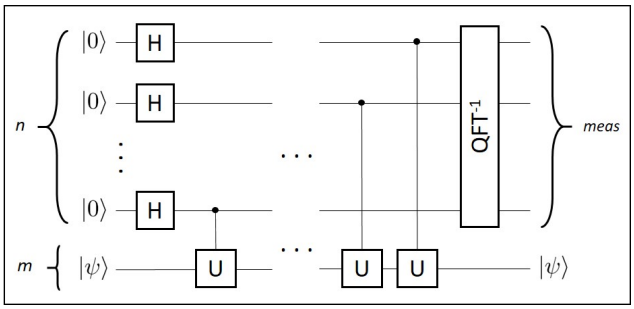


a 3. fejezetben fogjuk látni, többdimenziós valószínűségi változó várhatóértékének hatékony becslésében is. Most azt nézzük meg hogyan használható fázisbecslés meghatározására. A fázisbecslés hatékonysága szintén fontos lesz, ezen fog alapulni a 2. fejezetben vizsgált probléma eredménye.

Legyen adott egy m qubitra alkalmazható $\mathcal{U}$ kapu. Mivel $\mathcal{U} \in \mathbb{C}^{2^m \times 2^m}$ mátrix, beszélhetünk $|\psi\rangle$ sajátvektoráról illetve a hozzátartozó λ sajátértékről: $\mathcal{U}|\psi\rangle = \lambda|\psi\rangle$. Viszont azt is tudjuk, hogy $\mathcal{U}$ unitér, így $\lambda = e^{2\pi i\varphi}$ valamely $0 \leq \varphi < 1$ -re. Kvantumszámításban $|\psi\rangle$-t $\mathcal{U}$ sajátállapotának, φ-t a hozzátartozó sajátfázisának szoktuk nevezni, a feladatunk most φ hatékony becslése.

Állítás 1.4.2.: Legyen adott $\mathcal{U}$ és tegyük fel, hogy controlled-$\mathcal{U}$ kaput is elő tudjuk állítani (ez nem triviális tetszőleges kapura). Ekkor egy $|\psi\rangle$ sajátállapotát meg tudjuk határozni n tizedesjegy pontossággal n többlet qubitot használva és anélkül hogy elpusztítanánk $|\psi\rangle$ kvantumállapotot.

Legyen $\varphi = 0.\varphi_1\varphi_2...\varphi_n...$ tizedesalakban felírva. Vegyünk n többlet qubitot $|0\rangle$ alapállapotban, ezekre alkalmazzunk egy QFT-t (kvantum Fourier transzformációt), megfigyelhetjük hogy ez ekkor ugyanaz mintha mindegyik qubitra egy $\mathcal{H}$ kaput alkalmaznánk. Ezt követően alkalmazzunk n controlled-$\mathcal{U}$ kapukat, aminek a controll qubitjai sorra az n db segéd (többlet) qubit és a targetje a $|\psi\rangle$-hez tartozó regiszter. Végül alkalmazzunk egy inverz QFT-t az n qubithoz tartozó regiszterre és ezt mérjük is meg. Ekkor a kívánt eredményt kapjuk:



$$|0^n\rangle|\psi\rangle \xrightarrow{\mathcal{H}^n\otimes\mathcal{I}} \frac{1}{\sqrt{2^n}}\sum_{j=0}^{N-1}|j\rangle|\psi\rangle \xrightarrow{(\text{contr}-\mathcal{U})^n} \frac{1}{\sqrt{2^n}}\sum_{j=0}^{N-1}|j\rangle\mathcal{U}^j|\psi\rangle = \frac{1}{\sqrt{2^n}}\sum_{j=0}^{N-1}e^{2\pi i\frac{\varphi_1\ldots\varphi_n}{N}j}|j\rangle|\psi\rangle =$$

$$= \frac{1}{\sqrt{2^n}}\sum_{j=0}^{N-1}\omega_N^{(\varphi_1\ldots\varphi_n)j}|j\rangle|\psi\rangle \xrightarrow{QFT_N^{-1}} |\varphi_1\ldots\varphi_n\rangle|\psi\rangle$$

2. Várható érték becslése kvatumosan

Ebben a fejezetben egydimenziós diszkrét valós valószínűségi változó várható értékének a becsléséről lesz szó, az eloszlása szerint vett minták alapján. A cél az lesz, hogy minél kevesebb mintavétellel minél pontosabban tudjuk ezt megtenni. Látni fogjuk, hogy a problémára kétféleképpen is tekinthetünk, egyrészt úgy hogy rögzített pontosság eléréséhez mennyi mintát kell venni, másrészt úgy hogy a minta rögzített számossága esetén milyen pontosságot tudunk elérni. A problémát kvantum környezetben azért érdemes vizsgálni, mert a Monte Carlo módszereken alapuló algoritmusok alapját képezik, így kvantumos gyorsítás ezen algoritmusok gyorsításához vezet, erre fogunk látni egy pénzügyi példát a 4. fejezetben.

Robin Kothari és Ryan O'Donnell 2023-ban publikált eredménye [KO23] egydimenziós valószínűségi változó esetén négyzetes gyorsítást ér el - és ez optimális. Az algoritmusuk egy fázisbecslés lesz egy speciális, a Grover-algoritmusban használthoz hasonló kapura, viszont komplex fázisokkal. Tételüket visszavezetik egy döntési problémára, amit ezzel az algoritmussal tudnak megoldani. A fejezet elején a problémát ismertetem klasszikus esetben, majd a kvantumos mintavételről írok, végül ismertetem a Kothari-O'Donnell eredményt.

2.1. A várható érték becslési probléma

Legyen $X : \Omega \to \mathbb{R}$ véges diszkrét valós valószínűségi változó a $(\Omega, 2^\Omega, \mathbb{P})$ valószínűségi mezőn. Tekintsük az X várható értékét:

$$\mu := \mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}[\omega] X(\omega)$$

ezt szeretnénk megbecsülni. Gondoljunk most úgy X-re mint egy feketedobozra, amiből tudunk néhány mintát venni az eloszlása szerint, de csak a kapott mintákat használhatjuk fel arra, hogy becsüljünk. Jelöljük a becslést $\tilde{\mu}$-mal, a cél hogy nagy valószínűséggel közel legyünk a valódi várható értékhez. (A feketedoboz szemlélet túl erős megkötés, a 2.3. alfejezetben részletesebben írok a mintavételről.)

Definíció 2.1.1.: $\tilde{\mu}$ a várhatóérték ε-becslése, ha fennáll:

$$\mathbb{P}\big[|\tilde{\mu} - \mu| > \varepsilon\big] \leq \frac{1}{3}$$

Megjegyzés 2.1.2.: Elérhető, hogy a definícióban szereplő $1/3$ helyett δ álljon (ahol $0 < \delta < 1/3$). Ha adott egy a definíciónak eleget tevő eljárás, akkor annak $O\big(log(1/\delta)\big)$ ismétlésével, majd a kapott becslések mediánját véve a valószínűség δ-ra csökken. Ennek hátterében az áll, hogy ahhoz hogy a medián kívül essen $[\mu - \varepsilon, \mu + \varepsilon]$-on, a becslések többségének is kívül kell esnie.

A várható érték becslési problémát a mintavételek száma szerint szeretnénk mérni, ennek a nagyságrendjét fogjuk összehasonlítani klasszikus és kvantumos esetben. A probléma két paraméter relációjáról szól: hogy mekkora pontosságot érünk el és hogy hány mintát veszünk hozzá. Az alapján, hogy melyiket rögzítjük, kétféleképp fogalmazhatjuk meg a problémát:

Probléma I.): Legyen adott $\varepsilon > 0$. Mekkora az a legkisebb $n_\varepsilon \in \mathbb{Z}$, melyre létezik algoritmus ami n_ε mintát vesz és ε-becslést ad a várható értékre?

Probléma II.): Legyen adott n. Mekkora az a legkisebb $\varepsilon(n)$, melyre létezik algoritmus ami n mintát vesz és $\varepsilon(n)$-becslést ad a várható értékre?

2.2. Becslés klasszikus esetben

A klasszikus algoritmusok közül tekintsük a legkézenfekvőbbet a **II.)**-es problémára:

1. Vegyünk n mintát X-ből: $X_1, ..., X_n$.
2. $\tilde{\mu} := \frac{\sum X_i}{n}$

Állítás 2.2.1.: Ekkor $\varepsilon(n) = O\Big(\frac{\sigma}{\sqrt{n}}\Big)$, ahol σ az X szórása.

Biz.: Most $\tilde{\mu}$ a mintáktól függő valószínűségi változó, aminek szórása $\sigma(\tilde{\mu}) = \sigma(X)/\sqrt{n}$, ugyanis:

$$\sigma^2\big(\tilde{\mu}\big) = \sigma^2\Big(\frac{1}{n}\sum_{i=1}^{n} X_i\Big) = \Big(\frac{1}{n}\Big)^2 \sum_{i=1}^{n} \sigma^2\big(X_i\big) = \frac{1}{n}\sigma^2\big(X\big)$$

Alkalmazva a Csebisev egyenlőtlenséget $\tilde{\mu}$-ra $k > 0$ esetén azt kapjuk, hogy:

$$\mathbb{P}\big[|\tilde{\mu} - \mu| > k\big] \leq \frac{\sigma^2}{nk^2}$$

Tehát $k = \sqrt{3}\frac{\sigma}{\sqrt{n}}$ választással:

$$\mathbb{P}\Big[|\tilde{\mu} - \mu| > \sqrt{3}\frac{\sigma}{\sqrt{n}}\Big] \leq \frac{\sigma^2}{n \cdot 3\frac{\sigma^2}{n}} = \frac{1}{3}$$

Klasszikus esetben ez optimális is, [LM19] Theorem 1. szerint:

Tétel 2.2.2. (Lugosi-Mendelson): Legyen $n > 5$ egész és $2e^{-n/4} < \delta < 1/2$. Tetszőleges (klasszikus) n mintát használó várható érték becslő algoritmusra létezik eloszlás μ várható értékkel és σ szórással, melyre:

$$\mathbb{P}\left[|\tilde{\mu} - \mu| > \sigma\sqrt{\frac{log(1/\delta)}{n}}\right] \geq \delta$$

Megjegyzés 2.2.3.: Ez a tétel az erősebb δ-ás verzióval van kimondva, erre röviden kitértünk a 2.1.2. megjegyzésben is. A szakdolgozatom során a konstans $(1/3)$-os verziót használom, de fontos megjegyezni hogy ha δ-t is inputként kapjuk, akkor a mintavételek száma függ $1/\delta$-tól. Ilyen esetben az angol irodalomban "median of means"-nek nevezett algoritmusra hasonlóan a Csebisev egyenlőtlenségből $O\Big(\sigma\sqrt{\frac{log(1/\delta)}{n}}\Big)$ jön ki. A median of means alatt azt értjük fix n esetén, hogy kb. $log(1/\delta)$ egyenlő részre osztjuk n-et, minden részre számolunk átlagot és végül ezek mediánját válaszoljuk.

Láttuk tehát, hogy klasszikus esetben fix n esetén $O(\frac{\sigma}{\sqrt{n}})$-becslést tudunk elérni és ez a nagyságrend optimális. A teljesség kedvéért fogalmazzuk meg **I.)**-es probléma szerint is:

Következmény 2.2.4.: Adott $\varepsilon > 0$. Bármely klasszikus algoritmus, ami ε-becslést ad a várható értékre $\Omega\big(\frac{\sigma^2}{\varepsilon^2}\big)$ mintát vesz. Továbbá létezik klasszikus algoritmus, ami $O\big(\frac{\sigma^2}{\varepsilon^2}\big)$ mintavétellel elér ε-becslést.

2.3. Mintavétel kvantumos esetben

Ebben a részben megfoglamazzuk, mit értünk mintavétel alatt pontosan. A korábban írt feketedobozos szemlélet ugyan hasznos, hogy miként gondolhatunk rá klasszikus esetben, azonban túl erős megszorítás - mivel a lekérdezések számában mérjük az algoritmusokat, így természetesen kvantumos környezetben nem tudnánk semmi többletet elérni. Fogalmazzuk újra a mintavételt a következőképp:

Definíció 2.3.1.: Legyen továbbra is adott az $(\Omega, 2^{\Omega}, \mathbb{P})$ valószínűségi mezőnk. Legyen $\mathcal{P}$ olyan kvantum kapu, mely (megfelelő mennyiségű) kezdeti $|0\rangle$ állapotban lévő qubitet a következő állapotba visz:

$$\mathcal{P}|\vec{0}\rangle = \sum_{\omega \in \Omega} \sqrt{\mathbb{P}[\omega]}|\omega\rangle|garbage_{\omega}\rangle$$

Ekkor $\mathcal{P}$-t disztribúciós orákulumnak hívjuk, $|garbage_{\omega}\rangle$ pedig normalizált "szemét" vektor.

Megjegyzés 2.3.2.: Egy $|garbage_{\omega}\rangle$ mentes definíció sokkal kötöttebb lenne, a legtöbb esetben nem elvárható hogy tudjunk ilyen disztribúciós orákulumot készíteni. Klasszikus mintavételt szeretnénk tudni az általánosabb kvantumos mintavétellé alakítani. Például klasszikus esetben nem jelent különösebb gondot, hogy egy adott gráf lehetséges csúcscímkézései közül uniform eloszlás szerint vegyünk mintát. De, ha ismernénk disztribúciós orákulumot, ami ezek garbage-mentes szuperpozícióját adná, akkor megtudnánk hatékonyan oldani a gráf izomorfizmus problémát (lásd [OT01]).

Definíció 2.3.3.: Legyen adott egy $\mathcal{B}_X$ kapu, ami egy $|\omega\rangle$ állapothoz rendeli $X(\omega)$-t bináris alakban:

$$\mathcal{B}_X|\omega\rangle|\vec{0}\rangle = |\omega\rangle|X(\omega)\rangle$$

Ekkor $\mathcal{B}_X$-et bináris orákulumnak hívjuk. (Általánosan tetszőleges f függvényhez hasonlóan definiálhatunk $\mathcal{B}_f$ bináris orákulumot.)

Kvantumos esetben tekintsük egy mintavételnek azt, hogy vagy $\mathcal{P}$-t, vagy $\mathcal{B}_X$-et használjuk egyszer. Ezzel a megközelítéssel általánosítottuk a klasszikus mintavételt, hiszen egy mintát kapunk

először egy $\mathcal{P}$ és egy $\mathcal{B}_X$ hívással

$$|\vec{0}\rangle|\vec{0}\rangle \xrightarrow{\mathcal{P}} \sum_{\omega\in\Omega} \sqrt{\mathbb{P}[\omega]}|\omega\rangle|garbage_\omega\rangle|\vec{0}\rangle \xrightarrow{\mathcal{B}_X} \sum_{\omega\in\Omega} \sqrt{\mathbb{P}[\omega]}|\omega\rangle|garbage_\omega\rangle|X(\omega)\rangle$$

végül pedig az $|X(\omega)\rangle$ regiszter megmérésével.

2.4. Becslés kvantumos esetben

Most rátérünk Kothari és O'Donnell [KO23] módszerére, ami optimális négyzetes gyorsítás a klasszikus esethez képest. A módszerük egy döntési problémára való visszavezetés lesz illetve egy (a Grover algoritmusban használt kapuhoz hasonló) kapu sajátfázisának becslése.

Legyen most adott az X valószínűségi változó és a hozzá tartozó $\mathcal{P}$ és $\mathcal{B}_X$ orákulumok. Kvantumos esetben négyzetes gyorsításnál jobbat nem várhatunk, [Ham21] Theorem 4.6.2. szerint:

Tétel 2.4.1. (Hamoudi): Legyen $n > 1$ és $0 < \delta < 1$, úgy hogy $n \geq 2log(1/\delta)$. Legyen továbbá $\sigma > 0$ rögzített és tekintsük az X-ek azon $\mathbf{P}_\sigma$ családját, aminek a szórása σ. Ekkor $\Omega(n)$ mintavétel amit bármely kvantumalgoritmusnak végre kell hajtania, hogy minden $X \in \mathbf{P}_\sigma$ -ra olyan $\tilde{\mu}$ várhatóérték becslést adjon, melyre:

$$\mathbb{P}\left[|\tilde{\mu} - \mu| > \frac{\sigma log(1/\delta)}{n}\right] \leq \frac{1}{\delta}$$

Hasonlóan, ha a 2.1.1. definícióban szereplő 1/3-os változatos becslést tekintjük, akkor $\frac{\sigma log(1/\delta)}{n}$-et $O(\frac{\sigma}{n})$-re cserélhetjük.

Tétel 2.4.2. (Kothari-O'Donnell): Létezik kvantumalgoritmus, amely $O(n)$ mintát használ és az outputja $\tilde{\mu}$, melyre:

$$\mathbb{P}\Big[|\tilde{\mu} - \mu| > \frac{\sigma}{n}\Big] \leq \frac{1}{3}$$

2.4.1. Visszavezetés döntési feladatra

A szerzők egy döntési problémát fognak megoldani hatékonyan kvantumalgoritmussal és belátják, hogy az indukálja a 2.4.2.-es tételt. A visszavezetés több lépésben történik, viszont több tanulságos trükköt alkalmaznak, mint például a loglog módszer, az úgynevezett successive-halving, változó skálázás, illetve változó kvantilis szerinti megszorítása.

Tekintsük a következő döntési problémát:

1) IsMeanSmall(X, ε, c):

- A bemeneti X valós valószínűségi változó, $0 < \varepsilon \leq 1$ és $0 < c < 1$.
- Teljesül, hogy $\mathbb{E}[X^2] \leq 1$, továbbá legalább az egyik fennáll i) $|\mu| \leq c\varepsilon$ vagy ii) $\varepsilon \leq |\mu| \leq 2\varepsilon$.
- RETURN i) vagy ii) áll fenn.

(Ha a feltételek nem állnak fenn, akkor bármelyik válasz elfogadott.)

Állítás 2.4.1.1.: Ha az IsMeanSmall(X, ε, c) döntési feladatra létezik kvantumalgoritmus ami $O(1/\varepsilon)$ mintát használ, akkor a 2.4.2.-es tétel igaz.

Biz.: A visszavezetést 8 lépésben fogjuk megtenni. A visszavezetések során használjuk a jelölést $s := \sqrt{\mathbb{E}[X^2]}$, μ az X várhatóértékét, σ a szórását fogja jelölni. Ha a bizonyítás során új X', X'' változókat vezetünk be, akkor értelem szerűen $\mu', \mu'', \sigma', \sigma'', s', s''$ a megfelelő változóhoz tartozó értékeket jelölik.

2) IsEstimateClose($X, \varepsilon, \tilde{\mu}, c$):

- X valós valószínűségi változó, $-1 \leq \tilde{\mu} \leq 1$, $0 < \varepsilon \leq 1$ és $0 < c < 1$.
- $s \leq 1$ és legalább az egyik fennáll i) $|\tilde{\mu} - \mu| \leq c\varepsilon$ vagy ii) $\varepsilon \leq |\tilde{\mu} - \mu| \leq 2\varepsilon$.
- RETURN i) vagy ii) áll fenn.

Lemma 2.4.1.2.: Ha 1) megoldható $O(1/\varepsilon)$ mintával, akkor 2) is megoldható $O(1/\varepsilon)$ mintával.

Legyen $X' := \frac{X-\tilde{\mu}}{2}$. Ekkor felhasználva, hogy $(a-b)^2 \leq 2a^2 + 2b^2$:

$$(s')^2 = \mathbb{E}[(X')^2] = \frac{1}{4}\mathbb{E}\Big[\big(X-\tilde{\mu}\big)^2\Big] \leq \frac{1}{4}\mathbb{E}\Big[2X^2 + 2\tilde{\mu}^2\Big] = \frac{2s^2 + 2\tilde{\mu}}{4} \leq \frac{2\cdot 1 + 2\cdot 1}{4} = 1$$

Mivel $s' \leq 1$ teljesül, meghívhatjuk az 1)-es problémát IsMeanSmall(X', ε, c).

3) EstimateForSmallSecondMoment(X, ε):

- X valós valószínűségi változó és $0 < \varepsilon \leq 1$.
- Teljesül, hogy $s \leq 1$.
- RETURN $\tilde{\mu}$, amire $|\tilde{\mu} - \mu| \leq \varepsilon$.

Lemma 2.4.1.3.: Ha 2) megoldható $O(1/\varepsilon)$ mintával, akkor 3) is megoldható $O(1/\varepsilon)$ mintával.

A módszer egyfajta bináris keresés lesz. Az első lépésben tekintsük az $I_1 = [-1, 1]$ intervallumot, majd minden lépésben egy 2)-es hívással szükítsük ezt az intervallumot konstans faktorral, úgy hogy μ továbbra is beleessen. Mivel 2)-t exponenciálisan csökkenő ε értékekkel fogjuk hívni, meg fog maradni az $O(1/\varepsilon)$-os nagyságrendű mintavétel.

Legyen $0 < c < 1$ fix és minden j. lépésben határozzuk meg a_j, b_j-t és legyenek

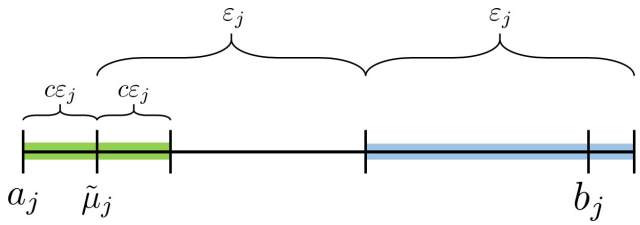
$$I_j := [a_j, b_j] \qquad \varepsilon_j := \frac{|I_j|}{2} \qquad \tilde{\mu}_j := a_j + c\varepsilon_j$$

Kezdetben $a_1 := -1$ és $b_1 := 1$ jó választás, hiszen $s \leq 1$-ből következik, hogy

$$\mu^2 = \mathbb{E}[X]^2 \leq \mathbb{E}[X^2] = s^2 \leq 1$$

A j. lépésben hívjuk meg IsEstimateClose$(X, \varepsilon_j, \tilde{\mu}_j, c)$-t, ami visszaadja hogy i) $|\tilde{\mu}_j - \mu| \leq c\varepsilon$ vagy ii) $\varepsilon \leq |\tilde{\mu}_j - \mu| \leq 2\varepsilon$, ekkor a következő esetek közül pontosan egy áll fenn:

a) $|\tilde{\mu}_j - \mu| \leq c\varepsilon$ és i) választ kaptunk.
b) $\varepsilon \leq |\tilde{\mu}_j - \mu| \leq 2\varepsilon$ és ii) választ kaptunk.
c) $|\tilde{\mu}_j - \mu| \not\leq c\varepsilon$ és $\varepsilon \not\leq |\tilde{\mu}_j - \mu| \not\leq 2\varepsilon$, így vagy i) vagy ii) választ kaptunk.



- Ha i) választ kaptunk, azt biztosan tudjuk, hogy $\mu \notin [a_j, \tilde{\mu}_j + c\varepsilon_j]$ (ábrán zöld) legyen tehát $a_{j+1} := a_j + 2c\varepsilon_j$ és $b_{j+1} := b_j$.
- Ha ii) választ kaptunk, akkor viszont azt tudjuk, hogy $\mu \notin [\tilde{\mu}_j + \varepsilon_j, b_j]$ (ábrán kék) legyen tehát $a_{j+1} := a_j$ és $b_{j+1} := a_j + (c+1)\varepsilon_j$

Amint az aktuális intervallum hossza $\leq \varepsilon$-ra csökken megállhatunk, ekkor $\tilde{\mu}_j$ jó becslés lesz. Minden lépésben az intervallum hossza legalább $(1-c')$-szeresére csökken valamely $0 < c' < 1$ értékre - hiszen az első esetben $|I_{j+1}| = (1-c)|I_j|$, a másodikban $|I_{j+1}| = \frac{1+c}{2}|I_j|$.

A mintavétel pedig továbbra is $O(1/\varepsilon)$, hiszen ha a t. lépésben állunk meg, akkor összesen $O(1/\varepsilon_1) + ... + O(1/\varepsilon_t) = O(1/\varepsilon_1 + ... + 1/\varepsilon_t)$ a mintavételek száma, amire

$$\frac{1}{\varepsilon_1} + ... + \frac{1}{\varepsilon_t} = 1 + \left(\frac{1}{1-c'}\right) + \left(\frac{1}{1-c'}\right)^2 + ... + \left(\frac{1}{1-c'}\right)^{t-1} = \frac{\left(\frac{1}{1-c'}\right)^t - 1}{\frac{1}{1-c'} - 1} \approx \frac{1}{(1-c')^t} \approx \frac{1}{\varepsilon}$$

4) EstimateForBernoulli(X, n):

- $X \in \{0,1\}$ valószínűségi változó és $n \in \mathbb{Z}^+$.
- RETURN $\tilde{\mu}$, amire $|\tilde{\mu} - \mu| \leq \frac{\sigma}{n}$.

Lemma 2.4.1.4.: Ha 3) megoldható $O(1/\varepsilon)$ mintával, akkor 4) megoldható $O(n)$ mintával.

Legyen $p := \mathbb{P}[X = 1]$, ekkor $\mu = p$ és $\sigma = \sqrt{p(1-p)}$. Feltehető, hogy $p \leq \frac{3}{4}$, ugyanis 3) hívásával, ha $\hat{\varepsilon} := \frac{1}{4}$, akkor $O(1)$ mintával legyen $\hat{\mu} :=$ EstimateForSmallSecondMoment$(X, \hat{\varepsilon})$. És ha $\hat{\mu} \leq \frac{1}{2}$, akkor p valóban nem lehet nagyobb $\frac{3}{4}$-nél. Ha viszont $\hat{\mu} > \frac{1}{2}$, akkor vegyük $X' := 1 - X$ -et és azzal számoljunk tovább (ekkor σ változatlan) a végén pedig ha $\tilde{\mu}$-t kapunk, akkor $1 - \tilde{\mu}$ -vel térjünk vissza.

Legyen $P := 3/4$ és folyamatosan tartsuk fenn, hogy $p \leq P$.

1. **HA** $P \leq \frac{1}{4n^2}$ **STOP**
2. $X' := \frac{X}{\sqrt{P}}$, $\varepsilon' := \frac{\sqrt{P}}{2}$.
3. Hívjuk meg 3)-at: $\tilde{\mu}' :=$ EstimateForSmallSecondMoment(X', ε')
   most $|\tilde{\mu}' - \mu'| \leq \sqrt{P}/4$, emiatt $|\sqrt{P}\tilde{\mu}' - \mu| \leq P/4$.
4. **HA** $\sqrt{P}\tilde{\mu}' \leq \frac{P}{2}$, akkor $p = \mu \leq \frac{3}{4}P$, így $P := \frac{3}{4}P$ (továbbra is jó felsőkorlát) és **GOTO 1.**
5. **KÜLÖNBEN** $\sqrt{P}\tilde{\mu}' > \frac{P}{2}$, ekkor $p = \mu \geq \frac{1}{4}P$. **STOP**

Két esetben állhattunk meg:

1. eset: $p \leq P \leq \frac{1}{4n^2}$, ekkor $\tilde{\mu} := 0$ jó becslés, mert

$$\sqrt{p} \leq \frac{1}{2n} \Longrightarrow p \leq \frac{\sqrt{p}}{2n}$$

Tehát kihasználva, hogy $\sigma = \sqrt{p(1-p)} \geq \sqrt{p \cdot \frac{1}{4}} = \frac{\sqrt{p}}{2}$:

$$|\tilde{\mu} - \mu| = |0 - \mu| = p \leq \frac{\sqrt{p}}{2n} \leq \frac{\sigma}{n}$$

2. eset: $p \geq \frac{1}{4}P$. Ekkor legyen $\hat{p} := \frac{P}{2}$, ez 2-approximáció p-re, mert:

$$p \geq \frac{1}{4}P \quad \Longrightarrow \quad 2p \geq \frac{1}{2}P = \hat{p}$$

és

$$p \leq P \quad \Longrightarrow \quad \frac{p}{2} \leq \frac{P}{2} = \hat{p}$$

Ebben az esetben legyen $X'' := \frac{X}{\sqrt{2\hat{p}}}$ és $\varepsilon'' := \frac{1}{4n}$.

$\tilde{\mu}'' :=$ EstimateForSmallSecondMoment(X'', ε''), amire $|\tilde{\mu}'' - \mu''| \leq \frac{1}{4n}$ és így $\tilde{\mu} := \sqrt{2\hat{p}} \cdot \tilde{\mu}''$ jó becslés:

$$|\tilde{\mu} - \mu| \leq \frac{\sqrt{2\hat{p}}}{4n} \leq \frac{\sqrt{4p}}{4n} = \frac{\sqrt{p}}{2n} \leq \frac{\sigma}{n}$$

Már csak a minták számát kell látni, hogy jó lesz. A második esetben a 3) hívása $O(1/\varepsilon'') = O(n)$. Ezen kívül pedig csak az kérdés, hogy összesen hány mintát használunk a 3. lépésben. Minden iterációban P csökken a 3/4-ére, így ε' is mindig csökken $\sqrt{3/4}$-edére. Az előbbi lemmában látott módszer itt is alkalmazható, mivel a j. iterációban $\varepsilon' = \frac{1}{4}\left(\sqrt{\frac{3}{4}}\right)^j \approx \left(\sqrt{\frac{3}{4}}\right)^j$, így most $\sqrt{\frac{3}{4}}$ ami az előbb $1 - c'$ volt. Továbbá, mivel P értéke végig nem kisebb, mint $\frac{1}{4n^2}$, így $\varepsilon' \geq \frac{1}{4}\sqrt{\frac{1}{4n^2}} = \frac{1}{8n}$ végig, tehát mint az előbb a szumma $\approx \frac{1}{\frac{1}{8n}} = O(n)$.

5) EstimateForBoundedVariable(X, n):

- $X \in [0, 1]$ valós valószínűségi változó és $n \in \mathbb{Z}^+$.
- RETURN $\tilde{\mu}$, amire $|\tilde{\mu} - \mu| \leq \frac{\sqrt{\mu}}{n}$.

Lemma 2.4.1.5.: Ha 4) megoldható $O(n)$ mintával, akkor 5) is megoldható $O(n)$ mintával.

Készítsünk egy $X' \in \{0, 1\}$ valószínűségi változót X-ből úgy, hogy egy $x' \sim X'$ mintavétel nézzen ki a következőképpen:

1. Először vegyünk egy mintát $x \sim X$-ből.
2. Készítsünk egy $X'' \in \{0, 1\}$ változót, aminek a várható értéke $\mu'' = x$.
3. RETURN $x' \sim X''$.

Ekkor X'-re alkalmazható 4) és a kívánt becslést adja, mivel $\mathbb{E}[X'] = \mathbb{E}[X] = \mu$ valamint

$$\sigma(X') = \sqrt{\mathbb{E}[(X')^2] - (\mathbb{E}[X'])^2} \leq \sqrt{\mathbb{E}[(X')^2]} \leq \sqrt{\mu}$$

6) EstimateForBoundedStddev(X, n, σ_B):

- X valós valószínűségi változó, $n \in \mathbb{Z}^+$ és $\sigma_B \geq 0$
- Teljesül, hogy $\sigma \leq \sigma_B$
- RETURN $\tilde{\mu}$, amire $|\tilde{\mu} - \mu| \leq \frac{\sigma_B}{n}$

Lemma 2.4.1.6.: Ha 3) megoldható $O(1/\varepsilon)$ mintával, akkor 6) megoldható $O(n)$ mintával.

Ha $\sigma_B = 0$, akkor X konstans és egy minta pontos lesz. Különben skálázzuk úgy X-et, hogy

$\sigma_B = 1/4$ legyen. (Tehát $X := X/(4 \cdot \sigma_B)$, az eredményként kapott $\tilde{\mu}$-t pedig majd értelemszerűen tudjuk visszaskálázni.)

Először vegyünk egy mintát $x \sim X$-ből. Legyen $X' := X - x$.

Tekintsük s'-t, felhasználva hogy $(a-b)^2 \leq 2a^2 + 2b^2$:

$$(s')^2 = \mathbb{E}[(X')^2] = \mathbb{E}[(X-x)^2] = \mathbb{E}[(X-\mu+\mu-x)^2] \leq 2\mathbb{E}[(X-\mu)^2] + 2\mathbb{E}[(x-\mu)^2]$$

A Csebisev egyenlőtlenségből kapjuk, hogy $|x-\mu| \leq 2\sigma$ legalább $3/4$ valószínűséggel:

$$(s')^2 \leq 2\mathbb{E}[(X-\mu)^2] + 2\mathbb{E}[(x-\mu)^2] \leq 2\sigma^2 + 2(2\sigma)^2 = 10\sigma^2 \leq 10\sigma_B^2 = \frac{10}{4^2} < 1$$

Mivel $s \leq 1$, alkalmazhatjuk 3)-at $\varepsilon' := \frac{1}{4n}$-nel. $\tilde{\mu} :=$ EstimateForSmallSecondMoment(X', ε)

$$|\tilde{\mu} - \mu| \leq \frac{1}{4n} = \frac{\sigma_B}{n}$$

7) ImprovedEstimateForBoundedVariable(X, n):

- $X \in [-1, 1]$ valós valószínűségi változó, $n \in \mathbb{Z}^+$ és $s \geq \frac{1}{n}$.
- RETURN $\tilde{\mu}$, amire $|\tilde{\mu} - \mu| \leq \frac{s}{n}$.

Lemma 2.4.1.7.: Ha 5) megoldható $O(n)$ mintával és 3) $O(1/\varepsilon)$ mintával, akkor 7) is megoldható $O(n)$ mintával.

Legyen $X' := X^2$, ekkor $X' \in [0, 1]$ és 5) alkalmazható, $\tilde{\mu}' :=$ EstimateForBoundedVariable$(X', 2n)$, ekkor

$$|\tilde{\mu}' - \mu'| \leq \frac{\sqrt{\mu'}}{2n}$$

és mivel $\mu' = \mathbb{E}[X'] = \mathbb{E}[X^2] = s^2$, ha $\hat{s} := \tilde{\mu}'$, kihasználva hogy $s \geq \frac{1}{n}$:

$$|\hat{s} - s^2| \leq \frac{s}{2n} \leq \frac{s^2}{2}$$

Ekkor $\hat{s}$ az s^2 2-approximációja, mert

$$\hat{s} \leq s^2 + \frac{s^2}{2} < 2s^2 \qquad \hat{s} \geq s^2 - \frac{s^2}{2} = \frac{s^2}{2}$$

Így $\tilde{s} := \sqrt{\hat{s}}$ az s-nek 2-approximációja ($\sqrt{2}$-approximáció is, de 2-approximáció elég nekünk).

Legyen most $X'' := \frac{X}{2\tilde{s}}$ és $\varepsilon'' := \frac{1}{4n}$, mivel

$$(s'')^2 = \frac{\mathbb{E}[X^2]}{4\tilde{s}^2} = \frac{s^2}{4\hat{s}^2} \leq \frac{s^2}{4(\frac{s}{2})^2} = 1$$

ezért 3) alkalmazható. Legyen $\tilde{\mu}'' :=$ EstimateForSmallSecondMoment(X'', ε'') és $\tilde{\mu} := 2\tilde{s}\tilde{\mu}''$.

$$|\tilde{\mu} - \mu| = 2\tilde{s}|\tilde{\mu}'' - \mu| \leq 2\tilde{s}\frac{1}{4n} \leq \frac{s}{n}$$

8) EstimateFor(X, n):

- X valós valószínűségi változó, $n \in \mathbb{Z}^+$.
- RETURN $\tilde{\mu}$, amire $|\tilde{\mu} - \mu| \leq \frac{s}{n}$.

Lemma 2.4.1.8.: Ha 7) megoldható $O(n)$ mintával, akkor 8) is megoldható $O(n)$ mintával.

A lemma bizonyításának ötlete, hogy ha $|X|$ megfelelő rendű kvantilisét vesszük és X-et aszerint megszorítva adunk becslést, akkor a megszorításon kívül eső értékek érdemben már nem befolyásolhatják a becslést.

Használva Hamoudi tételét, az 1.2.3.-as tételt konstans hibával, ha $|X|$ $\frac{1}{n^2}$-rendű kvantilisét szeretnénk megkapni, akkor $O(n)$ mintával olyan B becslést kapunk, amire: $Q(\frac{1}{n^2}) \leq B \leq Q(\frac{c}{n^2})$.

A kvantilis definíciójából következik, hogy $\mathbb{P}[|X| \geq B] \geq \frac{1}{n^2}$ és $\mathbb{P}[|X| > B] \leq \frac{c}{n^2}$.

Legyen $X' := \begin{cases} X & \text{ha } |X| \leq B \\ B & \text{ha } |X| > B \\ -B & \text{ha } |X| < B \end{cases}$

A 7)-es problémára akarunk visszavezetni. Ha $B = 0$, akkor könnyű dolgunk van, hisz $X' \equiv 0$

és $\mu := 0$ jó, különben legyen $X'' := \frac{X'}{B}$, így $X'' \in [-1, 1]$ és teljesül:

$$(s')^2 = \mathbb{E}[(X')^2] \geq B^2\mathbb{P}[|X'| \geq B] = B^2\mathbb{P}[|X'| \geq B] \geq \frac{B^2}{n^2}$$

Tehát $s'' = \frac{s'}{B} \geq \frac{1}{n}$ és így 7) hívható:

$\tilde{\mu}'' :=$ ImprovedEstimateForBoundedVariable(X'', n), amire $|\tilde{\mu}'' - \mu''| \leq \frac{s''}{n}$, azaz:

$$\left|\tilde{\mu}'' - \frac{\mu'}{B}\right| \leq \frac{s''}{n} = \frac{s'}{Bn} \qquad \Longrightarrow \qquad |B\tilde{\mu}'' - \mu'| \leq \frac{s'}{n}$$

Legyen $\tilde{\mu} := B\tilde{\mu}''$, azt állítjuk hogy ez jó becslés lesz. Egyrészt mivel X' az X megszorítása $[-B, B]$-re, így $s' \leq s$ és

$$|\tilde{\mu} - \mu'| \leq \frac{s'}{n} \leq \frac{s}{n}$$

Másrészt Cauchy-Schwarz miatt:

$$|\mu - \mu'| = \mathbb{E}[X - X'] \leq \mathbb{E}\Big[|X - X'|\Big] \leq \mathbb{E}\Big[||X| - B|\mathbb{1}_{\{|X|>B\}}\Big] \leq$$

$$\leq \sqrt{\mathbb{E}\Big[(|X| - B)^2\ \mathbb{1}_{\{|X|>B\}}\Big]}\sqrt{\mathbb{E}\Big[\mathbb{1}_{\{|X|>B\}}\Big]} \leq \sqrt{\mathbb{E}[|X|]^2}\sqrt{\mathbb{P}[|X| > B]} \leq s \cdot \frac{\sqrt{c}}{n} = O\Big(\frac{s}{n}\Big)$$

Tehát $|\tilde{\mu} - \mu| \leq |\tilde{\mu} - \mu'| + |\mu - \mu'| \leq O\Big(\frac{s}{n}\Big)$. (A konstans faktort pedig átvihetjük a mintavételek számába - ami úgy továbbra is $O(n)$ - hogy $\leq \frac{s}{n}$ legyen.)

9) OptimalEstimateFor(X, n):

- X valós valószínűségi változó, $n \in \mathbb{Z}^+$.
- RETURN $\tilde{\mu}$, amire $|\tilde{\mu} - \mu| \leq \frac{\sigma}{n}$.

Lemma 2.4.1.9.: Ha 8) megoldható $O(n)$ mintával, akkor 9) is megoldható $O(n)$ mintával.

Vegyünk egy $x \sim X$ mintát egyszer és legyen $X' := X - x$. Legyen $\tilde{\mu}' :=$ EstimateFor(X', n).

Ekkor, hasonlóan ahogy láttuk a 2.4.1.6.-os lemmában

$$(s')^2 = \mathbb{E}[(X')^2] = \mathbb{E}[(X-x)^2] = \mathbb{E}[(X-\mu+\mu-x)^2] \leq$$

$$\leq 2\mathbb{E}[(X-\mu)^2] + 2\mathbb{E}[(x-\mu)^2] \leq 2\sigma^2 + 2(2\sigma)^2 = 10\sigma^2$$

És így

$$|\tilde{\mu}' - \mu'| \leq \frac{s'}{n} \leq \frac{\sqrt{10}\sigma}{n} = O\Big(\frac{\sigma}{n}\Big)$$

És itt is a minták számát konstansszorosára növelve pontosan $\frac{\sigma}{n}$-et kapunk.

Beláttuk tehát a 2.4.1.1.-es állítást, azaz hogy ha az 1)-es problémát meg tudjuk oldani valamely $0 < c < 1$ konstansra $O(1/\varepsilon)$ mintával, akkor a 2.4.2.-es tétel igaz.

2.4.2. Kvantumalgoritmus a döntési problémára

Az előző visszavezetés nyomán azt szeretnénk belátni, hogy az 1)-es problémára $c = 1/2$ -del létezik $O(1/\varepsilon)$ mintát használó kvantum algoritmus. Maga az algoritmus egy kvantum fázisbecslés lesz, mielőtt az algoritmust leírom, az elemzéséről írok tömören.

Tétel 2.4.2.1. (Kothari-O'Donnell): Legyen adott egy X diszkrét valós valószínűségi változó, melyre teljesül, hogy $s := \sqrt{\mathbb{E}[X^2]} \leq 1$. Ha adott $0 < \varepsilon \leq 1$, akkor $O(1/\varepsilon)$ mintavétellel megkülönböztethető:

$$i)\ |\mu| \leq \frac{\varepsilon}{2} \qquad \text{vagy} \qquad ii)\ \varepsilon \leq |\mu| \leq 2\varepsilon$$

Legyen most adott az X valószínűségi változó és a 2.3.-ban definiált $\mathcal{P}$ és $\mathcal{B}_X$ orákulumok, ezekből $O(1)$ hívással készíthető a következő $\mathcal{U}$ kapu, sőt ennek a controlled-$\mathcal{U}$ változata is (a pontos előállítás [KO23] Remark 3.3. és Appendix A alatt):

$$\mathcal{U} := REFL_{\mathbb{P}} \cdot ROT_X$$

ahol

$$REFL_{\mathbb{P}} := \mathcal{P}(2|\vec{0}\rangle\langle\vec{0}| - \mathcal{I})\mathcal{P}^{\dagger}$$

és ROT_X pedig az a kapu, ami a következő komplex forgatást csinálja

$$ROT_X|\omega\rangle|garbage_\omega\rangle = e^{-2i\cdot arctan\left(X(\omega)\right)}|\omega\rangle|garbage_\omega\rangle$$

Megjegyzés 2.4.2.2.: Az $\mathcal{U}$ kapu meghatározására úgy könnyű gondolni, hogy a Grover algoritmus mintájára definiáljuk a mostani általánosabb problémára. Visszaemlékezve a Grover algoritmusra, láttuk hogy a $\mathcal{H}^{\otimes n}\mathcal{R}\mathcal{H}^{\otimes n}$ rész egy tükrözés $\mathcal{H}^{\otimes n}|0\rangle^n$-en keresztül, ami épp az uniform eloszlás szerinti $REFL_{\mathbb{P}}$. A Grover algoritmus problémájában viszont csak két lehetséges felvett érték ± 1 volt, ehelyett most tetszőleges valós értéke lehet $X(\omega)$-nak. Az ötlet, hogy vegyük a $2arctan(X)$-et, onnan jön hogy most minden lehetséges értéket egyértelműen fázisba szeretnénk elkódolni.

Jelölés 2.4.2.3.: $\theta \sim \Theta_U(|\sigma\rangle)$
Legyen $\mathcal{U}' \in \mathbb{R}^{n\times n}$ tetszőleges unitér mátrix, $|\sigma\rangle \in R^n$ pedig egy egységvektor. Tekintsük $\mathcal{U}'$ egy sajátfelbontását: $\mathcal{U}' = \sum_{j=1}^{n} e^{i\theta_j}|u_j\rangle\langle u_j|$. Ekkor $|\sigma\rangle$ kifejezve $\mathcal{U}'$ sajátbázisában: $|\sigma\rangle = \sum_{j=1}^{n} \hat{\sigma}_j|u_j\rangle$, ahol $\hat{\sigma}_j := \langle u_j|\sigma\rangle$. Mivel $|\sigma\rangle$ egységvektor és $\{|u_j\rangle\}_{j=1}^n$ bázis ortonormált, így $|\hat{\sigma}_1|^2, ..., |\hat{\sigma}_n|^2$ egy diszkrét valószínűségi eloszlást határoz meg $[n]$-en. Ekkor $\theta \sim \Theta_U(|\sigma\rangle)$ jelentse azt, hogy először eszerint az eloszlás szerint veszünk egy j indexet, majd visszaadjuk azt a θ_j-t ami a sajátfelbontásban a j indexhez tartozik.

Jelölés 2.4.2.4.: Legyen z komplex valószínűségi változó Ω eseményhalmazon, ekkor jelölje

$$|\mathbf{z}\rangle := \sum_{\omega\in\Omega} z(\omega)\cdot\sqrt{\mathbb{P}[\omega]}|\omega\rangle|garbage_\omega\rangle$$

Ezzel a jelöléssel tehát legyenek:

$$|\mathbf{1}\rangle := \sum_{\omega\in\Omega} 1\cdot\sqrt{\mathbb{P}[\omega]}|\omega\rangle|garbage_\omega\rangle \quad \left(=\mathcal{P}|\vec{0}\rangle\right) \qquad |\mathbf{1}+\mathbf{i}\mathbf{X}\rangle := \sum_{\omega\in\Omega}\left(1+iX(\omega)\right)\cdot\sqrt{\mathbb{P}[\omega]}|\omega\rangle|garbage_\omega\rangle$$

A [KO23] cikk egyik fő gondolata, hogy ha s kicsi, akkor ez a két vektor közel van egymáshoz. Egyrészt $|\mathbf{1}\rangle$ egységvektor és, ha s kicsi, akkor közel van egy egységvektorhoz, mert

$$\langle \mathbf{1}+\mathbf{iX}|\mathbf{1}+\mathbf{iX}\rangle = \mathbb{E}\Big[\overline{(1+iX)}(1+iX)\Big] = \mathbb{E}[1+X^2] = 1+s^2$$

Másrészt, ha s kicsi, akkor a Hölder egyenlőtlenség miatt $|\mu|$ is kicsi, hiszen $|\mu| = \big|\mathbb{E}[X]\big| \leq \mathbb{E}[|X|] \leq \sqrt{\mathbb{E}[|X|^2]} = s$, és

$$\big|\langle \mathbf{1}|\mathbf{1}+\mathbf{iX}\rangle\big| = \big|\mathbb{E}[1+iX]\big| = \sqrt{1+\mu^2}$$

Az $|\mathbf{1}\rangle$ egységvektor, viszont $|\mathbf{1}+\mathbf{iX}\rangle$ tipikusan nem (csak akkor az, ha $\mu = 0$). Jelölje a normáltját $|\mathbf{1}+\mathbf{iX}\rangle_{norm} = \frac{1}{\sqrt{1+s^2}}|\mathbf{1}+\mathbf{iX}\rangle$.

Állítás 2.4.2.5.: Ha $\theta \sim \Theta_U(|\mathbf{1}\rangle)$ és $\tilde{\theta} \sim \Theta_U(|\mathbf{1}+\mathbf{iX}\rangle_{norm})$, akkor

$$\mathbb{E}\left[sin\Big(\frac{\theta}{2}\Big)^{-2}\right] = \frac{1+s^2}{\mu} \quad \text{és} \quad \mathbb{E}\left[sin\Big(\frac{\tilde{\theta}}{2}\Big)^{2}\right] = \frac{\mu}{1+s^2}$$

Biz.: Először is általánosan egy $|\sigma\rangle$ egységvektorra $\hat{\sigma}_j = \langle u_j|\sigma\rangle$ esetén

$$\mathcal{U}|\sigma\rangle = \sum_j \hat{\sigma}_j e^{i\theta_j}|u_j\rangle \qquad \Longrightarrow \qquad \left(\frac{I-\mathcal{U}}{2}\right)^{\pm 1}|\sigma\rangle = \sum_j \hat{\sigma}_j \left(\frac{1-e^{i\theta_j}}{2}\right)^{\pm 1}|u_j\rangle$$

és így

$$\left|\left|\left(\frac{I-\mathcal{U}}{2}\right)^{\pm 1}|\sigma\rangle\right|\right|^2 = \mathbb{E}_{\theta\sim\Theta_{\mathcal{U}}(|\sigma\rangle)}\left[\left|\frac{1-e^{i\theta}}{2}\right|^{\pm 2}\right] =$$

$$= \mathbb{E}_{\theta\sim\Theta_{\mathcal{U}}(|\sigma\rangle)}\left[\left|\frac{1-cos(\theta)}{2}\right|^{\pm 2}\right] = \mathbb{E}_{\theta\sim\Theta_{\mathcal{U}}(|\sigma\rangle)}\left[sin\Big(\frac{\theta}{2}\Big)^{\pm 2}\right]$$

Speciálisan $\mathcal{U}|\mathbf{1}+\mathbf{iX}\rangle$ meghatározásához két tulajdonságot használunk ki:

a) $e^{-2i\cdot arctan(y)}(1+iy) = 1-iy$
b) $REFL_{\mathbb{P}}|\mathbf{z}\rangle = \big(\mathcal{P}(2|0\rangle\langle 0|-I)\mathcal{P}^\dagger\big)|\mathbf{z}\rangle = \big(2|\mathbf{1}\rangle\langle\mathbf{1}|-I\big)|\mathbf{z}\rangle = 2|\mathbf{1}\rangle\langle\mathbf{1}|\mathbf{z}\rangle - |\mathbf{z}\rangle = |2\mathbb{E}[\mathbf{z}]-\mathbf{z}\rangle$

Így

$$\mathcal{U}|\mathbf{1}+\mathbf{iX}\rangle = REFL_{\mathbb{P}}ROT_X|\mathbf{1}+\mathbf{iX}\rangle = REFL_{\mathbb{P}}e^{-2i\cdot arctan(X)}|\mathbf{1}+\mathbf{iX}\rangle =$$

$$= REFL_{\mathbb{P}}|\mathbf{1}-\mathbf{iX}\rangle = |\mathbf{1}+\mathbf{i}(\mathbf{X}-\mathbf{2}\mu)\rangle = |\mathbf{1}+\mathbf{iX}\rangle - 2i\mu|\mathbf{1}\rangle$$

Átrendezve:

$$\frac{I-U}{2}|\mathbf{1}+\mathbf{iX}\rangle = i\mu|\mathbf{1}\rangle \qquad \text{illetve} \qquad \Big(\frac{I-U}{2}\Big)^{-1}|\mathbf{1}\rangle = \frac{1}{i\mu}|\mathbf{1}+\mathbf{iX}\rangle$$

Tehát egyrészt

$$\mathbb{E}_{\tilde{\theta}\sim\Theta_{\mathcal{U}}(|\mathbf{1}+\mathbf{iX}\rangle_{norm})}\left[sin\Big(\frac{\tilde{\theta}}{2}\Big)^2\right] = \left|\left|\frac{I-\mathcal{U}}{2}|\mathbf{1}+\mathbf{iX}\rangle\right|\right|^2 = \left|\left|\frac{i\mu|\mathbf{1}\rangle}{\sqrt{1+\mathbb{E}[X^2]}}\right|\right|^2 = \frac{\mu^2}{1+s^2}$$

másrészt

$$\mathbb{E}_{\theta\sim\Theta_{\mathcal{U}}(|\mathbf{1}\rangle)}\left[sin\Big(\frac{\tilde{\theta}}{2}\Big)^{-2}\right] = \left|\left|\Big(\frac{I-\mathcal{U}}{2}\Big)^{-1}|\mathbf{1}\rangle\right|\right|^2 = \left|\left|\frac{1}{i\mu}|\mathbf{1}+\mathbf{iX}\rangle\right|\right|^2 = \frac{1+\mathbb{E}[X^2]}{\mu^2} = \frac{1+s^2}{\mu^2}$$

Az állításból látjuk, hogy várhatóan $sin(\theta/2)^2 \approx \frac{\mu^2}{1+s^2}$, ami indukája, hogy ha s kicsi akkor $|\theta/2| \approx |\mu|$, azaz $|\theta| \approx 2|\mu|$.

Egész pontosan a következő hozható ki a fenti állításból (lásd [KO23] Theorem 3.18 és Section 3.5):

Állítás 2.4.2.6.: Feltéve, hogy $s \leq \frac{1}{16}$, ha $\theta \sim \Theta_{\mathcal{U}}\big(|\mathbf{1}\rangle\big)$, akkor

$$\mathbb{P}\Big[\frac{4}{5}\cdot 2|\mu| \leq |\theta| \leq \frac{5}{4}\cdot 2|\mu|\Big] \geq 1-\frac{2}{9}$$

Ennek segítségével térjünk rá az algoritmusra. Idézzük vissza, hogy 2.4.2.1. tétel szerint adott ε-ra $O(1/\varepsilon)$ mintavétellel és legfeljebb 1/3 hibával akarunk tudni megkülönböztetni $|\mu| \leq \varepsilon/2$ és $\varepsilon \leq |\mu| \leq 2\varepsilon$ között. Most egyrészt egy megengedőbb $|\mu| \leq \varepsilon/2$ és $\varepsilon \leq |\mu|$ közti megkülönböztetést

adunk, másrészt a tételben csak $s \leq 1$-et tettük fel, míg 2.4.2.6.-ban $s \leq \frac{1}{16}$-ot. Ez utóbbi nem nagy probléma, X átskálázásával elérhető, hogy $s \leq \frac{1}{16}$ legyen és ez csak konstansszorosára növeli a minták számát.

Algoritmus 2.4.2.7. (Kothari-O'Donnell):

1. Csináljunk egy kvantum fázisbecslést az $\mathcal{U}$ kapura és a $\mathcal{P}|0\rangle$ állapotra $\varepsilon/6$ pontossággal úgy, hogy a hiba valószínűsége legfeljebb $1/9$ legyen. Legyen az output θ'. (A pontosság eléréséhez így $O(1/\varepsilon)$ controlled-$\mathcal{U}$ kapura van szükségünk, és a minták száma $= \mathcal{U}$ hívása $O(1/\varepsilon)$).
   Ekkor $\mathbb{P}[|\theta - \theta'| > \varepsilon/6] < 1/9$.
2. **HA** $\theta' < \frac{142}{100} \cdot \varepsilon$ **RETURN** $|\mu| \leq \varepsilon/2$
   KÜLÖNBEN **RETURN** $\varepsilon \leq |\mu|$

Az algoritmus helyességéhez: Most θ' legfeljebb $\varepsilon/6$ addíciós hibával tér el θ-tól, kivéve $1/9$ valószínűséggel. Az összhiba tehát $1/9 + 2/9 = 1/3$, ami kivételével teljesül, hogy:

$$\frac{4}{5}|\mu| - \frac{\varepsilon}{12} \leq \left|\frac{\theta'}{2}\right| \leq \frac{5}{4}|\mu| + \frac{\varepsilon}{12}$$

Ha $|\mu| \leq \varepsilon/2$, akkor $|\theta'/2| \leq 5/8 \cdot \varepsilon + \varepsilon/12 < (71/100) \cdot \varepsilon$.
Ha $\varepsilon \leq |\mu|$, akkor $|\theta'/2| \geq 4/5 \cdot \varepsilon - \varepsilon/12 > (71/100) \cdot \varepsilon$.

3. Többdimenziós változó várható értékének becslése

Most az előző fejezetben tárgyalt várható érték becslési problémát terjesztjük ki többdimenziós valószínűségi változó esetére. Magasabb ($d > 1$) dimenzióban rögtön kérdés lesz hogy milyen normában vizsgáljuk az eltérést a becslés és a valódi $\mu \in \mathbb{R}^d$ között. Ebben a fejezetben először az egydimenzióhoz képest felmerülő különbségekről írok, majd ismertetem a legjobb ismert és közel-optimális (polilogaritmikus faktoroktól eltekintve optimális) kvantumos módszert, Arjan Cornelissen, Yassine Hamoudi és Sofiene Jerbi [CHJ22] cikkéből. Először az algoritmus fő gondolatait fogom megmutatni néhány korlátozottabb példán, majd rátérek az algoritmusra ahol még feltesszük, hogy $||X||$ korlátos végül az utolsó részben arról írok, hogy ezt a szerzők hogy terjesztik ki korlátlan esetre kvantilisek segítségével.

3.1. A probléma magasabb dimenzióban

Idézzük fel a 2.1. részben megfogalmazott problémát. Legyen adott most egy d dimenziós diszkrét valós valószínűségi változó $X : \Omega \to \mathbb{R}^d$ az $(\Omega, 2^\Omega, \mathbb{P})$ valószínűségi mezőn. Megbecsülendő továbbra is:

$$\mu := \mathbb{E}[X] = \sum_{\omega \in \Omega} \mathbb{P}[\omega] X(\omega)$$

Most $\mu \in \mathbb{R}^d$, így a célunk hogy egy $\tilde{\mu} \in \mathbb{R}^d$-t adjunk, ami nagy valószínűséggel közel van μ-höz. A közelséget kettes normában fogjuk érteni, így tehát azt mondjuk, hogy

Definíció 3.1.1.: $\tilde{\mu}$ a várhatóértéknek ε-becslése, ha fennáll

$$\mathbb{P}[||\tilde{\mu} - \mu||_2 > \varepsilon] \leq \frac{1}{3}$$

Megjegyzés 3.1.2.: A definícióban hogy hangsúlyos legyen kiírtam a "2"-t a normához. A továbbiakban jellemzően nem írom ki, így $||.||$ értelemszerűen mindig $||.||_2$-t fog jelenteni.

Klasszikus eset, ha $d > 1$:
Láttuk, hogy klasszikus esetben $d = 1$ esetén fix n mintavétel esetén az elérhető $\varepsilon(n)$ becslési pontosság $O\left(\frac{\sigma}{\sqrt{n}}\right)$ volt. Ez $d > 1$ esetén $O\left(\sqrt{\frac{Tr[\Sigma]}{n}}\right)$ lesz, ahol Σ az X kovarianciamátrixa. Magasabb dimenzióban is alkalmazható a median-of-means algoritmus, ha az úgynevezett geometriai mediánt, azaz a mintáktól vett minimális euklideszi távolságot használjuk, akkor [Min15] Corollary 4.1. szerint

$$\mathbb{P}\left[||\tilde{\mu} - \mu|| > O\left(\sqrt{\frac{Tr[\Sigma]log(1/\delta)}{n}}\right)\right] \leq \delta$$

Tehát $\delta = \frac{1}{3}$-ra, $O\left(\sqrt{\frac{Tr[\Sigma]}{n}}\right)$. Ez jól láthatóan általánosítása az egydimenziós esetnek, hiszen $d = 1$ esetén $O\left(\sqrt{\frac{Tr[\Sigma]}{n}}\right) = O\left(\sqrt{\frac{Var[X]}{n}}\right) = O\left(\frac{\sigma}{\sqrt{n}}\right)$.

A mintavételről:
Kvantumos esetben a mintavételre továbbra is úgy gondolunk, mint 2.3. részben írtam, adottak lesznek $\mathcal{P}$ és $\mathcal{B}_X$ orákulumjaink. (Megjegyezhetjük, hogy ekkor a $\mathcal{B}_X$-ben szereplő $|X(\omega)\rangle = |X(\omega)_1\rangle \otimes ... \otimes |X(\omega)_d\rangle$, ahol $|X(\omega)_j\rangle$ a j-edik koordinátája.)

A közel-optimális algoritmus egyik kulcslépése, hogy valószínűségi orákulumból fázisorákulumot hatékonyan tudunk csinálni, [Gil19] 4.4.1. alapján:

Definíció 3.1.3.: (Valószínűségi orákulum) Legyen adott egy $p : X \to [0, 1]$ függvény. Ekkor $\mathcal{U}_p$ kaput valószínűségi orákulumnak hívjuk, ha a következő leképezést csinálja:

$$\mathcal{U}_p|0\rangle|x\rangle \to \left(\sqrt{p(x)}|1\rangle|\psi_x^{(1)}\rangle + \sqrt{1-p(x)}|0\rangle|\psi_x^{(0)}\rangle\right)|x\rangle$$

ahol $|\psi_x^{(0)}\rangle$ és $|\psi_x^{(1)}\rangle$ tetszőleges normalizált vektorok.

Definíció 3.1.4.: (Fázisorákulum) Legyen adott egy $f : X \to [-1, 1]$ függvény. Ekkor $\mathcal{O}_f$ kaput fázisorákulumnak hívjuk, ha a következő leképezést csinálja:

$$\mathcal{O}_f|0\rangle|x\rangle \to e^{if(x)}|0\rangle|x\rangle$$

Tétel 3.1.5. (Valószínűségi orákulum átalakítása fáziorákulummá): Legyen $\mathcal{U}_p$ a 3.1.3. definícióban megadott unitér és legyen adott $t \geq 0, \varepsilon' \in (0,1)$. Ekkor $\mathcal{U}_p$ és $\mathcal{U}_p^\dagger$ $O\big(t + log(1/\varepsilon')\big)$ használatával előállítható $\mathcal{O}_{p,t,\varepsilon'}$, melyre

$$\mathcal{O}_{p,t,\varepsilon'}|u\rangle|0\rangle \longrightarrow |u\rangle|\varphi_u\rangle \qquad \text{és} \qquad \big|\big||\varphi_u\rangle - e^{itp}|0\rangle\big|\big| \leq \varepsilon'$$

Következmény 3.1.6.: Ha adott egy $\mathcal{U}_f$ valószínűségi orákulum, akkor annak $\widetilde{O}(1)$ hívásával előállítható az f-hez tartozó $\mathcal{O}_f$ fázisorákulum.

Biz.: ([Gil19] alapján) Vegyük észre, hogy a $(\langle\vec{0}| \otimes \mathcal{I})\big(\mathcal{U}_p^\dagger(\mathcal{Z} \otimes \mathcal{I})\mathcal{U}_p\big)(|\vec{0}\rangle \otimes \mathcal{I}) = diag(1 - 2p(x))$. Tehát a $H := \mathcal{U}_p^\dagger(\mathcal{Z} \otimes \mathcal{I})\mathcal{U}_p$ Hamilton mátrix blokk-elkódolása a $diag(1 - 2p(x))$ mátrixnak. Ekkor az $e^{-iHt/2}$ Hamilton szimulációt véve, implementálható a kívánt fázisorákulum.

3.2. A kvantum algoritmus fő gondolata

Ebben a részben az általános többdimenziós probléma helyett sokkal megszorítottabb eseteket nézünk, amik előjelzik hogy mi lesz a végső algoritmus gondolati íve.

Probléma 1.) Legyen $d = 1$ és $X : \Omega \to [0,1]$.

A probléma sokkal kötöttebb, rögtön visszamegy az egydimenziós esethez, ráadásul még a valószínűségi változó értékkészlete is jócskán megvan szorítva. Mindenesetre erre a nagyon speciális esetre egy szimpla amplitúdó becslés elég lesz.

Induljunk ki a csupa nulla kezdőállapotból, ekkor egy $\mathcal{P}$ hívással, egy $\mathcal{B}_X$ hívással (és néhány

controlled-forgatással, lásd: [vA21] Lemma 3) kapjuk:

$$|0\rangle|0\rangle|0\rangle \to \sum_{\omega\in\Omega}\sqrt{\mathbb{P}[\omega]}|\omega\rangle|X(\omega)\rangle|0\rangle \to \sum_{\omega\in\Omega}\sqrt{\mathbb{P}[\omega]}|\omega\rangle|X(\omega)\rangle\big(\sqrt{1-X(\omega)}|0\rangle+\sqrt{X(\omega)}|1\rangle\big) =$$

$$= \sum_{\omega\in\Omega}\sqrt{\big(1-X(\omega)\big)\mathbb{P}[\omega]}|\omega\rangle|X(\omega)\rangle|0\rangle + \sum_{\omega\in\Omega}\sqrt{X(\omega)\mathbb{P}[\omega]}|\omega\rangle|X(\omega)\rangle|1\rangle =$$

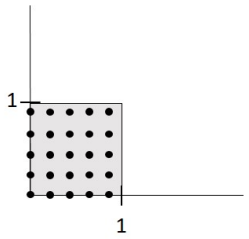
$$= \sqrt{1-\mathbb{E}[X]}|\psi_X^{(0)}\rangle|0\rangle + \sqrt{\mathbb{E}[X]}|\psi_X^{(1)}\rangle|1\rangle$$

Most a feltevés miatt $0 \le \mathbb{E}[X] \le 1$ és így $\sqrt{1-\mathbb{E}[X]}$ és $\sqrt{\mathbb{E}[X]}$ valódi amplitúdók. Tehát elég egy amplitúdó becslést csinálni az utolsó regiszter szerint, hogy megkapjuk $\sqrt{\mathbb{E}[X]} = \sqrt{\mu}$ becslését.

Probléma 2.) Legyen $d \ge 1$ és $X : \Omega \to \{x \in \mathbb{R}^d \mid x \ge 0, ||x||_1 \le 1\}$.

Ugyan X értékkészletét még mindig erősen megszorítjuk, ezen a problémán keresztül meg tudjuk mutatni magasabb dimenzióban hogyan tudjuk ezt hasonlóan megközelíteni. Talán a legszembetűnőbb probléma most az hogy $\mu = \mathbb{E}[X] \in \mathbb{R}^d$ nem egy szám, így önmagában nem csinálhatunk belőle amplitúdót mint az előbb. Az ötlet, hogy vegyünk előre rögzített pontokat és vizsgáljuk a várhatóérték ezekkel vett skalárszorzatát.

Vegyük a $G' := \{\frac{j}{2^m} \mid j = 0, ..., 2^{m-1}\}^d$ rácsot (azaz a $[0,1]^d$ hiperkocka egyenletes felosztása alkossa). Illetve definiáljunk egy $f : G' \to [0,1]$ függvényt, legyen $u \in G'$ esetén $f(u) := \langle u|\mathbb{E}[X]\rangle$.



Először is valóban bármely $u \in G'$ esetén $0 \leq f(u) \leq 1$. Hiszen $0 \leq X$ miatt $0 \leq \mathbb{E}[X]$ és mivel $||X|| \leq ||X||_1 \leq 1$, két legfeljebb 1 hosszú vektor skalárszorzata is legfeljebb 1. Tehát az előzőhöz hasonlóan járhatunk el, annyi különbséggel, hogy ahol az előbb $X(\omega) \in R$ volt, most $uX(\omega) \in R$ skalárszorzatot teszünk.

$$|0\rangle|0\rangle|0\rangle|u\rangle \to \sum_{\omega\in\Omega} \sqrt{\mathbb{P}[\omega]}|\omega\rangle|X(\omega)\rangle|0\rangle|u\rangle \to \sum_{\omega\in\Omega} \sqrt{\mathbb{P}[\omega]}|\omega\rangle|X(\omega)\rangle\big(\sqrt{1-uX(\omega)}|0\rangle + \sqrt{uX(\omega)}|1\rangle\big)|u\rangle$$

$$= \Big(\sum_{\omega\in\Omega} \sqrt{\big(1-X(\omega)\big)\mathbb{P}[\omega]}|\omega\rangle|X(\omega)\rangle|0\rangle + \sum_{\omega\in\Omega} \sqrt{uX(\omega)\mathbb{P}[\omega]}|\omega\rangle|uX(\omega)\rangle|1\rangle\Big)|u\rangle$$

$$= \Big(\sqrt{1-u\mathbb{E}[X]}|\psi_X^{(0)}\rangle|0\rangle + \sqrt{u\mathbb{E}[X]}|\psi_X^{(1)}\rangle|1\rangle\Big)|u\rangle = \Big(\sqrt{1-f(u)}|\psi_X^{(0)}\rangle|0\rangle + \sqrt{f(u)}|\psi_X^{(1)}\rangle|1\rangle\Big)|u\rangle$$

Mint láttuk $0 \leq f(u) \leq 1$, tehát az amplitúdók rendben vannak. Az újabb probléma viszont abból jön, hogy hiába tudnánk megbecsülni a $\sqrt{u\mathbb{E}[X]}$ amplitúdót, nekünk $\mathbb{E}[X]$ kell. Mindenesetre vegyük észre, hogy így $\mathcal{U}_f$ valószínűségi orákulumot kapunk.

Ehelyett legyen tehát a következő az algoritmus:

1. Az előzőek alapján készítsünk $\mathcal{U}_f$ orákulumot: $|0\rangle|u\rangle \xrightarrow{\mathcal{U}_f} \Big(\sqrt{f(u)}|1\rangle|\psi_x^{(1)}\rangle + \sqrt{1-f(u)}|0\rangle|\psi_x^{(0)}\rangle\Big)|u\rangle$

2. Felhasználva a 3.1.6. következményt, készítsünk $\mathcal{O}_f$ fázisorákulumot (ez $\widetilde{O}(1)$-szeres többletmintát jelent): $|0\rangle|u\rangle \xrightarrow{\mathcal{O}_f} e^{if(u)}|0\rangle|u\rangle$

3. Induljunk ki a G' rácson vett uniform szuperpozícióból

$$\frac{1}{\sqrt{|G'|}} \sum_{u\in G'} |u\rangle$$

4. Alkalmazzuk n-szer $\mathcal{O}_f$-et, így kapva:

$$\frac{1}{\sqrt{|G'|}} \sum_{u\in G'} |u\rangle \xrightarrow{(\mathcal{O}_f)^n} \frac{1}{\sqrt{|G'|}} \sum_{u\in G'} e^{inu\mathbb{E}[X]}|u\rangle$$

5. Felidézve, hogy a kvantum Fourier transzformáció a következő leképzést csinálja

$$QFT_{G'} : |k\rangle \longrightarrow \frac{1}{\sqrt{|G'|}} \sum_{u \in G'} e^{2\pi i u k} |k\rangle$$

alkalmazva tehát az inverzét, kapjuk az $\mathbb{E}[X]$-től függő állapotot:

$$\frac{1}{\sqrt{|G'|}} \sum_{u \in G'} e^{inu\mathbb{E}[X]} |u\rangle \xrightarrow{QFT_{G'}^{-1}} \left| round\Big(\frac{\mathbb{E}[X]n}{2\pi}\Big) \right\rangle$$

(Látszik, hogy minél nagyobb n, a kerekítés annál kevésbé befolyásolja az eredményt.)

6. Mérjük meg a kapott állapotot és legyen $\tilde{\mu}$ a $2\pi/n$-szerese.

Probléma 3) Legyen $d \geq 1$ és $X : \Omega \to \{x \in \mathbb{R}^d \mid ||x||_1 \leq 1\}$.

Most megnézzük, hogy lehet kezelni ha nem tesszük fel, hogy $X \geq 0$. Gondoljuk meg mi lenne a helyzet, ha $X \leq 0$ -t tennénk fel. Az ötlet az, hogy ha megtudjuk csinálni az $\mathcal{O}_f$ orákulumot, akkor az inverzét is (áramkör megfordítva), így

$$\frac{1}{\sqrt{|G'|}} \sum_{u \in G'} |u\rangle \xrightarrow{\left(\mathcal{O}_f^{-1}\right)^n} \frac{1}{\sqrt{|G'|}} \sum_{u \in G'} e^{-inu\mathbb{E}[X]} |u\rangle$$

Így tehát ugyanoda jutunk, (most $-u\mathbb{E}[X]$ ugyanaz, mint nemnegatív X-re $u\mathbb{E}[X]$ volt).

Bontsunk tehát két részre, legyenek

$$X(\omega)^+ := \begin{cases} X(\omega) & \text{ha } \langle u|X(\omega)\rangle > 0 \\ 0 & \text{különben} \end{cases} \qquad X(\omega)^- := \begin{cases} X(\omega) & \text{ha } \langle u|X(\omega)\rangle < 0 \\ 0 & \text{különben} \end{cases}$$

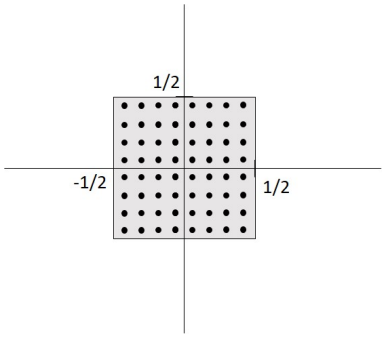
Ha tehát külön-külön felépítjük az X^+-hoz tartozó $\mathcal{O}_f^+$ orákulumot és az és X^--hoz tartozó $\mathcal{O}_f^-$ orákulumot, akkor vehetjük a szorzatukat $\mathcal{O}_f := \mathcal{O}_f^+ \mathcal{O}_f^-$.

3.3. Kvantumalgoritmus korlátos esetben

Ebben a részben veszünk egy jobb rácsot, mint amit az előző problémákban használtunk. Továbbá fel fogjuk tenni, hogy $||X|| \leq 1$ és adunk egy olyan becslést a várható értékre, aminek a végtelen normában való eltérésére megfelelő korlátot tudunk adni ahhoz, hogy a következő alfejezetben ezt kiterjeszthessünk a korlátlan esetre.

Legyen most a G rács a következő:

$$G := \left\{ \frac{j}{m} - \frac{1}{2} + \frac{1}{2m} \;\middle|\; j = 0, ..., m-1 \right\}^d$$



Ekkor egyrészt minden $u \in G$ a $\big(-1/2, 1/2\big)^d$ -be esik, másrészt szimmetrikus, ha $u \in G$ akkor $-u \in G$. A célunk ugyanaz lesz, mint az előző alfejezetben mutattuk, az $\langle u|X\rangle$ értékeket szeretnénk amplitúdókba elkódolni. Azonban az $||X|| \leq 1$ feltételből csak annyi következik, hogy $\langle u|X\rangle \leq \sqrt{d}$. Tekintsük a következő lemmákat:

Lemma 3.3.1.: Legyen $\alpha > 0$, $x \in \mathbb{R}^d$ vektor, ekkor

$$\mathbb{P}_{u\sim G}\Big[\alpha\big|\langle u|x\rangle\big| \geq ||x||\Big] \leq 2e^{-\frac{2}{\alpha^2}}$$

Biz.: Fel fogjuk használni a Hoeffding egyenlőtlenséget, miszerint ha $Z_1, ..., Z_n$ független valószínűségi változók $\mathbb{R}$ felett és $\forall i \ : \ i = 1, ..., n$ teljesül, hogy $a_i \leq Z_i \leq b_i$, akkor

$$\mathbb{P}\Big[\Big(\sum_{i=1}^{n} Z_i - \mathbb{E}\Big[\sum_{i=1}^{n} Z_i\Big] \geq t\Big)\Big] \leq exp\Bigg(\frac{-2t^2}{\sum\limits_{i=1}^{n}(b_i - a_i)^2}\Bigg)$$

Mivel a rácspontok szimmetrikusak 0-ra, így $\mathbb{E}_{u\sim G}\big[\langle u|x\rangle\big] = 0$. Másrészt $u_i \in \big(-\frac{1}{2}, \frac{1}{2}\big)$ miatt:

$$u_i x_i \in \Big(-\frac{x_i}{2}, \frac{x_i}{2}\Big) \quad \Longrightarrow \quad |\alpha u_i x_i| \in \Big[0, \Big|\frac{\alpha x_i}{2}\Big|\Big)$$

Alkalmazva tehát a lemmában szereplő valószínűségre a Hoeffding egyenlőtlenséget:

$$\mathbb{P}\Big[\alpha\big|\langle u|x\rangle\big| \geq ||x||\Big] = \mathbb{P}\Big[\Big|\sum_{i=1}^{d} \alpha\langle u|x\rangle\Big| - 0 \geq ||x||\Big] \leq exp\Bigg(\frac{-2||x||^2}{\sum\limits_{i=1}^{n}\Big|\frac{\alpha x_i}{2}\Big|^2}\Bigg) \leq 2exp\Bigg(-\frac{2||x||^2}{\alpha^2||x||^2}\Bigg) = 2e^{-\frac{2}{\alpha^2}}$$

Lemma 3.3.2.: Legyen $\alpha > 0, X : \Omega \to \mathbb{R}^d$ valószínűségi változó, ekkor

$$\mathbb{P}_{u\sim G}\Bigg[\alpha\mathbb{E}\Big[\big|\langle u|X\rangle\big|\Big] \geq \mathbb{E}\Big[||X||\Big]\Bigg] \leq \frac{\alpha}{2}$$

Biz.: Először becsüljük felül az $\mathbb{E}_u\Big[\big|\langle u|X\rangle\big|\Big]$ várható értéket a Cauchy-Schwarz egyenlőtlenséget használva

$$\mathbb{E}_u\Big[\big|\langle u|X\rangle\big|\Big] \leq \sqrt{\mathbb{E}_u\Big[\big|\langle u|X\rangle\big|^2\Big]} \leq \sqrt{\sum_{i=1}^{d}\mathbb{E}_{u_i}\Big[(u_i X_i)^2\Big]}$$

Mivel $u_i \in \big(-\frac{1}{2}, \frac{1}{2}\big)$, így $(2u_i X_i)^2 \leq X_i^2$ és

$$\mathbb{E}_u\Big[\big|\langle u|X\rangle\big|\Big] \leq \sqrt{\sum_{i=1}^{d}\Big(\frac{X_i}{2}\Big)^2} = \frac{||X||}{2}$$

Alkalmazva egy Markov egyenlőtlenséget:

$$\mathbb{P}_{u\sim G}\Big[\alpha\mathbb{E}\Big[\big|\langle u|X\rangle\big|\Big]\Big] \leq \frac{\mathbb{E}_u\Big[\alpha\mathbb{E}\Big[\big|\langle u|X\rangle\big|\Big]\Big]}{\mathbb{E}\Big[||X||\Big]} = \frac{\alpha\mathbb{E}\Big[\mathbb{E}_u\Big[\big|\langle u|X\rangle\big|\Big]\Big]}{\mathbb{E}\Big[||X||\Big]} \leq \frac{\alpha\mathbb{E}\Big[\frac{||X||}{2}\Big]}{\mathbb{E}\Big[||X||\Big]} = \frac{\alpha}{2}$$

A fenti lemmákból látjuk tehát, hogy az $u \in G$-k nagy részére tehát nem túl nagy az $\alpha\langle u|X\rangle$ érték, ahol az α szorzó majd egy - a mintavétel számától és a valószínűségi változó dimenziójától függő - $(0,1)$-beli érték lesz. Használjuk a [CHJ22]-ben bevezetett jelölést egy vektor norma szerinti levágására:

Jelölés 3.3.3.: Legyen $x \in \mathbb{R}^d$ és $a \leq b$ ekkor $[\![x]\!]_a^b$ alatt a következőt értjük:

$$[\![x]\!]_a^b := \begin{cases} x & a \leq ||x|| \leq b \\ 0 & \text{különben} \end{cases}$$

Az ötlet az, hogy ha $\mathbb{E}[\alpha\langle u|X\rangle]$ helyett $\mathbb{E}\Big[[\![\alpha\langle u|X\rangle]\!]_0^1\Big]$-t vesszük, azt már elkódolhatjuk amplitúdókba és az előző lemmák miatt viszont a lenullázott értékekből viszont nem lesz sok (és így a várható értéket nem befolyásolják jelentősen).

Tehát hasonlóan, ahogy Probléma 2.)-ben láttuk most olyan fázisorákulumot szeretnénk, amire

$$\frac{1}{\sqrt{|G|}}\sum_{u\in G}|u\rangle \longrightarrow \frac{1}{\sqrt{|G|}}\sum_{u\in G}e^{i\mathbb{E}\left[[\![\alpha\langle u|X\rangle]\!]_0^1\right]}|u\rangle$$

Ha ezt m-szer akarjuk egymás után alkalmazni, úgy mint az előző részben csináltuk akkor $\widetilde{O}(m)$ mintavételt használnánk. Azonban a szerzők megmutatták, hogy ha $||X|| \leq L \leq 1$, akkor lineáris amplitúdó amplifikációt alkalmazva $\widetilde{O}(m\sqrt{L})$-re javítható.

Tétel 3.3.4. (Fázisorákulum kevesebb mintával): Legyen adott az $X : \Omega \to \mathbb{R}^d$ valószínűségi változónk. Továbbá legyenek adottak az $(L, m, \alpha, \varepsilon)$ paraméterek, ekkor a következő függvény megvalósítható:

DirectionalMeanOracle$(X, L, m, \alpha, \varepsilon)$**:**

- INPUT: $||X|| \leq 1$, $L \in (0,1]$, $m \geq \frac{1}{L}$, $\alpha \in (0,1)$, $\varepsilon \in (0,1)$ és $\mathbb{E}\big[||X||\big] \leq L$.
- OUTPUT: $\mathcal{O}_f$, amire * $\mathcal{O}_f|u\rangle|0\rangle \approx e^{im\mathbb{E}\left[[\![\alpha\langle u|X\rangle]\!]_0^1\right]}|u\rangle|0\rangle$
- FELHASZNÁLT MINTA: $\widetilde{O}(m\sqrt{L} \cdot log^2(1/\varepsilon))$

* A kimenetnél "$\approx$" alatt azt értjük, hogy kettes normában ε-becslést adunk rá $\left(1 - \frac{\alpha}{2}\right)$ valószínűséggel.

Biz.: Egyrészt a Probléma 3.) mintájára szátválaszthatjuk X-et két új X^+ és X^- változóra, ahol

$$X(\omega)^+ := \begin{cases} X(\omega) & \text{ha } \alpha\langle u|X(\omega)\rangle > 0 \\ 0 & \text{különben} \end{cases} \qquad X(\omega)^- := \begin{cases} X(\omega) & \text{ha } \alpha\langle u|X(\omega)\rangle < 0 \\ 0 & \text{különben} \end{cases}$$

és elég O_f^+ orákulumot adni, ahol most a kimenetnél $\varepsilon/2$ becslést adunk. Hiszen O_f^- hasonlóan előállítható és végül $\mathcal{O}_f := \mathcal{O}_f^+\mathcal{O}_f^-$ ε becslése lesz a kívánt fázisorákulumnak.

Másrészt azt már láttuk, hogy a $\mathcal{P}$ és $\mathcal{B}_X$ orákulumokból hogyan készítsünk $\mathcal{U}_f^+$ valószínűségi orákulumot, most $f(u) := \mathbb{E}\Big[[\![\alpha\langle u|X\rangle]\!]_0^1\Big]$. Ezt követően azonban vizsgáljuk meg az L paraméter értékét.

- Ha $L \geq \frac{1}{4}$, akkor a felhasznált mintába L csak konstans szorzót jelent, tehát ugyanúgy elkészíthetjük $\mathcal{U}_f^+$-ból $\mathcal{O}_f^+$-t, mint az előző alfejezetben.

- Különben alkalmazzunk lineáris amplitúdó amplifikációt és csak utána konvertáljunk fázisorákulummá.

A javítás ahhoz képest, mint ha nem csinálnánk amplitúdó amplifikációt, azon múlik, hogy ha megnöveljük előre a fázist (az e kitevőjét) akkor már kisebb hatványa eléri a kívánt $e^{im\mathbb{E}\left[[\![\alpha\langle u|X\rangle]\!]_0^1\right]}$-et (persze az amplitúdó amplifikáció alkalmazásának is van költsége, de ezzel együtt is jobban megéri). Vegyük most az amplitúdó amplifikációra vonatkozó következő tételt, [GL20] Theorem 10., Lemma 11. szerint (QSVT alkalmazással):

Tétel 3.3.5. (Lineáris amplitúdó amplifikáció): Legyen $\mathcal{V}$ unitér, $\mathcal{P}$ pedig projekciós operátor. Ekkor $\exists\ \mathcal{V}_{t,\varepsilon}$ unitér, amire

$$\Big|\big|\big|\mathcal{P}\mathcal{V}_{t,\varepsilon'}|\vec{0}\rangle\big|\big| - t\big|\big|\mathcal{P}\mathcal{V}|\vec{0}\rangle\big|\big|\Big| \leq \varepsilon' \qquad \text{ha } t\big|\big|\mathcal{P}\mathcal{V}|\vec{0}\rangle\big|\big| \leq \frac{1}{2}$$

és ez implementálható $\mathcal{V}, \mathcal{V}^\dagger$ és $\mathcal{I} - 2\mathcal{P}$ $O(t \cdot log(1/\varepsilon'))$ alkalmazásával.

Visszatérve a bizonyításhoz, ha $L \leq \frac{1}{4}$, akkor $t = O(1/\sqrt{L})$ és $\varepsilon' = O(\varepsilon/(mL)$ paraméterekkel az előző tétellel $\mathcal{U}_f^+$-ból olyan $\mathcal{W}_f^+ \ : \ |u\rangle|\vec{0}\rangle|\vec{0}\rangle \longrightarrow \sqrt{1-p_u}|u\rangle|\psi_u^0\rangle|0\rangle + \sqrt{p_u}|u\rangle|\psi_u^1\rangle|1\rangle$ valószínűségi orákulumot készíthetünk, amire

$$\left|\sqrt{p_u} - \sqrt{\frac{\mathbb{E}\big[[\![\alpha\langle u|X\rangle]\!]_0^1\big]}{4L}}\right| \leq O\Big(\frac{\varepsilon}{mL}\Big)$$

(feltéve hogy $\mathbb{E}\big[[\![\alpha\langle u|X\rangle]\!]_0^1\big] \leq L$ - de ez a bemeneti $\mathbb{E}\big[||X||\big] \leq L$ feltétel miatt teljesül).

Végül csinálunk egy fázisorákulummá konvertálást $\mathcal{W}_f^+$-ből a 3.1.5. tétel alapján $t = O(mL), \varepsilon' = O(\varepsilon)$ paraméterekkel. Ekkor megmutatható hogy így valóban $\varepsilon/2$ becslést kapunk (lásd. [CHJ22] Proposition 3.2 Proof).

Nézzük még meg a minták számát. Az amplitúdó amplifikációnál $O\Big(\frac{1}{\sqrt{L}} log\big(\frac{mL}{\varepsilon}\big)\Big)$, míg a fázisorákulummá alakításnál $O\big(mL + log(\frac{1}{\varepsilon})\big)$ mintavételt csináltunk. Ez összesen tehát

$$O\left(\frac{log\big(\frac{mL}{\varepsilon}\big)}{\sqrt{L}}\Big(mL + log\Big(\frac{1}{\varepsilon}\Big)\Big)\right) \leq \widetilde{O}\left(m\sqrt{L} \cdot log\Big(\frac{1}{\varepsilon}\Big) + \frac{log^2\big(\frac{1}{\varepsilon}\big)}{\sqrt{L}}\right) \leq \widetilde{O}\left(m\sqrt{L} \cdot log^2\Big(\frac{1}{\varepsilon}\Big)\right)$$

Tétel 3.3.6. (Korlátos várhatóérték becslése): Legyen adott az $X : \Omega \to \mathbb{R}^d$ valószínűségi változónk. Továbbá legyenek adottak az (L, n, δ) paraméterek, ekkor a következő függvény megvalósítható:
BoundedEstimator(X, L, n, δ)**:**

- INPUT: $||X|| \leq 1$, $L \in (0, 1]$, $\delta \in (0, 1)$, $n \geq 1$ és $\mathbb{E}\big[||X||\big] \leq L$.
- OUTPUT: $\tilde{\mu}$, amire $\mathbb{P}\Big[\big|\big|\tilde{\mu} - \mu\big|\big|_\infty \leq \frac{\sqrt{L} log(d/\delta)}{n}\Big] \geq 1 - \delta$
- FELHASZNÁLT MINTA: $\widetilde{O}(n)$

Biz.: Az algoritmus úgy fog kinézni, ahogy várnánk. Egyrészt a 3.3.4.-es tétel alapján előállított fázisorákulumot fogjuk használni, másrészt 3.2. részben leírtuk az algoritmust lényegében.

1. Válasszunk megfelelő paramétereket a rácsfinomsághoz és a fázisorákulum előállításához. Legyenek $\alpha := \frac{1}{log(400\pi n\sqrt{d})}$, $m := 2^{\left\lceil log\left(\frac{8\pi n}{\alpha\sqrt{L}log(d/\delta)}\right)\right\rceil}$ és $\varepsilon := \frac{1}{25}$.
2. $|G\rangle := \frac{1}{m^{d/2}} \sum_{u\in G} |u\rangle$
3. $\mathcal{O}_f :=$ DirectionalMeanOracle$(X, L, m, \alpha, \varepsilon)$ és $|\psi\rangle := \mathcal{O}_f|G\rangle|\vec{0}\rangle$
4. $|\phi\rangle := (QFT_G)^{-1}|\psi\rangle$
5. **FOR** $k = 1, ..., O\big(log(d/\delta)\big)$:
   Ismételjük 2.-4. lépéseket és legyenek $\quad \tilde{v}^{(k)} := meas(|\phi\rangle) \quad \tilde{\mu}^{(k)} := \frac{2\pi}{\alpha} \cdot \tilde{v}^{(k)}$
6. **RETURN** $\tilde{\mu} :=$ medián$(\tilde{\mu}^{(1)}, ..., \tilde{\mu}^{(O(log(d/\delta)))})$ és itt a mediánt koordinátánként véve értjük.

Az eddigiek alapján világos, hogy ha $|\psi\rangle$ egyenlő lenne $|\psi'\rangle := \frac{1}{m^{d/2}} \sum_{u\in G} e^{im\alpha\langle u|\mathbb{E}[X]\rangle}|u\rangle|\vec{0}\rangle$ -val, akkor mint Probléma 2)-ben működne az elemzés. A probléma azzal van, hogy most $\mathbb{E}[\alpha\langle u|X\rangle]$ helyett a levágott $\mathbb{E}\big[[\![\alpha\langle u|X\rangle]\!]_0^1\big]$-gyel számoltunk, a *DirectionalMeanOracle* függvény ráadásul arra is csak egy ε becslést ad.

Belátható viszont ([CHJ22] Theorem 3.3 Proof), hogy ez a két állapot nincs is olyan távol egymástól, egész pontosan:

$$\big|\big||\psi'\rangle - |\psi\rangle\big|\big| \leq \frac{1}{12}$$

Illetve azt állítjuk, hogy az 5. lépésben mért $\tilde{v}^{(k)} \in \mathbb{R}^d$ minden $\tilde{v}_j^{(k)}$ koordinátájára $(j = 1, ..., d)$ teljesül, hogy $\big|\tilde{v}_j^{(k)} - \frac{\alpha}{2\pi}\mathbb{E}[X]_j\big| \leq \frac{4}{m}$. Ihletet véve a [GAW19] Lemma 20 bizonyításából most két fontos dolgot tudunk alkalmazni. Egyrészt most $|\psi'\rangle$-re:

$$|\psi'\rangle = \left(\left(\frac{1}{\sqrt{m}}\sum_{u\in G} e^{im\alpha\langle u_1|\mathbb{E}[X]_1\rangle}|u_1\rangle\right) \otimes ... \otimes \left(\frac{1}{\sqrt{m}}\sum_{u\in G} e^{im\alpha\langle u_d|\mathbb{E}[X]_1\rangle}|u_d\rangle\right)\right)|\vec{0}\rangle$$

és ekkor ha erre az ideális állapotra alkalmazunk inverz Fourier transzformációt, majd a kapott állapotot megmérve ha $\hat{v} = (\hat{v}_1, ..., \hat{v}_d)$ vektort kapunk, akkor (ahogy a hivatkozott bizonyításban

látjuk) tetszőleges K-ra igaz, (és így $K = 4$-re is) hogy:

$$\mathbb{P}\Big[|\hat{v}_i - \frac{\alpha}{2\pi}\mathbb{E}[X]| > \frac{K}{m}\Big] \leq \frac{1}{2(K-1)}$$

A másik fontos dolog az, hogy $2\big|\big||\psi'\rangle - |\psi\rangle\big|\big|$ felülbecsli bármely méréshez tartozó valószínűség különbségét ($|\psi'\rangle$ és $|\psi\rangle$ mérése között) ugyanis:

Ha a $\mathcal{P}$ projekciós operátor szerint mérjük meg $|\psi'\rangle$-t és $|\psi\rangle$-t is, akkor a mérések valószínűségének különbsége:

$$\big|\langle\psi'|\mathcal{P}|\psi'\rangle - \langle\psi|\mathcal{P}|\psi\rangle\big| = \big|Tr\big((|\psi'\rangle\langle\psi'| - |\psi\rangle\langle\psi|)\mathcal{P}\big)\big|$$

ennek a maximuma:

$$\max_{\mathcal{P}}\Big|Tr\big((|\psi'\rangle\langle\psi'| - |\psi\rangle\langle\psi|)\mathcal{P}\big)\Big| = \Big|\Big||\psi'\rangle\langle\psi'| - |\psi\rangle\langle\psi|\Big|\Big|_1 = 2\sqrt{1 - |\langle\psi'|\psi\rangle|^2}$$

és ha $|\psi'\rangle$ és $|\psi\rangle$ által közrezárt szög θ, akkor felhasználva hogy $2sin(\theta/2) = \big|\big||\psi'\rangle - |\psi\rangle\big|\big|$

$$2\sqrt{1 - |\langle\psi'|\psi\rangle|^2} = 2\sqrt{1 - |cos(\theta)|^2} = 2|sin(\theta)| = 4|sin(\theta/2)cos(\theta/2)| \leq 2\big|\big||\psi'\rangle - |\psi\rangle\big|\big|$$

Láttuk tehát, hogy ha az algoritmusban $|\psi'\rangle$ lenne $|\psi\rangle$ helyett, akkor legalább $5/6$ valószínűséggel teljesülne, hogy $|\hat{v}_i - \frac{\alpha}{2\pi}\mathbb{E}[X]| \leq \frac{4}{m}$. Ha az elemzésben (az algoritmusban szereplő) $|\psi\rangle$-re vizsgálódunk, akkor kihasználva hogy $\big|\big||\psi'\rangle - |\psi\rangle\big|\big| \leq \frac{1}{12}$ és hogy egy unitér transzformáció (QFT^{-1}) megtartja a távolságot, azt kapjuk hogy

$$\mathbb{P}\left[\left|\tilde{v}_i - \frac{\alpha}{2\pi}\mathbb{E}[X]\right| \leq \frac{4}{m}\right] \geq \frac{5}{6} - 2\big|\big||\psi'\rangle - |\psi\rangle\big|\big| \geq \frac{2}{3}$$

Így tehát m választása miatt:

$$\big|\big|\tilde{\mu} - \mu\big|\big|_\infty = \big|\big|\tilde{\mu} - \mathbb{E}[X]\big|\big|_\infty \leq \frac{2\pi}{\alpha}\cdot\frac{4}{m} \leq \frac{8\pi}{\alpha}\cdot\frac{1}{\frac{8\pi}{\alpha}\cdot\frac{n}{\sqrt{L}log(d/\delta)}} = \frac{\sqrt{L}log(d/\delta)}{n}$$

A felhasznált minták száma pedig $O(log(d/\delta))\cdot$ # minta(DirectionalMeanOracle) =

$$O\big(log(d/\delta)\big)\cdot\widetilde{O}\big(m\sqrt{L}\cdot log^2(1/\varepsilon)\big)=\widetilde{O}\Big(log(d/\delta)\frac{n}{\alpha\sqrt{L}log(d/\delta)}\sqrt{L}log^2(1/\varepsilon)\Big)=\widetilde{O}\Big(\frac{n}{\alpha}\Big)=\widetilde{O}(n)$$

3.4. A kvantumalgoritmus kiterjesztése korlátlan esetre

Végül az előző fejezet eredményét felhasználva az általános ($||X||\leq 1$ feltétel nélküli) probléma megoldására térünk rá. Az algortimus ötlete az lesz, hogy kvantilisek segítségével exponenciálisan csökkenő levágásait vizsgáljuk a valószínűségi változónak. Látni fogjuk, hogy amennyiben $n\geq d$ az algoritmus közel-optimális, $\widetilde{O}\Big(\frac{\sqrt{d\cdot Tr(\Sigma)}}{n}\Big)$ becslést fog tudni adni $\widetilde{O}(n)$ mintát használva.

Tétel 3.4.1. (Alsó korlát I. [CHJ22] Theorem 3.7): Legyen $n\leq d$, illetve legyen $\sigma>0$ rögzített és jelölje $\mathbf{P}_\sigma$ az X-ek azon családját, aminek a Σ kovarianciamátrixára $Tr(\Sigma)=\sigma^2$. Ekkor bármilyen várhatóérték becslő kvantumalgortimushoz, ami legfeljebb n mintát használ, létezik $X\in\mathbf{P}_\sigma$ melyre:

$$\mathbb{P}\left[||\tilde{\mu}-\mu||\geq\Omega\left(\sqrt{\frac{Tr(\Sigma)}{n}}\right)\right]\geq\frac{2}{3}$$

Tétel 3.4.2. (Alsó korlát II. [CHJ22] Theorem 3.8): Legyen $n>d$, illetve $\sigma>0$ rögzített és jelölje $\mathbf{P}_\sigma$ az X-ek azon családját, aminek a Σ kovarianciamátrixára $Tr(\Sigma)=\sigma^2$. Ekkor bármilyen várhatóérték becslő kvantumalgortimushoz, ami legfeljebb n mintát használ, létezik $X\in\mathbf{P}_\sigma$ melyre:

$$\mathbb{P}\left[||\tilde{\mu}-\mu||\geq\Omega\left(\frac{\sqrt{d\cdot Tr(\Sigma)}}{n}\right)\right]\geq\frac{2}{3}$$

Gondoljunk vissza a magasabb dimenziós klasszikus esetre. A 3.4.1.-es tétel alapján kvantumos esetben nem várhatunk nagyságrendi gyorsítást, amennyiben a dimenzió felülbecsli a megengedett minták számát. Azonban $n>d$ esetben a 3.3.6.-os tétel kiterjesztése polilogaritmikus faktoroktól eltekintve optimális $\sqrt{d/n}$-es javítást fog adni nekünk.

Tétel 3.4.3. (Cornelissen-Hamoudi-Jerbi): Legyen adott $X : \Omega \to \mathbb{R}^d$ diszkrét valószínűségi változó és jelölje Σ a kovarianciamátrixát. Ekkor létezik kvantumalgoritmus, amely a $\mu = \mathbb{E}[X]$ várhatóértéket becsli $\tilde{\mu}$-vel, $\widetilde{O}(n)$ mintát használ és teljesül rá, hogy:

$$\mathbb{P}\left[||\tilde{\mu} - \mu|| \leq \begin{cases} \sqrt{\frac{Tr(\Sigma)}{n}} & \text{ha } n \leq d \\ \frac{\sqrt{d \cdot Tr(\Sigma)}}{n} & \text{ha } n > d \end{cases}\right] \geq \frac{2}{3}$$

A tétel bizonyításához először végtelen normában becsüljünk felül:

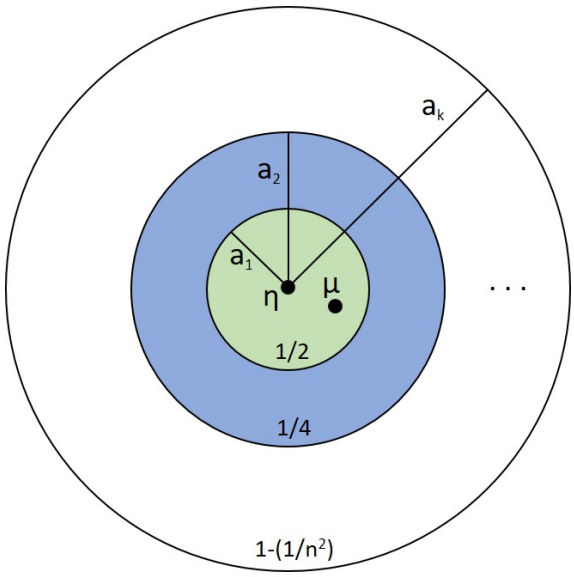
Tétel 3.4.4. (Kiterjesztés korlátlan esetre): Legyen adott az $X : \Omega \to \mathbb{R}^d$ diszkrét valószínűségi változónk. Továbbá legyenek adottak (n, δ) paraméterek, ekkor a következő függvény megvalósítható:

NearOptimalEstimator(X, n, δ)**:**

- INPUT: X tetszőleges, $\delta \in (0, 1)$, $n \geq log(d/\delta)$
- OUTPUT: $\tilde{\mu}$, amire $\mathbb{P}\left[\left|\left|\tilde{\mu} - \mu\right|\right|_\infty \leq \widetilde{O}\left(\frac{\sqrt{Tr(\Sigma)}}{n}\right)\right] \geq 1 - \delta$
- FELHASZNÁLT MINTA: $\widetilde{O}(n)$

Biz.: Az lesz az ötlet hogy kevés mintával először klasszikusan megbecsüljük X várható értékét η-val, ezt követően pedig vesszük azt az η középpontú d-dimenziós gömböt, amibe körülbelül X lehetséges értékeinek fele esik bele és arra csinálunk egy becslést az előző alfejezet alapján. Ezt követően X maradék lehetséges értékeinek a felét becsüljük és így tovább. Az algoritmus során egyrészt exponenciálisan csökken a még meg nem becsült rész, másrészt mindig alkalmazunk egy "normálást" a megfelelő kvantilis szerint hogy alkalmazhassuk a korlátos tételünket.

1. Legyenek $k := O\left(log\left(\frac{n}{log(d/\delta)}\right)\right)$ és $n' := O\left(nk\frac{log(kd/\delta)}{log(d/\delta)}\right)$.
2. Tetszőleges optimális klasszikus várhatóérték becslő algoritmussal $O(log(1/\delta))$ mintával becsüljük meg X-et és legyen a becslés η. Ekkor
$$\mathbb{P}\left[||\eta - \mu|| > \sqrt{Tr(\Sigma)}\right] \leq \delta$$
3. Legyen $Y := X - \eta$ valószínűségi változó.



4. **FOR** $j = 1, ..., k$:
   Legyen a_j az $||Y||$ 2^{-j}-rendű kvantilisének becslése. 1.2.3.: $Q_{||Y||}\left(\frac{1}{2^j}\right) \leq a_j \leq Q_{||Y||}\left(\frac{c}{2^j}\right)$.
   Legyen $Y_j := \frac{1}{a_j}[\![Y]\!]_{a_{j-1}}^{a_j} \quad (a_0 := 0)$
   HA $a_{j-1} = a_j$, akkor $\tilde{\mu}_j := 0$
   KÜLÖNBEN $\tilde{\mu}_j :=$ BoundedEstimator$(Y_j, 2^{-(j-1)}, n', O(\delta/k))$

5. RETURN $\tilde{\mu} := \eta + \sum_{j=1}^{k} a_j \tilde{\mu}_j$

Ha $\tilde{\mu}_Y = \sum_{j=1}^{k} a_j \tilde{\mu}_j$, akkor az algoritmus helyességéhez elég belátni, hogy $||\tilde{\mu}_Y - \mu_Y||_\infty \leq \widetilde{O}\left(\frac{\sqrt{\mathbb{E}[||Y||^2]}}{n}\right)$, ha ez teljesül akkor

$$||\tilde{\mu} - \mu||_\infty = ||\tilde{\mu}_Y - \mu_Y||_\infty \leq \widetilde{O}\left(\frac{\sqrt{\mathbb{E}[||Y||^2]}}{n}\right)$$

felhasználva, hogy nagy valószínűséggel $||\eta - \mu|| \leq \sqrt{Tr(\Sigma)}$

$$\mathbb{E}\left[||Y||^2\right] = \mathbb{E}\left[||X - \eta||^2\right] + ||\eta - \mu||^2 = Tr(\Sigma) + ||\eta - \mu||^2 \leq 2Tr(\Sigma)$$

és így

$$||\tilde{\mu} - \mu||_\infty \leq \widetilde{O}\left(\frac{\sqrt{Tr(\Sigma)}}{n}\right)$$

Lemma 3.4.5.: Minden $j = 1, ..., k$-ra $a_j \leq \sqrt{\frac{\mathbb{E}\left[||Y||^2\right] 2^j}{c}}$, ahol c a kvantilis becslésben szereplő globális konstans.

Biz.: Egyrészt egy Csebisev egyenlőtlenséget alkalmazva:

$$\mathbb{P}\big[||Y|| \geq a_j\big] = \mathbb{P}\big[||Y||^2 \geq a_j^2\big] \leq \frac{\mathbb{E}\big[||Y||^2\big]}{a_j^2}$$

Másrészt a kvantilis 1.2.2. definíciójából adódóan:

$$\mathbb{P}\big[||Y|| \geq a_j\big] \geq \mathbb{P}\Big[||Y|| \geq Q\Big(\frac{c}{2^j}\Big)\Big] \geq \frac{c}{2^j}$$

Tehát:

$$\frac{c}{2^j} \leq \frac{\mathbb{E}\big[||Y||^2\big]}{a_j^2} \quad \Longrightarrow \quad a_j \leq \sqrt{\frac{\mathbb{E}\big[||Y||^2\big] 2^j}{c}}$$

Lemma 3.4.6.: $\big|\big|[\![Y]\!]_{a_k}^{\infty}\big|\big|_\infty \leq \sqrt{\frac{\mathbb{E}\left[||Y||^2\right]}{2^k}}$

Biz.: Ismét a kvantilis definícióját használva:

$$\mathbb{P}\big[||Y|| > a_k\big] \leq \mathbb{P}\big[||Y|| \geq a_k\big] \leq \mathbb{P}\Big[||Y|| > Q\Big(\frac{1}{2^k}\Big)\Big] \leq \frac{1}{2^k}$$

Tehát valóban:

$$\big|\big|[\![Y]\!]_{a_k}^{\infty}\big|\big|_\infty \leq \big|\big|[\![Y]\!]_{a_k}^{\infty}\big|\big| \leq \sqrt{\mathbb{E}\big[||Y||^2\big]\mathbb{P}[||Y|| > a_k]} \leq \sqrt{\frac{\mathbb{E}\big[||Y||^2\big]}{2^k}}$$

Visszatérve tehát a tétel bizonyításához

$$||\tilde{\mu}_Y - \mu_Y||_\infty = \sum_{j=1}^{k} a_j ||\tilde{\mu}_j - \mu_j||_\infty + \left|\left|[\![Y]\!]_{a_k}^{\infty}\right|\right|_\infty$$

Felhasználva 3.4.5. és 3.4.6. lemmákat, illetve hogy az algoritmusban megadott paraméterekkel a korlátos várhatóértékbecslésre $||\tilde{\mu}_j - \mu_j||_\infty \leq \widetilde{O}\left(\frac{\sqrt{2^{-j}}}{n'}\right)$:

$$||\tilde{\mu}_Y - \mu_Y||_\infty \leq \sum_{j=1}^{k} \sqrt{\frac{\mathbb{E}\left[||Y||^2\right]2^j}{c}} \cdot \widetilde{O}\left(\frac{\sqrt{2^{-j}}}{n'}\right) + \sqrt{\frac{\mathbb{E}\left[||Y||^2\right]}{2^k}} =$$

$$= \widetilde{O}\left(\sum_{j=1}^{k} \frac{\sqrt{\mathbb{E}\left[||Y||^2\right]}}{nk}\right) + \frac{\sqrt{\mathbb{E}\left[||Y||^2\right]}}{2^{k/2}} = \widetilde{O}\left(\frac{\sqrt{\mathbb{E}\left[||Y||^2\right]}}{n} + \frac{\sqrt{\mathbb{E}\left[||Y||^2\right]}}{2^{log(n)}}\right) = \widetilde{O}\left(\frac{\sqrt{\mathbb{E}\left[||Y||^2\right]}}{n}\right)$$

Hátravan még hogy belássuk, a felhasznált minták száma valóban $\widetilde{O}(n)$. Az algoritmusban három helyen használunk mintavételt:

1. A klasszikus mintavételnél a 2. lépésben. Ez $O(log(1/\delta))$ minta.
2. A kvantilis becslésnél a 4. lépésben. Ezt sorra meghívjuk a $\frac{1}{2^j}$-rendű kvantilisekre $j = 1, ..., k$. Mivel az 1.2.3.-es tétel szerint $O\left(\frac{log(k/\delta)}{\sqrt{1/2^j}}\right)$ mintavétel szükséges ehhez, így a minták száma itt:

$$O\Big(\sum_{j=1}^{k} 2^{j/2} log(k/\delta)\Big) = O\left(\frac{2^{\frac{k+1}{2}} - 1}{2^{\frac{1}{2}} - 1} log(k/\delta)\right) = \widetilde{O}\big(2^{k/2}\big) = \widetilde{O}\big(2^{log(n)/2}\big) = \widetilde{O}\big(n^{1/2}\big)$$

3. A korlátos becslés hívásakor a 4. lépés végén. Ez pedig $k \cdot \widetilde{O}(n') = \widetilde{O}(n \cdot log^2(n)) = \widetilde{O}(n)$.

Összesen valóban $\widetilde{O}(n)$ mintát használtunk.

Visszatérve a 3.4.3.-es tételhez, $n \leq d$ esetén a klasszikus algoritmus is optimális (a 3.4.1. tétel alapján), $n > d$ esetben pedig a 3.4.4. tétel alapján nagy valószínűséggel:

$$\left|\left|\tilde{\mu} - \mu\right|\right| \leq \sqrt{d}\left|\left|\tilde{\mu} - \mu\right|\right|_\infty \leq \widetilde{O}\Big(\frac{\sqrt{dTr(\Sigma)}}{n}\Big)$$

4. A várható érték becslésének egy pénzügyi alkalmazása

Végül a várhatóérték becslési probléma 2.1. egy alkalmazásáról szeretnék írni. Az amerikai opció problémája egy pénzügyi optimális megállási probléma. A [LBRS23] cikkben ennek a problémának a klasszikus, Monte Carlo módszert használó - a gyakorlatban is alkalmazott - megoldásáról írnak, majd arról hogy mindez hogyan ültethető kvantumos környezetbe és azzal mekkora gyorsítás érhető el. A gyorsítás hátterében az áll, hogy a várhatóérték becslési probléma kvantumos környezetben kevesebb mintavétellel megoldható. A cikk megírásakor Montanaro [Mon15] cikkének eredményét használták, ami $O\big(\frac{\sigma}{\varepsilon} log(1/\delta) log^{3/2}\big(\frac{\sigma}{\varepsilon}\big) loglog\big(\frac{\sigma}{\varepsilon}\big)\big) = \widetilde{O}\big(\frac{\sigma}{\varepsilon}\big)$ mintát használ a várhatóérték ε-becsléséhez. Ehelyett a kvantumos részt most a jobb, Kothari és O'Donnell $O\big(\frac{\sigma}{\varepsilon}\big)$ mintát használó 2.4.2. eredményével írom le.

4.1. Az amerikai opció probléma

A pénzügyi szektorban az úgynevezett amerikai opció az egyik legkézzelfoghatóbb példa optimális megállási idő feladatra. Az opció vételi vagy eladási jogot biztosít egy termékre a vásárlójának egy előre meghatározott áron, de valamilyen jövőbeli időpontban. Amerikai opció esetén, a beváltásának időpontja tetszőleges lehet egy meghatározott futamidőn belül. A befektető vételi jog esetén akkor jár jól, ha a beváltáskor magasabb a piaci értéke a terméknek, mint amikor az opciót megvette.

Például van egy 1 éves futamidejű amerikai típusú vételi opciónk egy portfólióra, ami 100 részvényt tartalmaz. Ez azt jelenti, hogy az opció vételétől kezdve bármelyik nap beválthatjuk, de minden nap változik az összes részvény értéke valamennyit. A profit (vagy ahogy később hivatkozni fogunk rá *payoff*) ekkor az opció megvétele és beváltása között a portfolió értékének megváltozása lesz.

Tegyük fel, hogy a piac lehetséges változását valószínűségi eloszlásokkal modellezzük. A problémát ekkor egy optimális megállási problémaként fogalmazhatjuk meg:

Feladat: Adott egy sztochasztikus folyamat, minden időpontban különböző profittal, mikor érdemes megállni, hogy ennek a profitnak a várható értéke minél nagyobb legyen?

A feladat modellezése:

- Legyen adott egy $(X_t)_{t=0}^T$ Markov-lánc egy Ω mintatérrel és $E \subseteq \mathbb{R}^d$ eseménytérrel. A példában a $t = 0, ..., T$ diszkrét időpontok felelnek meg a napoknak és X_t írja le, hogy a t napon hogyan alakultak a részvényeink.

- Adott még ezen kívül egy $(Z_t)_{t=0}^T$ sztochasztikus folyamat is - amit innentől *payoff* folyamatnak hívunk. A payoff értéke adott t-re csak X_t-től függ: $Z_t := z_t(X_t)$, ahol $z_t \in L_2(E)$.
$\big(L_2(E) :=$ négyzetesen integrálható Borel függvények halmaza.$\big)$

- Jelöljük U_t-vel a következőt:

$$U_t := \begin{cases} Z_T & \text{ha } t = T \\ \max\{Z_t, \mathbb{E}(U_{t+1}|X_t)\} & \text{különben} \end{cases}$$

  Ekkor U_t a várható payoff-ot írja le, ha a t időpontig eljutunk (addig még nem értékesítettük az opciónkat). Ez a felírás azt mutatja, hogy akkor érdemes t-ben értékesíteni, ha a várható payoff később már nem nagyobb. Ha van egy $X'_0, ..., X'_T$ mintánk, akkor a hozzá tartozó $U'_T, U'_{T-1}, ..., U'_0$ értékeket ebben a sorrenben meg tudjuk határozni.

- Mivel $\mathbb{E}(U_{t+1}|X_t)$ X_t mérhető, ezért van olyan Borel mérhető $f_t(x)$, melyre $f_t(X_t) = \mathbb{E}(U_{t+1}|X_t)$, és jelölje u_t:

$$u_t := \begin{cases} z_T & \text{ha } t = T \\ \max\{z_t, f_t\} & \text{különben} \end{cases}$$

  Ekkor $U_t = u_t(X_t)$.

- Legyen $\tau_t := \min\{k \geq t \mid U_k = Z_k\}$.

  A τ_t $(t = 0, ..., T)$ értékek megállási idők. Az U_t megadásából látszik, hogy minden t-hez van olyan u, amire ez fennáll - ez épp az az időpont, amikor már később legfeljebb ugyanakkora payoff-ra számíthatunk. Tehát, ha t-ig eljutottunk, akkor τ_t-ben érdemes megállni.

τ_t felírható az előzőhöz hasonló módon:

$$\tau_t := \begin{cases} T & \text{ha } t = T \\ t \cdot \mathbb{1}\{Z_t \geq \mathbb{E}(Z_{\tau_{t+1}}|X_t)\} + \tau_{t+1} \cdot \mathbb{1}\{Z_t < \mathbb{E}(Z_{\tau_{t+1}}|X_t)\} & \text{különben} \end{cases}$$

- A feladat során nem csak a megállási időt, hanem az elérhető profitot is meg szeretnénk kapni. Tehát a feladat célja az (U_0, τ_0) pár kiszámítása, ezt u_t függvények becslésén keresztül tudjuk elérni.

4.2. A klasszikus LSM módszer

Klasszikus esetben az úgynevezett Least Squares Monte Carlo vagy LSM módszer, amit alkalmazni fogunk. Az algoritmus ötlete, hogy egyrészt az U_t, τ_t értékeket az előbbi megadásukkal, mint dinamikus programmal számoljuk ki. Ezt úgy fogjuk tudni megtenni, hogy (X_t)-ből veszünk N független minta szimulációt:

$$(X_t^{(1)})_{t=0}^T, ..., (X_t^{(N)})_{t=0}^T$$

A $Z_t^{(i)} = z_t(X_t^{(i)})$ -k a hozzájuk tartozó payoff-ok, és $t = T, ..., 0$ fogunk számolni a minta alapján.

Legyen $\{e_{t,k}\}_{k=1}^m$ lineárisan független $L_2(E)$-beli "bázisfüggvények" halmaza minden t-re, röviden e_t-vel jelöljük az m dimenziós vektort: $e_t(\cdot) := \big(e_{t,1}(\cdot), ..., e_{t,m}(\cdot)\big)$. Ezeket az e_t-ket $(t = 0, ..., T)$ hívjuk approximációs sémának.

Jelölés 4.2.1.: ha $a \in \mathbb{R}^m$, akkor $a \cdot e_t := a_1 \cdot e_{t,1} + ... + a_m \cdot e_{t,m}$

Most tehát ebben az $\{e_{t,k}\}_{k=1}^m$ bázisban az $\mathbb{E}(U_{t+1}|X_t)$ -et szeretnénk megbecsülni $\alpha_t \cdot e_t(X_t)$-vel, vegyük az L_2 norma szerinti legközelebbit:

$$\alpha_t = \arg\min_a \mathbb{E}\big((U_{t+1} - a \cdot e_t(X_t))^2\big)$$

Állítás 4.1.: Ha adott $\{e_{t,k}\}_{k=1}^{m}$ $(t = 0, ..., T)$ approximációs séma, akkor legyen A $m \times m$-es kovariancia mátrix melyre $\big(A_t\big)_{i,j} = \mathbb{E}\big(e_{t,i}(X_t)e_{t,j}(X_t)\big)$ és $b_t = \mathbb{E}\big(U_{t+1}e_t(X_t)\big)$. Ekkor $\alpha_t \approx A_t^{-1} \cdot b_t$.

Biz.:

$$\mathbb{E}\big(U_{t+1}\big) \approx \mathbb{E}\big(\alpha_t \cdot e_t(X_t)\big) = \mathbb{E}\big(\alpha_{t,1} \cdot e_{t,1}(X_t) + ... + \alpha_{t,m} \cdot e_{t,m}(X_t)\big)$$

$$\mathbb{E}\big(U_{t+1}\big) \approx \mathbb{E}\big(\alpha_{t,1} \cdot e_{t,1}(X_t)\big) + ... + \mathbb{E}\big(\alpha_{t,m} \cdot e_{t,m}(X_t)\big)$$

Tehát valóban

$$b_t = \mathbb{E}\big(U_{t+1}\big)\mathbb{E}\big(e_t(X_t)\big) \approx \mathbb{E}\big(\alpha_{t,1} \cdot e_{t,1}(X_t)\big)\mathbb{E}\big(e_t(X_t)\big) + ... + \mathbb{E}\big(\alpha_{t,m} \cdot e_{t,m}(X_t)\big)\mathbb{E}\big(e_t(X_t)\big) =$$

$$\mathbb{E}\big(\alpha_{t,1} \cdot e_{t,1}(X_t)\big)\mathbb{E}\big(e_{t,1}(X_t) + ... + e_{t,m}(X_t)\big) + ... + \mathbb{E}\big(\alpha_{t,m} \cdot e_{t,m}(X_t)\big)\mathbb{E}\big(e_{t,1}(X_t) + ... + e_{t,m}(X_t)\big) =$$

$$\alpha_{t_1}\mathbb{E}\big(e_{t,1}(X_t)\sum_{k=1}^{m} e_{t,1}(X_t)\big) + ... + \alpha_{t_m}\mathbb{E}\big(e_{t,m}(X_t)\sum_{k=1}^{m} e_{t,1}(X_t)\big) = \alpha_t \cdot A_t$$

Az algorimtus során mindent becsülni fogunk a minta alapján, tehát például A_t számolásakor $\mathbb{E}\big(e_{t,j}(X_t)e_{t,k}(X_t)\big)$ helyett a minták alapján vett átlagot számoljuk: $\frac{1}{N}\sum_{i=1}^{N}\big(e_{t,j}(X_t^{(i)})e_{t,k}(X_t^{(i)})\big)$. Ilyenkor A_t helyett az $\widetilde{A}_t$ jelölést használjuk (és hasonlóan jelöljük a többi értéket is amikor a minta alapján vett átlaggal becslünk.)

Klasszikus algoritmus:

1. Vegyünk minta szimulációkat: $(X_t^{(i)})_{t=0}^T \quad i = 1, ..., N$
2. Számoljuk ki a $Z_t^{(i)}$ és $e_{t,k}(X_t^{(i)})$ értékeket minden $t = 0, ..., T \ ; \ i = 1, ..., N \ ; \ k = 1, ..., m$ -re
3. Adjuk meg $\widetilde{A}_t$ -t minden $t = 0, ..., T$ -re és az inverzüket is számoljuk ki.
4. $\tilde{u}_T = z_T \qquad \forall \ i = 1, ..., N$

   FOR $t = (T-1), ..., 0$:
   $$\tilde{\alpha}_t = \widetilde{A}_t^{-1} \frac{1}{N} \sum_{i=1}^{N} \tilde{u}_{t+1}(X_{t+1}^{(i)}) \cdot e_t(X_t^{(i)})$$
   $$\tilde{u}_t := \max\{z_t, \tilde{\alpha}_t \cdot e_t\}$$
5. **RETURN**
   $$\widetilde{U}_0 = \frac{1}{N} \sum_{i=1}^{N} \tilde{u}_0(X_0^{(i)})$$

4.3. Kvantumalgoritmus az amerikai opcióra

A kvantumalgoritmusban a várhatóérték becslést szubrutinként használjuk, hogy megbecsüljünk néhány értéket, melyek (X_t)-től függnek. Jelöljük $\hat{N}$-pal a kvatnumos esetben szükséges mintavételek számát, tudjuk tehát hogy $\hat{N} = O(\sigma/\epsilon)$.

Mint láttuk a klasszikus esetben egy mintavételnek tekintettünk egy teljes $(X_t^{(i)})_{t=0}^T$ mintaszimulációt. Ennek a kvantumos megfelelője tehát 2.3. alapján az a $\mathcal{P}$ és $\mathcal{B}_X$ lesz, amire

$$\mathcal{B}_X\mathcal{P}|0\rangle = \sum \sqrt{p(x)}|x\rangle \ , \text{ ahol } p(x) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2|X_1 = x_1)...\mathbb{P}(X_T = x_T|X_{T-1} = x_{T-1})$$

Alkalmazásukkal tehát a Markov-lánc lehetséges realizációinak szuperpozícióját kaphatjuk (melyet ha megmérünk akkor az eloszlás szerinti mintavételt kapunk).

Kvantumalgoritmus

1. Becsüljünk meg minden $\mathbb{E}[e_{t,k}(X_t)e_{t,l}(X_t)]$ várhatóértékét ($\hat{N}$ mintával) $k,l = 1,...,m$
2. Ez alapján határozzuk meg $\widetilde{A}_t$ -ket, és az inverzüket is
3. $\tilde{u}_T := z_T$
4. **FOR** $t = (T-1),...,0$:
   Becsüljük meg $\mathbb{E}[\tilde{u}_{t+1} \cdot e_{t,k}(X_t)]$ értékeket ($k = 1,...,m$) és határozzuk meg $\tilde{b}_t$ -t
   $\tilde{\alpha}_t = \widetilde{A}_t^{-1} \cdot \tilde{b}_t$
   $\tilde{u}_t := \max\{z_t, \tilde{\alpha}_t \cdot e_t\}$
5. **RETURN** $\tilde{u}_0(\mathbb{E}[X_0])$

Futásidő összehasonlítás:

Jelölje $\mathbf{T}_{samp}$ azt az időt, amit egy mintavételre kell fordítani.

Klasszikus esetben N mintavételt csinálunk ez $O(\mathbf{T}_{samp} \cdot N)$, majd $Z_t^{(i)}$ és $e_{t,k}(X_t^{(i)})$ kiszámolása ugyan még megy $O(N \cdot T \cdot m)$ időben, de $\widetilde{A}_t$ kiszámolásához kell $O(N \cdot T \cdot m^2)$. Az $\widetilde{A}_t$ inverzének kiszámolása $O(T \cdot m^\omega)$, végül az $\tilde{\alpha}_t$ értékek kiszámolása $O(N \cdot T \cdot m)$, így a futásidő:

$$O(\mathbf{T}_{samp} \cdot N + N \cdot T \cdot m^2 + T \cdot m^\omega)$$

Kvantumos esetben pedig az $\mathbb{E}[e_{t,k}(X_t)e_{t,l}(X_t)]$ értékek becslésére $O(\mathbf{T}_{samp} \cdot \hat{N} \cdot T \cdot m^2)$ idő kell. Az $\widetilde{A}_t$ inverzeknek kiszámolása most is $O(T \cdot m^\omega)$. A $\tilde{b}_t$ becslések $O(\mathbf{T}_{samp} \cdot \hat{N} \cdot T \cdot m)$ időben mennek. Végül az utolsó lépésben is csinálunk egy várhatóértékbecslést, de az csak $O(\mathbf{T}_{samp} \cdot \hat{N})$. A futásidő tehát:

$$O(\mathbf{T}_{samp} \cdot \hat{N} \cdot T \cdot m^2 + T \cdot m^\omega)$$

Meggondolandó, hogy a két algoritmusban lényegében ugyanazon lépéseket és becsléseket csináljuk (például $\widetilde{A}_t$-t klasszikus esetben átlaggal, kvantumos esetben az optimális Kothari-O'Donnell módszerrel becsüljük.). Így ha ugyanolyan pontosságot akarunk elérni, mivel tudjuk hogy egydimenziós

valószínűségi változót ugyanolyan pontossággal négyzetesen kevesebb mintával tudunk becsülni kvantumos esetben mint klasszikusban, így $\hat{N} := \sqrt{N}$. Látjuk továbbá, hogy klasszikus esetben elég volt csak az algoritmus elején venni mintákat, azokat újra felhasználtuk ellentétben a kvantumos esettel. Illetve ha feltesszük hogy egy mintavétel konstans időben megtehető akkor a kvantumos algortimus $O(\sqrt{N} \cdot T \cdot m^2 + T \cdot m^\omega)$ a klasszikus $O(N \cdot T \cdot m^2 + T \cdot m^\omega)$ helyett. Tehát ha kis időtartamra nem túl nagy approximációs sémával dolgozva tekintjük az amerikai opció problémát (azaz $m << N$ és $T << N$), akkor négyzetes gyorsítást érhetünk el.

Felhasznált irodalom

[CHJ22] Arjan Cornelissen, Yassine Hamoudi, and Sofiene Jerbi. Near-optimal quantum algorithms for multivariate mean estimation. *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, 2022. arXiv: 2111.09787 34, 42, 44, 45, 47

[dW19] Ronald de Wolf. *Quantum Computing: Lecture Notes*. QuSoft, CWI and University of Amsterdam, 2019. arXiv: 1907.09415 6

[GAW19] András Gilyén, Srinivasan Arunachalam, and Nathan Wiebe. Optimizing quantum optimization algorithms via faster quantum gradient computation. *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2019. arXiv: 1711.00465 45

[Gil19] András Gilyén. *Quantum singular value transformation & its algorithmic applications*. PhD thesis, University of Amsterdam, 2019. 35, 36

[GL20] András Gilyén and Tongyang Li. Distributional property testing in a quantum world. *Proceedings of the 11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, 2020. arXiv: 1902.00814 43

[Ham21] Yassine Hamoudi. *Quantum Algorithms for the Monte Carlo Method*. PhD thesis, Université de Paris, 2021. 8, 20

[KO23] Robin Kothari and Ryan O'Donnell. Mean estimation when you have the source code; or, quantum Monte Carlo methods. *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1186–1215, 2023. arXiv: 2208.07544 16, 20, 29, 31, 32

[LBRS23] Alessandro Luongo, Jinge Bao, Patrick Rebentrost, and Miklos Santha. Quantum algorithm for stochastic optimal stopping problems with applications in finance. *Working paper, National University of Singapore*, 2023. arXiv: 2111.15332 52

[LM19] Gábor Lugosi and Shahar Mendelson. Mean estimation and regression under heavy-tailed distributions: A survey. *Foundations of Computational Mathematics*, pages 1145–1190, 2019. arXiv: 1906.04280 18

[Min15] Stanislav Minsker. Geometric median and robust estimation in Banach spaces. *Bernoulli*, pages 2308–2335, 2015. arXiv: 1308.1334 35

[Mon15] Ashley Montanaro. Quantum speedup of Monte Carlo methods. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 2015. arXiv: 1504.06987 52

[OT01] Tatsuaki Okamoto and Keisuke Tanaka. Graph Non-Isomorphism Has a Succinct Quantum Certificate. *Tokyo Institute of Technology. Department of Information Sciences*, 2001. 19

[vA21] Joran van Apeldoorn. Quantum probability oracles & multidimensional amplitude estimation. *16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021)*, 2021. 37