

TDK-DOLGOZAT

Általános irányú Pauli csatorna tomográfiája

Virosztek Dániel, matematikus MSc hallgató

Konzulensek:

Dr. Hangos Katalin
tudományos tanácsadó
Folyamatirányítási Kutatócsoport
MTA SZTAKI

Dr. Ruppert László
Analízis Tanszék
BME TTK



BME
2012

Tartalomjegyzék

1. Bevezetés	3
1.1. Kvantum állapot, kvantum mérés	4
1.2. Pauli csatorna a qubit esetben	5
1.3. Feladatkitűzés	7
2. Az egyszerűsített eset	8
2.1. A becslési eljárás tényezői	9
2.2. A csatornamátrix becslése	10
2.3. A csatornaparaméterek becslése	11
3. A csatornabecslés hatékonyságát leíró mennyiségek	13
3.1. Alapozó lemmák	13
3.2. Nemlineáris paraméterbecslések közelítése	17
3.3. A $\phi = 0$ paraméterű csatorna vizsgálata	19
4. Optimalizálás az egyszerűsített esetben	20
4.1. A csatornamátrixot legjobban becslő eljárás	20
4.2. A kontrakcióparamétereket legjobban becslő eljárás	21
4.3. A szögparamétert legjobban becslő eljárás	22
5. Általános qubit Pauli csatorna tomográfiája	24
5.1. Az input qubitok és a mérési irányok kiválasztása	24
5.2. Az output qubitok Bloch-koordinátáinak becslése	25
5.3. A csatornamátrix becslése	25
5.4. A csatornaparaméterek becslése a csatornamátrixra vonatkozó becslésből	26
6. A paraméterbecslés hatékonysága	27
6.1. A szög- és kontrakcióparaméterek becslésének linearizálása	28
6.2. A csatornamátrix leghatékonyabb becslése	31
6.3. A kontrakcióparaméterek leghatékonyabb becslése	31
6.4. A szögparaméterek leghatékonyabb becslése	32
7. Általános irányú Pauli csatorna prímmhatvány-szintű rendszere	33
7.1. Kölcsonösen torzítatlan bázisrendszerek	34
7.2. Kvázi-merőleges részalgebrák kapcsolata kölcsonösen torzítatlan bázisrendszerekkel	35
7.3. Pauli csatorna általánosított szögparaméterrel, prímmhatvány-szintű rendszerre	36
8. Összefoglalás	38

1. Bevezetés

Modern IT-technológiák széles körében hasznosíthatóak különböző kvantum jelenségek, ezért ezen jelenségek pontos leírása kucsfontosságú. Egy széles körben vizsgált jelenség a kvantum csatorna, amely a kvantum rendszer állapotát alakítja át. A kvantum csatornák paraméterbecslésének (más szóval a *csatornatomográfiának*) jelentős szerepe van a kvantum információfeldolgozásban.

A kvantum információelméletben felmerülő matematikai problémák egy jelentős része paraméterbecslések hatékonyságvizsgálatával kapcsolatos. Egy becslési eljárás hatékonyságát többféle mennyiség jellemzi, a leggyakoribb a *Fisher-információ* és a klasszikus valószínűségi mennyiségek vizsgálata [1, 2]. Az említett célfüggvények analízisére sokféle matematikai eszköz alkalmas; a klasszikus szélsőérték-keresés éppúgy, mint például a lineáris algebra.

A csatornatomográfia direkt módszere kellően általánosan röviden leírható: ismert kvantum állapotokra (*input* állapotokra) hattatjuk a csatornát, majd az így keletkezett *output* állapotokat mérések segítségével megbecsüljük. Az input állapotok és a becült output állapotok összehasonlítása a csatornára vonatkozó információt ad.

Hatékony csatornatomográfiához kísérlettervezésre van szükség. A kísérletterv tartalmazza az optimális input állapotokat, az optimális mérési eljárást, mellyel az output állapotokat mérjük, valamint egy hatékony becslést, mellyel a mért adatokból a csatornát becsljük.

A kvantummechanikában a mérés nem determinisztikus, valószínűségi jelleggel bír [3, 4], így az input kvantum rendszernek több, egymástól független kópiájára van szükség, a becslések pedig a különböző kópiák vizsgálatából származó adatokból készített statisztikák.

A csatornatomográfia jól megalapozott elmélettel bír, a tomográfias eljárások kimerítő leírása olvasható [5]-ben. A csatornák egy széles osztályát alkotják a *Pauli csatornák*, melyek elméleti és gyakorlati szempontból is fontosak. A Pauli csatornák tomográfiájának bőséges irodalma van, azonban a téma nehézsége miatt a dolgozatok általában speciális esetekkel foglalkoznak. Ezek a speciális esetek leggyakrabban ismertnek és adottnak tekintik a Pauli csatorna kontrakciós irányait, és csak az egyes irányokban történő kontrakció nagyságát becslük.

Jelen dolgozatban konstruálunk egy paraméterbecslési sémát, amely *qubit* Pauli csatorna becslésére alkalmas. A cél a bemutatott becslési eljárások közül az optimális módszer kiválasztása. Ennek érdekében valószínűségi jellegű célfüggvényeket definiálunk, majd az esetek többségében analitikus módszerekkel megoldjuk az optimalizálási feladatot. Abban az esetben, amikor az analitikus optimalizálás nem vezet eredményre, numerikus vizsgálatok alapján fogalmazunk meg sejtéseket. Végül prímmhatvány-szintű rendszereken ható *általánosított* Pauli csatornákkal (lásd [6]) foglalkozunk: a maximális kommutatív részalgebraikkal adott csatornákra kiterjesztjük a szögparaméterezést. Az általánosított Pauli csatorna fogalma a közelmúlt tudományos eredménye, ezért ennek a csatornafajtának eddig nem definiálták a kontrakciós irányait, így a javasolt szögparaméterezés egy kezdeti lépés lehet az ilyen csatornák tomográfiájában.

1.1. Kvantum állapot, kvantum mérés

Az alábbiakban a kvantum információelmélet alapfogalmainak egy rövid áttekintését adjuk a [3, 4] művek alapján.

A véges kvantum rendszerek állapotait *sűrűségi mátrixok* reprezentálják, vagyis olyan ρ $n \times n$ -es mátrixok, melyek kielégítik az alábbi feltételeket:

$$\text{Tr}(\rho) = 1 \quad (1.1)$$

$$\rho \geq 0. \quad (1.2)$$

Az n -szintű rendszer állapotterét $\mathcal{S}(\mathbb{C}^n)$ -nel jelöljük. Az $n \times n$ -es önadjungált mátrixok tere ($\mathbf{M}_n^{s.a.}(\mathbb{C})$) n^2 dimenziós vektortér \mathbb{R} felett, mely az $\langle x, y \rangle := \text{Tr}x^*y$ skaláris szorzattal Hilbert-tér. Ennek a térnek rögzíthetjük egy ortogonális bázisát úgy, hogy az egyik bázisvektor az identitás legyen: $\mathcal{F} = \{F_0 = I, F_1, F_2, \dots, F_{n^2-1}\}$. Ekkor a $\rho \in \mathcal{S}(\mathbb{C}^n)$ állapot kifejezhető az \mathcal{F} bázis szerint:

$$\rho = \frac{1}{n} (x_0 I + x_1 F_1 + \dots + x_{n^2-1} F_{n^2-1}). \quad (1.3)$$

A $\text{Tr}(\rho) = 1$ feltétel miatt $x_0 = 1$, az $(x_1, x_2, \dots, x_{n^2-1}) \in \mathbb{R}^{n^2-1}$ vektort a ρ állapot *Bloch-vektorának* hívjuk. Tehát $\mathbf{M}_n^{s.a.}(\mathbb{C})$ egy (I -t tartalmazó, merőleges) bázisának rögzítése egy $\mathcal{S}(\mathbb{C}^n) \rightarrow \mathbb{R}^{n^2-1}$ beágyazást indukál. Egy ilyen beágyazás az állapotter ábrázolásának fontos eszköze, az inverzét *Bloch-paraméterezésnek* nevezzük.

A kvantummechanikában a mérés valószínűségi természetű. A mérhető mennyiségek, az *obszervábilisok* önadjungált mátrixok. Tehát az obszervábilisok felírhatók spektrális alakban. Ha $B \in \mathbf{M}_n^{s.a.}(\mathbb{C})$, akkor

$$B = \sum_j \lambda_j P_j, \quad (1.4)$$

ahol λ_j -vel a B különböző sajátértékeit jelöljük, P_j pedig a λ_j -hez tartozó sajátaltérre való vetítés. A mérés lehetséges kimenetelei a B sajátértékei, a λ_j sajátérték $\text{Tr}\rho P_j$ valószínűséggel adódik, ahol ρ a rendszer állapota.

A kvantum mérés megváltoztatja a rendszer állapotát. Ha a mérés kimenetele λ_j , akkor a mérés után a

$$\rho_j = \frac{P_j \rho P_j}{\text{Tr} P_j \rho P_j}$$

sűrűségi mátrix írja le a rendszer állapotát.

Egy adott obszervábilishez tartozó spektrális projekciók $\{P_j\}_{j \in J}$ rendszerére teljesül, hogy

$$(\forall j \in J) P_j \geq 0; \sum_{j \in J} P_j = I. \quad (1.5)$$

Általában, ha korlátos operátorok egy $\{P_j\}_{j \in J}$ rendszere teljesíti az (1.5)-ben előírt feltételeket, akkor $\{P_j\}_{j \in J}$ -t *pozitív operátor értékű mértéknek* (röviden POVM-nek) hívjuk.

Hasonlóan az obszervábilishoz, a POVM is kvantum mérést reprezentál. A mérés lehetséges kimenetelei ez esetben a POVM elemei. Ha M a POVM egy eleme, akkor annak a valószínűsége, hogy M -et mérjük, éppen $\text{Tr}\rho M$, ahol ρ a rendszer állapota. A POVM-ek egy fontos speciális osztályát alkotják a kételemű POVM-ek. Minden kételemű POVM $\{M, I - M\}$ alakú, ahol $0 \leq M \leq I$. Még speciálisabb esetben, ha M projekció, akkor az $\{M, I - M\}$ POVM-et *Neumann-mérésnek* nevezzük.

Jelölés. A dolgozatban \mathbb{C}^n Hilbert-tér standard bázisának elemeit a $|0\rangle, \dots, |n-1\rangle$, a duális tér megfelelő elemeit a $\langle 0|, \dots, \langle n-1|$ szimbólumokkal is jelöljük.

1.2. Pauli csatorna a qubit esetben

A legegyszerűbb kvantum rendszer a kvantum bit (röviden: *qubit*), melynek Hilberttere $\mathcal{H} = \mathbb{C}^2$, így az állapotokat 2×2 -es sűrűségi mátrixok reprezentálják. A 2×2 -es önadjungált mátrixok terét a *Pauli bázis* feszíti ki; ha $C \in \mathbf{M}_2^{s.a.}(\mathbb{C})$, akkor

$$C = \frac{1}{2}(x_0\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3), \quad (1.6)$$

ahol

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

a Pauli mátrixok, $x_0, x_1, x_2, x_3 \in \mathbb{R}$. Könnyen ellenőrizhető, hogy C pontosan akkor lesz sűrűségi mátrix ha $x_0 = 1$, $x_1^2 + x_2^2 + x_3^2 \leq 1$.

Tehát a Bloch-paraméterezés ezúttal egy nagyon hasznos geometriai szemléletet ad: az állapotteret azonosíthatjuk az \mathbb{R}^3 tér (zárt) egységgömbjével.

A kvantum csatorna a rendszer állapotterén ható nyomtartó, teljesen pozitív leképezés. A Pauli csatornák a kvantum csatornák egy jól ismert családját alkotják a qubit esetben. Heurisztikusan, a Pauli csatorna olyan leképezés, amely az állapottér bizonyos irányában összehúzó.

1. Definíció (Qubit Pauli csatorna). *Legyen $\{v_1, v_2, v_3\}$ tetszőleges ortonormált bázis $\mathbf{M}_2^{s.a.}(\mathbb{C})$ $\sigma_0 = I$ -ra merőleges alterén. Legyenek $\lambda_1, \lambda_2, \lambda_3$ valós számok, melyekre*

$$1 \pm \lambda_3 \geq |\lambda_1 \pm \lambda_2| \quad (1.7)$$

teljesül. Ekkor az

$$\mathcal{E} : \mathcal{S}(\mathbb{C}^2) \rightarrow \mathbf{M}_2^{s.a.}(\mathbb{C}); \quad \rho = \frac{1}{2} \left(I + \sum_{i=1}^3 \theta_i v_i \right) \mapsto \mathcal{E}(\rho) = \frac{1}{2} \left(I + \sum_{i=1}^3 \lambda_i \theta_i v_i \right) \quad (1.8)$$

leképezést qubit Pauli csatornának nevezzük. Az $\{\frac{1}{2}(I + tv_i) : t \in \mathbb{R}\} \subset \mathbf{M}_2^{s.a.}(\mathbb{C})$ affin altereket ($i \in \{1, 2, 3\}$) csatornairányoknak, a $\lambda_1, \lambda_2, \lambda_3$ számokat kontrakcióparamétereknek hívjuk.

Az imént definiált Pauli csatorna nyomtartó, hiszen az identitásra merőleges altéren hat. A leképezés teljesen pozitív tulajdonságát az (1.7) feltétel garantálja [6]. Tehát valójában $\text{Range}(\mathcal{E}) \subset \mathcal{S}(\mathbb{C}^2) \subset \mathbf{M}_2^{s.a.}(\mathbb{C})$. Érdekes megjegyezni, hogy az (1.7) feltétel a három változójában szimmetrikus, és pontosan akkor teljesül, ha $(\lambda_1, \lambda_2, \lambda_3) \in \mathbf{Conv}((1, 1, 1), (1, -1, -1), (-1, 1, -1), (-1, -1, 1)) \subset \mathbb{R}^3$, ahol \mathbf{Conv} a konvex burkot jelöli. Így többek között $-1 \leq \lambda_i \leq 1$, $i \in \{1, 2, 3\}$. Az általánosság megszorítása nélkül feltehetjük (és mostantól fel is tesszük), hogy $\lambda_1 \geq \lambda_2 \geq \lambda_3$.

A Pauli csatorna hatása szemléletesebbé válik, ha az állapotokra mint Bloch vektorokra tekintünk. A már említett Bloch-paraméterezés a

$$\rho : \overline{\mathbf{B}}^3 \rightarrow \mathcal{S}(\mathbb{C}^2); \quad \theta = (\theta_1, \theta_2, \theta_3) \mapsto \rho(\theta) = \frac{1}{2}(I + \theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3) \quad (1.9)$$

bijekció, ahol $\overline{\mathbf{B}}^3 \subset \mathbb{R}^3$ a zárt egységgömb, amit ebben a kontextusban Bloch-gömbnek is szokás nevezni. Megmutatjuk, hogy a Bloch-gömb modellben a Pauli csatornáknak megfelelő leképezések jól karakterizálhatók.

2. Definíció (Csatornamátrix). *Az $A : \overline{\mathbf{B}}^3 \rightarrow \overline{\mathbf{B}}^3$ leképezést az \mathcal{E} Pauli csatorna csatornamátrixának nevezzük, ha*

$$\mathcal{E} \circ \rho = \rho \circ A. \quad (1.10)$$

1. Lemma. *Minden Pauli csatornához egyértelműen létezik csatornamátrix, mely az \mathbb{R}^3 tér egy lineáris leképezése. A csatornamátrixok ortogonális bázisban diagonalizálhatók.*

Bizonyítás. A ρ paraméterezés bijektív, így az (1.10) defináló egyenlőség a

$$\rho^{-1} \circ \mathcal{E} \circ \rho = A \quad (1.11)$$

alakban írható, ez az alak mutatja A létezését és egyértelműségét.

Mivel $\{v_j\}_{j=1}^3$ ortonormált bázis, ezért a

$$\underline{\sigma} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix}, \quad \underline{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \quad (1.12)$$

jelöléseket bevezetve

$$\underline{\sigma} = R\underline{v} \quad (1.13)$$

valamely $R \in \mathbf{O}(3, \mathbb{R})$ mátrixra.

Az $r_{ij} = [R]_{ij}$ jelöléssel élve:

$$\begin{aligned} \mathcal{E}(\rho(\theta)) &= \mathcal{E}\left(\frac{1}{2} \left(I + \sum_{i=1}^3 \theta_i \sigma_i\right)\right) = \mathcal{E}\left(\frac{1}{2} \left(I + \sum_{i=1}^3 \theta_i \sum_{j=1}^3 r_{ij} v_j\right)\right) = \\ &= \mathcal{E}\left(\frac{1}{2} \left(I + \sum_{j=1}^3 \left(\sum_{i=1}^3 \theta_i r_{ij}\right) v_j\right)\right) = \frac{1}{2} \left(I + \sum_{j=1}^3 \lambda_j \left(\sum_{i=1}^3 \theta_i r_{ij}\right) v_j\right) = \\ &= \frac{1}{2} \left(I + \sum_{j=1}^3 \lambda_j \left(\sum_{i=1}^3 \theta_i r_{ij}\right) \left(\sum_{k=1}^3 r_{kj} \sigma_k\right)\right) = \frac{1}{2} \left(I + \sum_{k=1}^3 \left(\sum_{i=1}^3 \left(\sum_{j=1}^3 \lambda_j r_{ij} r_{kj}\right) \theta_i\right) \sigma_k\right). \end{aligned} \quad (1.14)$$

Vagyis

$$\left((\rho^{-1} \circ \mathcal{E} \circ \rho)(\theta)\right)_k = \sum_{i=1}^3 \left(\sum_{j=1}^3 \lambda_j r_{ij} r_{kj}\right) \theta_i. \quad (1.15)$$

Ebből az alakból látható, hogy A lineáris, azaz A egy 3×3 -as valós mátrix, sőt, a mátrixelemek is leolvashatók:

$$[A]_{ki} = \sum_{j=1}^3 \lambda_j r_{ij} r_{kj} = \sum_{j=1}^3 [R]_{kj} \lambda_j [R^T]_{ji}. \quad (1.16)$$

Tehát

$$A = R\Lambda R^T, \quad (1.17)$$

ahol

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

□

Legyen $\mathbf{E} = \{\text{Diag}(\varepsilon_1, \varepsilon_2, \varepsilon_3) : \varepsilon_i \in \{1, -1\} \forall i \in \{1, 2, 3\}\}$. Ekkor \mathbf{E} az $\mathbf{O}(3, \mathbb{R})$ csoport egy (\mathbb{Z}_2^3 -mal izomorf és nem normális) részcsoportha. Könnyen látható, hogy $R_1 \Lambda(\lambda_1, \lambda_2, \lambda_3) R_1^T \equiv R_2 \Lambda(\lambda_1, \lambda_2, \lambda_3) R_2^T$ pontosan akkor, ha $R_2 = R_1 E$ valamely $E \in \mathbf{E}$ mátrixra ($R_1, R_2 \in \mathbf{O}(3, \mathbb{R})$). Tehát ortogonális mátrixok a csatornaparaméterezés szempontjából pontosan akkor megkülönböztethetetlenek, ha az \mathbf{E} részcsoportha ugyanazon *baloldali mellékosztályában* vannak.

Így tehát az $\mathbf{O}(3, \mathbb{R})/\mathbf{E}$ egy parametrizációját és a fenti lemma eredményét használva paraméterezni tudjuk a csatornamátrixokat.

2. Lemma. *Valamennyi qubit Pauli csatornát reprezentáló A mátrix előáll*

$$A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) = R_z R_y R_x \Lambda R_x^{-1} R_y^{-1} R_z^{-1} \quad (1.18)$$

alakban, ahol

$$R_z(\phi_z) = \begin{pmatrix} \cos \phi_z & -\sin \phi_z & 0 \\ \sin \phi_z & \cos \phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad R_y(\phi_y) = \begin{pmatrix} \cos \phi_y & 0 & -\sin \phi_y \\ 0 & 1 & 0 \\ \sin \phi_y & 0 & \cos \phi_y \end{pmatrix},$$

$$R_x(\phi_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi_x & -\sin \phi_x \\ 0 & \sin \phi_x & \cos \phi_x \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

$0 \leq \phi_z, \phi_y, \phi_x < \pi$, és $\phi_y = \frac{\pi}{2} \Rightarrow \phi_z = 0$, valamint a $\lambda_1 \geq \lambda_2 \geq \lambda_3$ számok kielégítik (1.7)-et.

Ez a paraméterezés szürjektív, de nem bijektív: abban az esetben, ha valamely kontrakciópáraméterek megegyeznek, a csatornamátrix némely szögparaméterétől függetlenné válik. A paraméterbecslési feladat megoldása közben ki fog derülni, hogy a paramétertartomány miként szűkíthető addig, amíg a paraméterezés injektívvé válik. Az eredményt (5.15) rögzíti.

1.3. Feladatkitűzés

A tomográfiás feladat nem más, mint ismeretlen Pauli csatorna *valamennyi* paraméterének becslése. A csatornaparaméterek becslésének rögzített és ismert csatornairányok esetén jelentős irodalma van [5], azonban iránybecsléssel csak néhány dolgozat foglalkozik [1, 2], a csatornairány precíz matematikai definíciója pedig újdonság.

Egy olyan becslési eljárást fogunk mutatni, amely a már említett *direkt* módszerek közé sorolandó. Vagyis az eljárás alapötlete az, hogy kiválasztunk néhány input qubitot, ezeknek vesszük független kópiáit, a kópiák mindegyikére hattatjuk a Pauli csatornát, majd az így keletkezett output qubitok Bloch koordinátáira *mérések* segítségével becsléseket adunk, ezekből a becslésekből pedig a csatornaparaméterekre vonatkozó becsléseket készítünk.

Ennek az alapötletnek a matematikai kidolgozásával foglalkoznak a következő fejezetek. Előbb tekintjük a paraméterbecslési feladat egy egyszerűsítését, melynek lényege, hogy a Pauli csatorna némely paramétereit ismertnek tekintjük, rögzítjük. Így az ismeretlen paraméterek száma csökken, a feladat technikailag egyszerűsödik. Az egyszerűbb probléma elemzése olyan tapasztalatokkal gazdagít, amelyek az általános probléma vizsgálatokor hasznosnak bizonyulnak. Az egyszerűsített eset tárgyalása után rátérünk az eredeti problémára.

2. Az egyszerűsített eset

Tegyük fel, hogy a csatornairányokat kijelölő $\{v_j\}_{j=1}^3 \subset \mathbf{M}_2^{s.a.}(\mathbb{C})$ ortonormált bázis és a $\lambda_1, \lambda_2, \lambda_3$ kontrakcióparaméterek meghatározta \mathcal{E} Pauli csatornáról a következőket tudjuk:

$$v_3 = \frac{1}{\sqrt{2}}\sigma_3 \text{ és } \lambda_3 = 0. \quad (2.1)$$

Tekintsük az állapottér következő részét:

$$\mathcal{S}'(\mathbb{C}^2) = \left\{ \frac{1}{2}(I + \theta_1\sigma_1 + \theta_2\sigma_2) : \theta_1, \theta_2 \in \mathbb{R}, \theta_1^2 + \theta_2^2 \leq 1 \right\}. \quad (2.2)$$

A konstrukcióból adódik, hogy $\rho^{-1}(\mathcal{S}'(\mathbb{C}^2))$ a Bloch-gömb és az e_1, e_2 standard bázisvektorok kifeszítette sík metszete. Tehát $\rho^{-1}(\mathcal{S}'(\mathbb{C}^2)) = \overline{\mathbf{B}}^2 \subset \mathbb{R}^2$. Könnyen ellenőrizhető, hogy

$$\text{Range}(\mathcal{E}) \subset \mathcal{S}'(\mathbb{C}^2), \text{ és } \text{Range}(\mathcal{E}|\mathcal{S}'(\mathbb{C}^2)) = \text{Range}(\mathcal{E}), \quad (2.3)$$

vagyis a leképezés megszorításával a képtér nem csökken. Tehát elegendő az $\mathcal{E}' := \mathcal{E}|\mathcal{S}'(\mathbb{C}^2)$ leképezést vizsgálni, amely a következőképpen írható le:

$$\mathcal{E}' : \mathcal{S}'(\mathbb{C}^2) \rightarrow \mathcal{S}'(\mathbb{C}^2); \rho = \frac{1}{2} \left(I + \sum_{i=1}^2 \theta_i v_i \right) \mapsto \mathcal{E}'(\rho) = \frac{1}{2} \left(I + \sum_{i=1}^2 \lambda_i \theta_i v_i \right). \quad (2.4)$$

Az 1. definíció mintájára az \mathcal{E}' leképezéshez is lehet csatornamátrixot definiálni.

3. Definíció. Az $A : \overline{\mathbf{B}}^2 \rightarrow \overline{\mathbf{B}}^2$ leképezést az \mathcal{E}' egyszerűsített Pauli csatorna csatornamátrixának nevezzük, ha

$$\mathcal{E}' \circ \rho = \rho \circ A. \quad (2.5)$$

Az 1. lemma bizonyítását szinte lépésről lépésre megismételve igazolható, hogy A az \mathbb{R}^2 tér egy ortogonális bázisban diagonalizálható leképezése. Ebből adódik, hogy a 2. lemma analogonja is kimondható az egyszerűsített esetben.

3. Lemma. Bármely egyszerűsített Pauli csatornát reprezentáló A mátrix előáll

$$A(\lambda_1, \lambda_2, \phi) = R(\phi)\Lambda(\lambda_1, \lambda_2)R(\phi)^{-1} \quad (2.6)$$

alakban, ahol

$$R(\phi) = \begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}, \quad \Lambda(\lambda_1, \lambda_2) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

$0 \leq \phi < \pi$, és $\lambda_1 = \lambda_2 \Rightarrow \phi = 0$, valamint a $\lambda_1 \geq \lambda_2, \lambda_3 = 0$ számok kielégítik (1.7)-et.

A paramétertartomány $\lambda_1 = \lambda_2 \Rightarrow \phi = 0$ szűkítése miatt a fenti paraméterezés bijektív.

2.1. A becslési eljárás tényezői

Az \mathcal{E}' egyszerűsített Pauli csatorna tomográfiájához két input qubitot és két különböző Neumann-mérést választunk.

Az input állapotok legyenek $\mathcal{S}'(\mathbb{C}^2)$ -ben, hiszen ez \mathcal{E}' értelmezési tartománya. $\mathcal{S}'(\mathbb{C}^2)$ definíciója alapján az input qubitok \mathbb{R}^2 -beli Bloch-vektorokkal írhatóak le, ezeket jelöljük $\underline{\theta}_1, \underline{\theta}_2$ -vel. Érdemes $\underline{\theta}_1$ -et és $\underline{\theta}_2$ -t merőlegesnek és egységnyi hosszúnak választani [1], [2]. Ezek alapján

$$[\underline{\theta}_1, \underline{\theta}_2] = R(\vartheta), \quad (2.7)$$

ahol R ugyanazt a forgatást jelöli, mint (2.6)-ben, és $\vartheta \in \mathbb{R}$. Tehát a kiválasztott input állapotok egyetlen valós szögparaméterrel jellemezhetők.

Végiggondoljuk, hogy a Neumann-mérések esetében is hasonló a helyzet.

4. Lemma. *A qubit esetben az $\{M, I-M\}$ POVM pontosan akkor nemtriviális Neumann-mérés, ha M egyrangú projekció. Az egyrangú projekciók éppen az egységnyi hosszú Bloch-vektorral reprezentálható sűrűségi mátrixok.*

Bizonyítás. Az állítás első fele nyilvánvaló. Másrészt M egy-rangú projekció $\Leftrightarrow M = M^*$ és M két sajátértéke $\mu_1 = 1, \mu_2 = 0 \Leftrightarrow M = M^*$ és $\text{Tr}M = 1, \text{Det}M = 0$, ez utóbbi 3 tulajdonság pedig jól láthatóan pontosan akkor teljesül, ha M a Bloch-gömb felszínén van. \square

A lemma alapján a Neumann-mérések egységnyi hosszú Bloch-vektorokkal ábrázolhatók. A Neumann-mérés *irányán* a reprezentáns Bloch-vektor irányát értjük.

Mivel \mathcal{E}' egy σ_3 -ra merőleges affin altéren hat, ezért olyan Neumann-méréseket választunk, hogy mindkét POVM mindkét eleme merőleges legyen σ_3 -ra.

5. Lemma. *Ha $M \in \mathbf{M}_2(\mathbb{C})$ és $\langle M, \sigma_3 \rangle = \text{Tr}(M\sigma_3) = 0$, akkor $\langle I-M, \sigma_3 \rangle = \text{Tr}((I-M)\sigma_3) = 0$. Továbbá, ha M sűrűségi mátrix, melyet az \mathbf{m} Bloch-vektor reprezentál, akkor $I-M$ -nek a $-\mathbf{m}$ Bloch-vektor felel meg.*

Bizonyítás. Egyrészt $\text{Tr}((I-M)\sigma_3) = \text{Tr}(\sigma_3) - \text{Tr}(M\sigma_3) = 0 - 0 = 0$. Másrészt $M = \frac{1}{2}I + \frac{1}{2}\sum_{i=1}^3 m_i\sigma_i \Rightarrow I-M = \frac{1}{2}I - \frac{1}{2}\sum_{i=1}^3 m_i\sigma_i$. \square

A lemma alapján elegendő (és persze szükséges) M -et σ_3 -ra merőlegesnek választani, és az is láthatóvá vált, hogy egy egységnyi hosszú Bloch-vektor és az ellentettje ekvivalens Neumann-mérést reprezentál. Ha $M = \frac{1}{2}I + \frac{1}{2}\sum_{i=1}^3 m_i\sigma_i$, akkor $\text{Tr}(M\sigma_3) = m_3$, tehát a Neumann-méréseket ábrázoló vektorokra vonatkozó feltétel egyszerűen $m_3 = 0$.

Tehát a mérések is \mathbb{R}^2 -beli vektorokkal azonosíthatóak, és ismert, hogy a két reprezentáns vektort (jelöljük ezeket $\underline{m}_1, \underline{m}_2$ -vel) érdemes merőlegesnek választani [1]. Tehát

$$[\underline{m}_1, \underline{m}_2] = R(\tau), \quad (2.8)$$

ahol $\tau \in \mathbb{R}$.

2.2. A csatornamátrix becslése

A csatornatomográfia első lépése a (2.5) egyenlettel definiált A csatornamátrix becslése a 2.1. alfejezetben jellemzett input állapotok és mérések segítségével.

6. Lemma. *Ha az $\{M, I - M\}$ Neumann-mérést a $\rho(\theta)$ állapotban lévő rendszeren elvégezzük, akkor*

$$\text{Prob}(M\text{-et mérjük}) = \frac{1}{2}(1 + m \cdot \theta), \quad (2.9)$$

ahol $m = (m_1, m_2, m_3)$ az M -et jelölő Bloch vektor.

Bizonyítás.

$$\begin{aligned} \text{Prob}(M\text{-et mérjük}) &= \text{Tr}(\rho(\theta)M) = \\ &= \text{Tr}\left(\frac{1}{2}(I + \theta_1\sigma_1 + \theta_2\sigma_2 + \theta_3\sigma_3)\frac{1}{2}(I + m_1\sigma_1 + m_2\sigma_2 + m_3\sigma_3)\right) = \\ &= \frac{1}{4}2(1 + m \cdot \theta) = \frac{1}{2}(1 + m \cdot \theta), \end{aligned} \quad (2.10)$$

hiszen a Pauli bázis ortogonális, a bázisvektorok normanégyzete 2. \square

Ezt a lemmát majd a teljes becslési feladat vizsgálatokor is felhasználjuk.

Jelölje $\underline{\theta}_1^*$ és $\underline{\theta}_2^*$ a $\mathcal{E}'(\rho(\underline{\theta}_1))$ és a $\mathcal{E}'(\rho(\underline{\theta}_2))$ output állapotok Bloch-vektorait. Ekkor a csatornamátrix definíciója szerint

$$[\underline{\theta}_1^*, \underline{\theta}_2^*] = A(\lambda_1, \lambda_2, \phi) [\underline{\theta}_1, \underline{\theta}_2]. \quad (2.11)$$

Tehát (2.6) és (2.7) alapján

$$[\underline{\theta}_1^*, \underline{\theta}_2^*] = R(\phi)\Lambda(\lambda_1, \lambda_2)R(\phi)^{-1}R(\vartheta). \quad (2.12)$$

Az első feladat az output állapotok Bloch-koordinátáinak becslése. Ehhez a j . inputnak tekintjük $2N$ kópiáját, ezeken hattatjuk az \mathcal{E}' leképezést. Az így kapott $2N$ output qubiton Neumann-méréseket végzünk, mindkét mérési irányban N -et. N a becslési eljárás fontos paramétere, a továbbiakban *mérésszámnak* nevezzük. Jelöljük N_{ij}^+ -val azon események számát, amikor a j . output állapotban az i . irányban mérve (vagyis az $\{M_i, I - M_i\}$ POVM-et mérve) M_i a mérés kimenetele. Legyen $x_{ij} = \underline{m}_i \cdot \underline{\theta}_j^*$, $X = \{x_{ij}\}_{i,j=1}^2$. A 6. lemmából és a kópiák függetlenségéből következik, hogy

$$N_{ij}^+ = \text{Binom}\left(N, \frac{1 + x_{ij}}{2}\right). \quad (2.13)$$

A binomiális eloszlás tulajdonságaiból adódik, hogy

$$\mathbf{E}(N_{ij}^+) = \frac{N}{2}(1 + x_{ij}), \quad \mathbf{Var}(N_{ij}^+) = N\frac{1 + x_{ij}}{2}\left(1 - \frac{1 + x_{ij}}{2}\right) = \frac{N}{4}(1 - x_{ij}^2). \quad (2.14)$$

Ebből az alakból látszik, hogy x_{ij} -re torzítatlan becslés adható, melynek varianciája $\frac{1}{N}$ -es nagyságrendben lecseng:

$$\hat{x}_{ij} := \frac{2}{N}N_{ij}^+ - 1. \quad (2.15)$$

Bizonyítás.

$$\mathbf{E}(\hat{x}_{ij}) = \frac{2}{N} \frac{N}{2} (1 + x_{ij}) - 1 = x_{ij}, \quad \mathbf{Var}(\hat{x}_{ij}) = \frac{4}{N^2} \frac{N}{4} (1 - x_{ij}^2) = \frac{(1 - x_{ij}^2)}{N}. \quad (2.16)$$

□

Az

$$M = [\underline{m}_1, \underline{m}_2], \quad \Theta = [\underline{\theta}_1, \underline{\theta}_2], \quad \Theta^* = [\underline{\theta}_1^*, \underline{\theta}_2^*], \quad \hat{X} = \{\hat{x}_{ij}\}_{i,j=1}^2 \quad (2.17)$$

jelöléseket bevezetve, definíció szerint

$$X = M^T \Theta^*, \quad (2.18)$$

(2.18)-ba (2.8)-at és (2.12)-t behelyettesítve azt kapjuk, hogy

$$X = R(\tau)^T R(\phi) \Lambda(\lambda_1, \lambda_2) R(\phi)^{-1} R(\vartheta). \quad (2.19)$$

A forgatások szögparaméterei összeadódnak, így

$$X = R(\phi - \tau) \Lambda(\lambda_1, \lambda_2) R(\vartheta - \phi). \quad (2.20)$$

M ortogonális, így (2.18) átrendezhető az $MX = \Theta^*$ alakba, tehát az output qubitok koordinátáira vonatkozó becslést adjuk meg a

$$\hat{\Theta}^* := M \hat{X}. \quad (2.21)$$

egyenlettel. A (2.17)-ben bevezetett jelölésekkel (2.11) a

$$\Theta^* = A(\lambda_1, \lambda_2, \phi) \Theta \quad (2.22)$$

alakot ölti. Ennek az egyenletnek az $A = \Theta^* \Theta^{-1}$ ekvivalens alakja mutatja, hogy mi legyen a csatornamátrixra vonatkozó becslés:

$$\hat{A} := \hat{\Theta}^* \Theta^{-1}, \quad (2.23)$$

vagyis

$$\hat{A} := M \hat{X} \Theta^{-1} = R(\tau) \hat{X} R(\vartheta)^{-1}. \quad (2.24)$$

2.3. A csatornaparaméterek becslése

A csatornaparaméterezés invertálásával kapjuk a csatornamátrixra vonatkozó becslésből a paraméterek (esetünkben λ_1, λ_2 és ϕ) becslését. A paraméterezést a (2.6) egyenlet írja le, amely kézzelfoghatóbbá válik, ha elvégezzük a mátrixszorzást:

$$A(\lambda_1, \lambda_2, \phi) = \begin{pmatrix} \lambda_1 \cos \phi^2 + \lambda_2 \sin \phi^2 & (\lambda_1 - \lambda_2) \sin \phi \cos \phi \\ (\lambda_1 - \lambda_2) \sin \phi \cos \phi & \lambda_1 \sin \phi^2 + \lambda_2 \cos \phi^2 \end{pmatrix}. \quad (2.25)$$

Meg kell tehát konstruálni az $A : \mathcal{D} \rightarrow \mathbf{M}_2(\mathbb{R}); (\lambda_1, \lambda_2, \phi) \mapsto A(\lambda_1, \lambda_2, \phi)$ leképezés ($\mathcal{D} \subset \mathbb{R}^3$ a 3. lemmában meghatározott paramétertartomány) inverzének egy kiterjesztését, vagyis egy olyan

$$T : \mathbf{M}_2(\mathbb{R}) \rightarrow \mathbb{R}^3; \hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi}) \quad (2.26)$$

leképezést, melyre

$$T \circ A = \text{Id}_{\mathcal{D}} \quad (2.27)$$

teljesül. Az inverz kiterjesztése azért lényeges, mert az \hat{A} valószínűségi változó $\text{Range}(A)$ -n kívül eső értékeket is felvehet.

Vezessük be az

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = A, \quad \begin{pmatrix} \hat{a}_{11} & \hat{a}_{12} \\ \hat{a}_{21} & \hat{a}_{22} \end{pmatrix} = \hat{A} \quad (2.28)$$

jelöléseket. T konstrukciójának első lépése, hogy \hat{A} -t szimmetrizáljuk: $\hat{A}_s := \frac{1}{2} (\hat{A} + (\hat{A})^T)$. A csatornamátrixok szimmetrikusak, így a $\text{Range}(A)$ -beli elemeket ez a lépés fixen hagyja. (2.6) mutatja, hogy az $A(\lambda_1, \lambda_2, \phi)$ mátrix sajátértékei a λ_1, λ_2 paraméterek. Ezért a kontrakcióparamétereket becsüljük \hat{A}_s sajátértékeivel:

$$\begin{aligned} \hat{\lambda}_{1,2} &= \frac{1}{2} \left(\text{Tr} \hat{A}'_s \pm \sqrt{(\text{Tr} \hat{A}'_s)^2 - 4 \text{Det} \hat{A}'_s} \right) = \\ &= \frac{1}{2} \left(\hat{a}_{11} + \hat{a}_{22} \pm \sqrt{(\hat{a}_{11} + \hat{a}_{22})^2 - 4 \left(\hat{a}_{11} \hat{a}_{22} - \left(\frac{\hat{a}_{12} + \hat{a}_{21}}{2} \right)^2 \right)} \right) \\ &= \frac{1}{2} \left(\hat{a}_{11} + \hat{a}_{22} \pm \sqrt{(\hat{a}_{11} + \hat{a}_{22})^2 - 4 \hat{a}_{11} \hat{a}_{22} + (\hat{a}_{12} + \hat{a}_{21})^2} \right) \\ &= \frac{1}{2} \left(\hat{a}_{11} + \hat{a}_{22} \pm \sqrt{(\hat{a}_{11} - \hat{a}_{22})^2 + (\hat{a}_{12} + \hat{a}_{21})^2} \right). \end{aligned} \quad (2.29)$$

\hat{A}_s szimmetrikus, ezért \mathbb{R}^2 valamely merőleges bázisában diagonális. Ha $\hat{\lambda}_1 > \hat{\lambda}_2$, akkor egyértelműen létezik $\hat{\phi} \in [0, \pi)$, hogy $\hat{A}_s = R(\hat{\phi}) \Lambda(\hat{\lambda}_1, \hat{\lambda}_2) R(\hat{\phi})^{-1}$. Ha $\hat{\lambda}_1 = \hat{\lambda}_2$, akkor \hat{A}_s skalármátrix, így $\hat{A}_s = R(\hat{\phi}) \Lambda(\hat{\lambda}_1, \hat{\lambda}_2) R(\hat{\phi})^{-1} \forall \hat{\phi} \in [0, \pi)$.

(2.25)-ből leolvasható, hogy ha $\lambda_1 > \lambda_2$ és $\phi \notin \{0, \frac{\pi}{2}\}$, akkor

$$\frac{a_{11} - a_{22}}{a_{12} + a_{21}} = \cot(2\phi) \text{ és } \phi > \frac{\pi}{2} \Leftrightarrow a_{12} = a_{21} < 0. \quad (2.30)$$

Ha $\lambda_1 > \lambda_2$ és $\phi \in \{0, \frac{\pi}{2}\}$, akkor $\phi = 0 \Leftrightarrow a_{11} > a_{22}$, $\phi = \frac{\pi}{2} \Leftrightarrow a_{11} < a_{22}$. Ha $\lambda_1 = \lambda_2$, akkor a paramétertartomány megválasztása miatt $\phi = 0$. Tehát a szögparaméter becslése az alábbi esetszétválasztással adható meg.

1. Ha \hat{A}_s nem diagonális (azaz $\hat{a}_{12} + \hat{a}_{21} \neq 0$), akkor

$$\hat{\phi} = \frac{1}{2} \left(\text{arccot} \left(\frac{\hat{a}_{11} - \hat{a}_{22}}{\hat{a}_{12} + \hat{a}_{21}} \right) + \mathbf{1}\{\hat{a}_{12} + \hat{a}_{21} < 0\} \pi \right), \quad (2.31)$$

ahol arccot értékészletét $(0, \pi)$ -nek választjuk.

2. Ha \hat{A}_s diagonális, akkor $\hat{a}_{11} = \hat{a}_{22} \Rightarrow \hat{\phi} = 0$ (ez a skalármátrix esete). Továbbá $\hat{a}_{11} > \hat{a}_{22} \Rightarrow \hat{\phi} = 0$, $\hat{a}_{11} < \hat{a}_{22} \Rightarrow \hat{\phi} = \frac{\pi}{2}$.

3. A csatornabecslés hatékonyságát leíró mennyiségek

Az előző fejezetben leírt paraméterbecslés hatékonysága függ a becslési eljárás tényezőinek (Neumann-mérések, input qubitok, mérésszám) megválasztásától, és a becslés tárgyától, a Pauli csatornától is. Az input állapotokat és a méréseket a 2.1. alfejezetben bevezetett szögparaméterekkel írjuk le, így egy becslési stratégia a (τ, ϑ, N) paraméterhármasval jellemezhető. A Pauli csatornát a $(\lambda_1, \lambda_2, \phi)$ paraméterek írják le. A becslési stratégia hatékony megválasztása kísérlettervezésre ad lehetőséget.

Egy becslés hatékonyságát természetes módon jellemzi az átlagos négyzetes hibája. Ebben a dolgozatban három ilyen jellegű mennyiséggel foglalkozunk. Az első kettő a szögparaméter illetve a kontrakcióparaméterek becslésének átlagos négyzetes hibája. A harmadik a becslült paraméterekhez tartozó csatornamátrix és a valódi csatornamátrix átlagos négyzetes távolsága. Vezessük tehát be a következő mennyiségeket:

$$g_1(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{E} \left(\text{dist}(\hat{\phi}, \phi)^2 \right) \quad (3.1)$$

$$g_2(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{E} \left((\hat{\lambda}_1 - \lambda_1)^2 + (\hat{\lambda}_2 - \lambda_2)^2 \right) \quad (3.2)$$

$$g_3(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{E} \left(\|A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi}) - A(\lambda_1, \lambda_2, \phi)\|^2 \right), \quad (3.3)$$

ahol $\|\cdot\|$ a Hilbert-Schmidt norma. A paraméterezés szempontjából ekvivalens szögeket azonosítjuk, és a becslés hibáját is ennek megfelelően mérjük. Vagyis

$$\text{dist}(\hat{\phi}, \phi) := \inf\{|\hat{\phi} - (\phi + k\pi)| : k \in \mathbb{Z}\}. \quad (3.4)$$

A paramétertartományt úgy választottuk, és a $\hat{\phi}$ becslést úgy konstruáltuk, hogy $\phi, \hat{\phi} \in [0, \pi)$, ezért valójában

$$\inf\{|\hat{\phi} - (\phi + k\pi)| : k \in \mathbb{Z}\} = \min\{|\hat{\phi} - (\phi + k\pi)| : k \in \{-1, 0, 1\}\}. \quad (3.5)$$

Tehát $\text{dist}(\hat{\phi}, \phi)$ könnyen számolható.

A kísérlettervezési feladat nem más, mint az adott $A(\lambda_1, \lambda_2, \phi)$ egyszerűsített Pauli csatornát leghatékonyabban becsülő stratégia megkeresése. Ha a mérésszámot caeteris paribus növeljük, a paraméterbecslés hatékonysága nő, ezért az optimumot fix N mellett keressük.

Tehát a cél a

$$(\tau, \vartheta) \mapsto g_i(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) \quad (i \in \{1, 2, 3\}) \quad (3.6)$$

függvények minimalizálása adott $(\lambda_1, \lambda_2, \phi, N)$ értékek mellett.

3.1. Alapozó lemmák

A g_1, g_2, g_3 függvények analízise összetett feladat, ezért először segédállításokat foglalkozunk meg.

7. Lemma (Forgatásinvariancia). *Legyen $\delta \in \mathbb{R}$ tetszőleges. Becsüljük az $A_1 = A(\lambda_1, \lambda_2, \phi)$ csatorna mátrixát és paramétereit a τ, ϑ, N paraméterű becslési stratégiával, a becslések legyenek $\hat{A}_{1,s}, \hat{\lambda}_1^1, \hat{\lambda}_2^1, \hat{\phi}^1$. Az $A_2 = A(\lambda_1, \lambda_2, \phi + \delta)$ csatornát becsüljük a $\tau + \delta, \vartheta + \delta, N$ paraméterű eljárással, az így kapott becslések legyenek $\hat{A}_{2,s}, \hat{\lambda}_1^2, \hat{\lambda}_2^2, \hat{\phi}^2$. Ekkor*

$$\hat{A}_{2,s} = R(\delta)\hat{A}_{1,s}R(\delta)^{-1}, \quad (3.7)$$

következésképpen

$$\hat{\lambda}_1^1 = \hat{\lambda}_1^2, \quad \hat{\lambda}_2^1 = \hat{\lambda}_2^2, \quad (3.8)$$

$$\hat{\phi}^2 = \hat{\phi}^1 + \delta \pmod{\pi} \quad (3.9)$$

és

$$\|\hat{A}_{1,s} - A_1\|^2 = \|\hat{A}_{2,s} - A_2\|^2. \quad (3.10)$$

Bizonyítás. (2.20)-ből látszik, hogy

$$X(\phi, \tau, \vartheta, \lambda_1, \lambda_2) = X(\phi + \delta, \tau + \delta, \vartheta + \delta, \lambda_1, \lambda_2). \quad (3.11)$$

(2.13), (2.15) és (2.17) mutatja, hogy a paraméterbecslések forrása, az \hat{X} valószínűségi változó eloszlása csak X -től függ. Tehát a két becslési eljárás esetében \hat{X} azonos eloszlású ($\hat{X}_1 \stackrel{d}{=} \hat{X}_2$). Ezért, az $\hat{X} = \hat{X}_1 = \hat{X}_2$ jelölést bevezetve, (2.24) alapján

$$\hat{A}_2 = R(\tau + \delta)\hat{X}R(\vartheta + \delta)^{-1} = R(\delta)R(\tau)\hat{X}R(\vartheta)^{-1}R(\delta)^{-1} = R(\delta)\hat{A}_1R(\delta)^{-1}. \quad (3.12)$$

Ebből következik, hogy a szimmetrizált becslésekre

$$\hat{A}_{2,s} = R(\delta)\hat{A}_{1,s}R(\delta)^{-1}, \quad (3.13)$$

teljesül, vagyis $\hat{A}_{1,s}$ és $\hat{A}_{2,s}$ hasonlóak. Hasonló mátrixok sajátértékei megegyeznek, így

$$\hat{\lambda}_1^1 = \hat{\lambda}_1^2, \quad \hat{\lambda}_2^1 = \hat{\lambda}_2^2. \quad (3.14)$$

Másrészt

$$\hat{A}_{1,s} = R(\hat{\phi}^1)\Lambda(\hat{\lambda}_1, \hat{\lambda}_2)R(\hat{\phi}^1)^{-1}, \quad (3.15)$$

így

$$\hat{A}_{2,s} = R(\delta)R(\hat{\phi}^1)\Lambda(\hat{\lambda}_1, \hat{\lambda}_2)R(\hat{\phi}^1)^{-1}R(\delta)^{-1} = R(\hat{\phi}^1 + \delta)\Lambda(\hat{\lambda}_1, \hat{\lambda}_2)R(\hat{\phi}^1 + \delta)^{-1}. \quad (3.16)$$

Ebből az eredményből és az

$$\hat{A}_{2,s} = R(\hat{\phi}^2)\Lambda(\hat{\lambda}_1, \hat{\lambda}_2)R(\hat{\phi}^2)^{-1} \quad (3.17)$$

összefüggésből következik, hogy

$$\hat{\phi}^2 = \hat{\phi}^1 + \delta \pmod{\pi}. \quad (3.18)$$

Végül $A_2 = R(\delta)A_1R(\delta)^{-1}$, így

$$\begin{aligned} \|\hat{A}_{2,s} - A_2\|^2 &= \|R(\delta)\hat{A}_{1,s}R(\delta)^{-1} - R(\delta)A_1R(\delta)^{-1}\|^2 = \\ &= \|R(\delta) \left(\hat{A}_{1,s} - A_1 \right) R(\delta)^{-1}\|^2 = \|\hat{A}_{1,s} - A_1\|^2, \end{aligned} \quad (3.19)$$

mert az ortogonális mátrixszal konjugálás nem változtatja a Hilbert-Schmidt normát. \square

A lemma egyszerű következménye, hogy

$$\left(\hat{\lambda}_i^1 - \lambda_i\right)^2 = \left(\hat{\lambda}_i^2 - \lambda_i\right)^2 \quad (i \in \{1, 2\}), \quad \text{dist}(\hat{\phi}^1, \phi) = \text{dist}(\hat{\phi}^2, \phi + \delta). \quad (3.20)$$

A (3.15) és (3.17) egyenletek mutatják, hogy

$$\hat{A}_{i,s} = A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi}^i) \quad (i \in \{1, 2\}), \quad (3.21)$$

így a lemma (3.10) állítása az

$$\|A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi}^1) - A(\lambda_1, \lambda_2, \phi)\|^2 = \|A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi}^2) - A(\lambda_1, \lambda_2, \phi + \delta)\|^2 \quad (3.22)$$

alakot ölti. Azonos valószínűségi változók várható értéke megegyezik, így ha a (3.20) és (3.22) egyenleteket összevetjük a g_1, g_2, g_3 függvények definíciójával, azt kapjuk, hogy

$$g_i(\lambda_1, \lambda_2, \phi + \delta, \tau + \delta, \vartheta + \delta, N) = g_i(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) \quad (\forall \delta \in \mathbb{R}, i \in \{1, 2, 3\}). \quad (3.23)$$

Tehát valamennyi célfüggvény invariáns a szögparaméterek összehangolt megváltoztatására. Ez azt jelenti, hogy az egyik szögparaméter rögzíthető: elegendő a $\phi = 0$ szögparamétrú csatornákat vizsgálni, hiszen

$$g_i(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = g_i(\lambda_1, \lambda_2, 0, \tau - \phi, \vartheta - \phi, N) \quad (i \in \{1, 2, 3\}). \quad (3.24)$$

A továbbiakban előfordul, hogy a számolást áttekinthetőbbé teszi, ha a 2×2 -es mátrixokat a mátrixegységek alkotta bázis szerint kifejtjük, és a reprezentáns vektorokkal dolgozunk. Egy alapvető számolási szabályt rögzít a következő állítás, amely direkt számolással bizonyítható.

8. Lemma. *Legyen $B, J \in \mathbf{M}_n(\mathbb{R})$, ekkor az $L_B R_J : \mathbf{M}_n(\mathbb{R}) \rightarrow \mathbf{M}_n(\mathbb{R})$; $X \mapsto BXJ$ leképezés lineáris, és $L_B R_J$ mátrixa $B \otimes J^T \in \mathbf{M}_{2n}(\mathbb{R})$. Vagyis, ha az*

$$x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn} \in \mathbb{R}$$

számokat az $X = \sum_{i,j=1}^n x_{ij} E_{ij}$ egyenlet definiálja (ahol $\{E_{ij}\}_{i,j=1}^n$ a mátrixegységekből álló bázis), és hasonlóan, az

$$y_{11}, \dots, y_{1n}, y_{21}, \dots, y_{2n}, \dots, y_{n1}, \dots, y_{nn} \in \mathbb{R}$$

számok a $BXJ = \sum_{i,j=1}^n y_{ij} E_{ij}$ egyenlettel adottak, akkor

$$(y_{11}, \dots, y_{1n}, y_{21}, \dots, y_{2n}, \dots, y_{n1}, \dots, y_{nn})^T = B \otimes J^T (x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn})^T. \quad (3.25)$$

A következő állítás első felét a első torzítatlansági lemma bizonyításában is használni fogjuk, a második felének pedig későbbi bizonyításokban lesz fontos szerepe.

9. Lemma. *Ortogonalis mátrixok tenzorszorzata is ortogonalis, és ha ortogonalis mátrix elemeit négyzetre emeljük, akkor duplán sztochasztikus mátrixot kapunk.*

Bizonyítás. Ha $O_1, O_2 \in \mathbf{M}_n(\mathbb{R})$ ortogonálisak, akkor

$$(O_1 \otimes O_2)(O_1 \otimes O_2)^T = (O_1 \otimes O_2)(O_1^T \otimes O_2^T) = O_1 O_1^T \otimes O_2 O_2^T = I_n \otimes I_n = I_{2n}. \quad (3.26)$$

Teljesen hasonlóan belátható, hogy

$$(O_1 \otimes O_2)^T (O_1 \otimes O_2) = I_{2n}. \quad (3.27)$$

Tehát $O_1 \otimes O_2$ ortogonális, így többek között $(O_1 \otimes O_2)^{-1} = (O_1^T \otimes O_2^T)$.

A duplán sztochasztikus tulajdonság abból adódik, hogy ortogonális mátrix minden sora és minden oszlopa egységnyi hosszúságú vektor (az euklidészi normában). \square

10. Lemma (Torzítatlanság I). *A csatornamátrix becslése torzítatlan, azaz $\mathbf{E}(\hat{a}_{ij}) = a_{ij}$, ahol $\{a_{ij}\}_{i,j=1}^2$ az $A(\lambda_1, \lambda_2, \phi)$ csatornamátrix elemei, $\{\hat{a}_{ij}\}_{i,j=1}^2$ pedig a (2.24)-ban definiált \hat{A} becslés elemei.*

Bizonyítás. (2.16)-ban láttuk, hogy $\mathbf{E}(\hat{x}_{ij}) = x_{ij}$. A 8. lemma alapján (2.24) így írható fel a reprezentáns vektorok segítségével:

$$\hat{A} = (R(\tau) \otimes R(\vartheta)) \hat{X}, \quad (3.28)$$

vagyis

$$\begin{pmatrix} \hat{a}_{11} \\ \hat{a}_{12} \\ \hat{a}_{21} \\ \hat{a}_{22} \end{pmatrix} = \begin{pmatrix} \cos \tau \cos \vartheta & -\cos \tau \sin \vartheta & -\sin \tau \cos \vartheta & \sin \tau \sin \vartheta \\ \cos \tau \sin \vartheta & \cos \tau \cos \vartheta & -\sin \tau \sin \vartheta & -\sin \tau \cos \vartheta \\ \sin \tau \cos \vartheta & -\sin \tau \sin \vartheta & \cos \tau \cos \vartheta & -\cos \tau \sin \vartheta \\ \sin \tau \sin \vartheta & -\sin \tau \cos \vartheta & \cos \tau \sin \vartheta & \cos \tau \cos \vartheta \end{pmatrix} \begin{pmatrix} \hat{x}_{11} \\ \hat{x}_{12} \\ \hat{x}_{21} \\ \hat{x}_{22} \end{pmatrix}. \quad (3.29)$$

(Jelölésben nem különböztetjük meg a 2×2 -es mátrixot és a reprezentáns vektorát.) A várható érték linearitása miatt

$$\mathbf{E}(\hat{A}) = (R(\tau) \otimes R(\vartheta)) \mathbf{E}(\hat{X}) = (R(\tau) \otimes R(\vartheta)) X. \quad (3.30)$$

Másrészt az oszlopvektoros formalizmusban (2.19) az $X = (R(\tau)^T \otimes R(\vartheta)^T) A$ alakot ölti, így $(R(\tau)^T \otimes R(\vartheta)^T)^{-1} = R(\tau) \otimes R(\vartheta)$ miatt

$$A = (R(\tau) \otimes R(\vartheta)) X. \quad (3.31)$$

(3.30) és (3.31) összevetésével kapjuk a bizonyítandó állítást. \square

Ha $i \neq k$ vagy $j \neq l$, akkor \hat{x}_{ij} és \hat{x}_{kl} függetlenek, hiszen független mérések kimeneteleiből számíthatóak. Tehát (3.29) megmutatja, hogyan állnak elő az \hat{a}_{ij} becslések független valószínűségi változók lineáris kombinációjaként:

$$\hat{a}_k = \sum_{l \in H} c_{kl}(\tau, \vartheta) \hat{x}_l, \quad (3.32)$$

ahol $H = \{11, 12, 21, 22\}$, $k \in H$.

11. Lemma. *Legyen $\psi = \sum_{k \in H} d_k \hat{a}_k$ ($d_k \in \mathbb{R}$). Ekkor*

$$\mathbf{Var}(\psi) = \sum_{l \in H} \left(\sum_{k \in H} d_k c_{kl}(\tau, \vartheta) \right)^2 \frac{1 - x_l^2}{N}. \quad (3.33)$$

Bizonyítás. (3.32) alapján

$$\psi = \sum_{k \in H} d_k \hat{a}_k = \sum_{k \in H} d_k \sum_{l \in H} c_{kl}(\tau, \vartheta) \hat{x}_l = \sum_{l \in H} \left(\sum_{k \in H} d_k c_{kl}(\tau, \vartheta) \right) \hat{x}_l. \quad (3.34)$$

(2.16)-ban kiszámoltuk, hogy $\mathbf{Var}(\hat{x}_{ij}) = \frac{1}{N} (1 - x_{ij}^2)$. Innen az összeadandók függetlenségéből következik a bizonyítandó állítás. \square

3.2. Nemlineáris paraméterbecslések közelítése

A (3.1)-ben és (3.2)-ben definiált g_1, g_2 függvények kezelése azért nehéz, mert a $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi}$ paraméterbecslések nem lineárisak (lásd (2.29), (2.31)). Azonban lineáris becsléseket kapunk, ha a $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi}$ függvényeket (melyek $\mathbf{M}_2(\mathbb{R}) \rightarrow \mathbb{R}$ leképezések, argumentumuk a becsült csatornamátrix, vagyis \hat{A}) az elsőrendű Taylor-polinomjukkal közelítjük. A linearizálás bázispontja legyen az *argumentum várható értéke*. A 11. lemma alapján $\mathbf{Var}(\hat{a}_{ij}) = \underline{O}\left(\frac{1}{N}\right)$, így ha N elég nagy, az \hat{A} valószínűségi változó eloszlása az $\mathbf{E}(\hat{A})$ egy kis környezetére koncentrálódik, tehát ekkor az $\mathbf{E}(\hat{A})$ bázispontú Taylor-polinom jó közelítés.

A $T : \mathbf{M}_2(\mathbb{R}) \rightarrow \mathbb{R}^3$; $\hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi})$ paraméterbecslés az

$$A : \mathbb{R}^3 \supset \mathcal{D} \rightarrow \mathbf{M}_2(\mathbb{R}); (\lambda_1, \lambda_2, \phi) \mapsto A(\lambda_1, \lambda_2, \phi) \quad (3.35)$$

csatornaparaméterezés balinverze (2.27), ezért a deriváltjára

$$dT(A(\lambda_1, \lambda_2, \phi)) = (dA(\lambda_1, \lambda_2, \phi))^{-1} \quad (3.36)$$

teljesül. A deriváltját a

$$dA = \begin{pmatrix} \frac{\partial a_{11}}{\partial \lambda_1} & \frac{\partial a_{11}}{\partial \lambda_2} & \frac{\partial a_{11}}{\partial \phi} \\ \frac{\partial a_{22}}{\partial \lambda_1} & \frac{\partial a_{22}}{\partial \lambda_2} & \frac{\partial a_{22}}{\partial \phi} \\ \frac{\partial a_{12}}{\partial \lambda_1} & \frac{\partial a_{12}}{\partial \lambda_2} & \frac{\partial a_{12}}{\partial \phi} \end{pmatrix} \quad (3.37)$$

alakban fogjuk felírni. (2.25)-ből rövid számolás után adódik, hogy

$$dA(\lambda_1, \lambda_2, \phi) = \begin{pmatrix} \cos^2 \phi & \sin^2 \phi & (\lambda_2 - \lambda_1) \sin 2\phi \\ \sin^2 \phi & \cos^2 \phi & (\lambda_1 - \lambda_2) \sin 2\phi \\ \frac{1}{2} \sin 2\phi & -\frac{1}{2} \sin 2\phi & (\lambda_1 - \lambda_2) \cos 2\phi \end{pmatrix}. \quad (3.38)$$

A 7. lemma eredménye (forgatásinvariancia) miatt kitüntetett szerepe van a $\phi = 0$ szögparaméternek, így érdemes megjegyezni, hogy (3.38) alapján

$$dA(\lambda_1, \lambda_2, 0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda_1 - \lambda_2) \end{pmatrix}. \quad (3.39)$$

T konstrukciójából adódik, hogy

$$T = T|_{\mathbf{M}_2^s(\mathbb{R})} \circ S, \quad (3.40)$$

ahol $\mathbf{M}_2^s(\mathbb{R}) \subset \mathbf{M}_2(\mathbb{R})$ a szimmetrikus mátrixok altere, S pedig az $\hat{A} \mapsto \frac{1}{2} (\hat{A} + (\hat{A})^T)$ szimmetrizálás. Vezessük be az $\hat{a}_{12,s} = \frac{1}{2}(\hat{a}_{12} + \hat{a}_{21})$ jelölést. A (3.39)-beli mátrix invertálásával kapjuk, hogy az $A(\lambda_1, \lambda_2, 0)$ helyen

$$\frac{\partial \hat{\lambda}_1}{\partial \hat{a}_{11}} = \frac{\partial \hat{\lambda}_2}{\partial \hat{a}_{22}} = 1, \quad \frac{\partial \hat{\phi}}{\partial \hat{a}_{12,s}} = \frac{1}{\lambda_1 - \lambda_2}, \quad (3.41)$$

és minden más parciális derivált eltűnik. Alkalmazzuk a láncszabályt a (3.40) egyenletre, ezzel a paraméterbecslések deriváltját kapjuk.

12. Lemma. *A bevezetett jelöléseket megtartva*

$$\frac{\partial \hat{\lambda}_1}{\partial \hat{a}_{11}} = \frac{\partial \hat{\lambda}_2}{\partial \hat{a}_{22}} = 1, \quad \frac{\partial \hat{\phi}}{\partial \hat{a}_{12}} = \frac{\partial \hat{\phi}}{\partial \hat{a}_{21}} = \frac{1}{2(\lambda_1 - \lambda_2)} \quad (3.42)$$

és dT minden más komponense zérus az $A(\lambda_1, \lambda_2, 0)$ helyen.

A 10. lemma szerint $\mathbf{E}(\hat{A}) = A$, ezért a paraméterbecslések $\tilde{\lambda}_i, \tilde{\phi}$ -vel jelölt $\mathbf{E}(\hat{A})$ bázispontú lineáris közelítései a következő becslések:

$$\tilde{\lambda}_i = \hat{\lambda}_i(A) + \frac{\partial \hat{\lambda}_i}{\partial \hat{a}_{11}}(A) \cdot (\hat{a}_{11} - a_{11}) + \cdots + \frac{\partial \hat{\lambda}_i}{\partial \hat{a}_{22}}(A) \cdot (\hat{a}_{22} - a_{22}) \quad (i \in \{1, 2\}), \quad (3.43)$$

$$\tilde{\phi} = \hat{\phi}(A) + \frac{\partial \hat{\phi}}{\partial \hat{a}_{11}}(A) \cdot (\hat{a}_{11} - a_{11}) + \cdots + \frac{\partial \hat{\phi}}{\partial \hat{a}_{22}}(A) \cdot (\hat{a}_{22} - a_{22}). \quad (3.44)$$

Tömörebb jelöléssel:

$$\tilde{\lambda}_i = \hat{\lambda}_i(A) + \left\langle \text{grad} \hat{\lambda}_i(A), \hat{A} - A \right\rangle, \quad \tilde{\phi} = \hat{\phi}(A) + \left\langle \text{grad} \hat{\phi}(A), \hat{A} - A \right\rangle. \quad (3.45)$$

A paraméterbecslések elsőrendű közelítésével definiálhatjuk a g_1, g_2 függvények egy közelítését:

$$\tilde{g}_1(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{E} \left(\tilde{\phi} - \phi \right)^2, \quad (3.46)$$

$$\tilde{g}_2(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{E} \left((\tilde{\lambda}_1 - \lambda_1)^2 + (\tilde{\lambda}_2 - \lambda_2)^2 \right). \quad (3.47)$$

13. Lemma (Torzítatlanság II). *A $\tilde{\lambda}_i$ ($i \in \{1, 2\}$), $\tilde{\phi}$ linearizált becslések torzítatlanok.*

Bizonyítás.

$$\begin{aligned} \mathbf{E}(\tilde{\lambda}_i) &= \mathbf{E}(\hat{\lambda}_i(A)) + \mathbf{E} \left(\left\langle \text{grad} \hat{\lambda}_i(A), \hat{A} - A \right\rangle \right) = \hat{\lambda}_i(A) + \left\langle \text{grad} \hat{\lambda}_i(A), \mathbf{E}(\hat{A} - A) \right\rangle = \\ &= \hat{\lambda}_i(A) + 0 = \lambda_i. \end{aligned} \quad (3.48)$$

Itt kihasználtuk, hogy a csatornamátrix becslése torzítatlan (10. lemma), valamint azt, hogy a paraméterbecslés a csatornaparaméterezés balinverze, tehát

$$\hat{\lambda}_i(A(\lambda_1, \lambda_2, \phi)) = \lambda_i \quad (i \in \{1, 2\}). \quad (3.49)$$

Mivel $\hat{\phi}(A(\lambda_1, \lambda_2, \phi)) = \phi$, így teljesen hasonlóan belátható, hogy

$$\mathbf{E}(\tilde{\phi}) = \mathbf{E}(\hat{\phi}(A)) + \mathbf{E} \left(\left\langle \text{grad} \hat{\phi}(A), \hat{A} - A \right\rangle \right) = \phi. \quad (3.50)$$

□

Következmény.

$$\tilde{g}_1(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{Var}(\tilde{\phi}), \quad (3.51)$$

$$\tilde{g}_2(\lambda_1, \lambda_2, \phi, \tau, \vartheta, N) = \mathbf{Var}(\tilde{\lambda}_1) + \mathbf{Var}(\tilde{\lambda}_2). \quad (3.52)$$

(3.45) mutatja, hogy a linearizált paraméterbecslések az A csatornamátrix elemeire vonatkozó becslések lineáris kombinációi (additív konstanstól eltekintve). Az ilyen típusú becslések szórása a 11. lemma (3.33) eredménye alapján számolható. Tehát \tilde{g}_1 és \tilde{g}_2 zárt alakban felírható.

3.3. A $\phi = 0$ paraméterű csatorna vizsgálata

A *forrotásinvarianciának* nevezett jelenség (lásd (3.24)) miatt elegendő a $\phi = 0$ paraméterű csatornák paraméterbecsléseinek hatékonyságát vizsgálni. A 12. lemma (3.42) eredménye alapján a $\phi = 0$ esetben (3.45) a következő alakban írható:

$$\tilde{\lambda}_1 = \hat{\lambda}_1(A(\lambda_1, \lambda_2, 0)) + (\hat{a}_{11} - a_{11}), \quad (3.53)$$

$$\tilde{\lambda}_2 = \hat{\lambda}_2(A(\lambda_1, \lambda_2, 0)) + (\hat{a}_{22} - a_{22}), \quad (3.54)$$

$$\tilde{\phi} = \hat{\phi}(A(\lambda_1, \lambda_2, 0)) + \frac{1}{2(\lambda_1 - \lambda_2)} (\hat{a}_{12} - a_{12} + \hat{a}_{21} - a_{21}). \quad (3.55)$$

Tehát ebben az esetben

$$\mathbf{Var}(\tilde{\lambda}_1) = \mathbf{Var}(\hat{a}_{11}), \quad \mathbf{Var}(\tilde{\lambda}_2) = \mathbf{Var}(\hat{a}_{22}), \quad (3.56)$$

$$\mathbf{Var}(\tilde{\phi}) = \frac{1}{4(\lambda_1 - \lambda_2)^2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}). \quad (3.57)$$

A szokásos jelöléseket megtartva $A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi}) = \hat{A}_s$, így

$$\begin{aligned} & \|A(\lambda_1, \lambda_2, \phi) - A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi})\|^2 = \|A(\lambda_1, \lambda_2, \phi) - \hat{A}_s\|^2 = \\ & = (\hat{a}_{11} - a_{11})^2 + (\hat{a}_{22} - a_{22})^2 + \left(\frac{\hat{a}_{12} + \hat{a}_{21}}{2} - a_{12}\right)^2 + \left(\frac{\hat{a}_{12} + \hat{a}_{21}}{2} - a_{21}\right)^2 = \\ & = (\hat{a}_{11} - a_{11})^2 + (\hat{a}_{22} - a_{22})^2 + 2\left(\frac{\hat{a}_{12} - a_{12}}{2} + \frac{\hat{a}_{21} - a_{21}}{2}\right)^2, \end{aligned} \quad (3.58)$$

hiszen $a_{12} = a_{21}$. \hat{A} torzítatlanságából ($\mathbf{E}(\hat{a}_{ij}) = a_{ij}$) és (3.58)-ből következik, hogy

$$\mathbf{E}\left(\|A(\lambda_1, \lambda_2, \phi) - A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\phi})\|^2\right) = \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \frac{1}{2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}). \quad (3.59)$$

Ez utóbbi állítás tetszőleges ϕ esetén igaz.

Összegzés. A minimalizálandó $\tilde{g}_1, \tilde{g}_2, g_3$ célfüggvények a $\phi = 0$ esetben:

$$\tilde{g}_1(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) = \frac{1}{4(\lambda_1 - \lambda_2)^2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}), \quad (3.60)$$

$$\tilde{g}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) = \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}), \quad (3.61)$$

$$g_3(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) = \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \frac{1}{2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}). \quad (3.62)$$

4. Optimalizálás az egyszerűsített esetben

Ebben a fejezetben a 3. fejezet bevezetésében leírt kísérlettervezési feladat megoldásával foglalkozunk, tehát az iménti összegzésben szereplő mennyiségeket minimalizáljuk.

4.1. A csatornamátrixot legjobban becsülő eljárás

14. Tétel.

$$g_3(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) \geq \frac{1}{N} (3 - (\lambda_1^2 + \lambda_2^2)), \quad (4.1)$$

és (4.1) egyenlőséggel teljesül, ha $\tau = \vartheta = 0$.

Bizonyítás.

$$\begin{aligned} g_3(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) &= \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \frac{1}{2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) = \\ &= \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{12}) + \mathbf{Var}(\hat{a}_{21}) - \frac{1}{2} \mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}). \end{aligned} \quad (4.2)$$

A 11. lemma alapján (3.29)-ből következik, hogy

$$\begin{pmatrix} \mathbf{Var}(\hat{a}_{11}) \\ \mathbf{Var}(\hat{a}_{12}) \\ \mathbf{Var}(\hat{a}_{21}) \\ \mathbf{Var}(\hat{a}_{22}) \end{pmatrix} = \frac{1}{N} \begin{pmatrix} \cos^2 \tau \cos^2 \vartheta & \cos^2 \tau \sin^2 \vartheta & \sin^2 \tau \cos^2 \vartheta & \sin^2 \tau \sin^2 \vartheta \\ \cos^2 \tau \sin^2 \vartheta & \cos^2 \tau \cos^2 \vartheta & \sin^2 \tau \sin^2 \vartheta & \sin^2 \tau \cos^2 \vartheta \\ \sin^2 \tau \cos^2 \vartheta & \sin^2 \tau \sin^2 \vartheta & \cos^2 \tau \cos^2 \vartheta & \cos^2 \tau \sin^2 \vartheta \\ \sin^2 \tau \sin^2 \vartheta & \sin^2 \tau \cos^2 \vartheta & \cos^2 \tau \sin^2 \vartheta & \cos^2 \tau \cos^2 \vartheta \end{pmatrix} \begin{pmatrix} 1 - x_{11}^2 \\ 1 - x_{12}^2 \\ 1 - x_{21}^2 \\ 1 - x_{22}^2 \end{pmatrix}. \quad (4.3)$$

A (4.3)-beli együtthatómátrix a 9. lemma szerint duplán sztochasztikus, így

$$\mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{12}) + \mathbf{Var}(\hat{a}_{21}) = \frac{1}{N} (4 - (x_{11}^2 + x_{22}^2 + x_{12}^2 + x_{21}^2)). \quad (4.4)$$

Másrészt (ugyancsak a 11. lemma szerint)

$$\mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) = \sum_{l \in H} (c_{12,l}(\tau, \vartheta) - c_{21,l}(\tau, \vartheta))^2 \frac{1 - x_l^2}{N}. \quad (4.5)$$

$\sum_{l \in H} (c_{12,l}(\tau, \vartheta) - c_{21,l}(\tau, \vartheta))^2 = 2$, hiszen merőleges, egység hosszú vektorok különbségének norma-négyzete. $\frac{1}{N}(1 - x_l^2) \leq \frac{1}{N} (\forall l \in H)$, így

$$\sum_{l \in H} (c_{12,l}(\tau, \vartheta) - c_{21,l}(\tau, \vartheta))^2 \frac{1 - x_l^2}{N} \leq \frac{2}{N}, \quad (4.6)$$

vagyis $\frac{1}{2} \mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) \leq \frac{1}{N}$.

Végül (2.20) alapján:

$$\begin{aligned} x_{11}^2 + x_{12}^2 + x_{21}^2 + x_{22}^2 &= \text{Tr} X X^T = \\ &= \text{Tr} (R(-\tau) \Lambda(\lambda_1, \lambda_2) R(\vartheta) R(\vartheta)^T \Lambda(\lambda_1, \lambda_2)^T R(-\tau)^T) = \lambda_1^2 + \lambda_2^2, \end{aligned} \quad (4.7)$$

ezzel a (4.1) egyenlőtlenséget beláttuk. Könnyű ellenőrizni, hogy a $\tau = \vartheta = 0$ esetben $\mathbf{Var}(\hat{a}_{11}) = \frac{1}{N}(1 - \lambda_1^2)$, $\mathbf{Var}(\hat{a}_{22}) = \frac{1}{N}(1 - \lambda_2^2)$, $\mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) = \frac{2}{N}$, így a minimum felvétetik. \square

Megjegyzés. Többet is állíthatunk: (4.1) akkor és csak akkor teljesül egyenlőséggel, ha $\tau = \vartheta = 0 \pmod{\frac{\pi}{2}}$. (Feltéve, hogy a csatorna nem elfajuló: $\lambda_1 > \lambda_2$.)

Bizonyítás. A $\mathbf{Var}(\hat{a}_{11})$, $\mathbf{Var}(\hat{a}_{22})$, $\mathbf{Var}(\hat{a}_{12} + \hat{a}_{21})$ mennyiségek a (3.33) képlet szerint számolhatóak τ, ϑ, X és N ismeretében, az $x_{ij}(\lambda_1, \lambda_2, \tau, \vartheta)$ ($i, j \in \{1, 2\}$) elemek pedig a (2.19) egyenlettel adóttak, ezért $g_3(\lambda_1, \lambda_2, 0, \tau, \vartheta, N)$ explicit módon kifejezhető: a szükséges algebrai átalakításokat elvégezve adódik, hogy

$$\begin{aligned} & g_3(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) = \\ &= \frac{1}{16N} (48 - 13(\lambda_1^2 + \lambda_2^2) - 2\lambda_1\lambda_2 - (\lambda_1 - \lambda_2)^2 (\cos(4\tau) + \cos(4\vartheta)) - (\lambda_1 + \lambda_2)^2 \cos(4(\vartheta - \tau))). \end{aligned} \quad (4.8)$$

Fix $\lambda_1 > \lambda_2$, N mellett ez a kifejezés pontosan akkor minimális, ha $\cos(4\tau) = \cos(4\vartheta) = 1$, vagyis akkor és csak akkor, ha $\tau = \vartheta = 0 \pmod{\frac{\pi}{2}}$. Ellenőrzésképpen meggondolható, hogy (4.8) minimuma $\frac{1}{N} (3 - (\lambda_1^2 + \lambda_2^2))$. \square

Gyakorlati következmény. Tehát az optimális becsléshez az input állapot, a Neumann-mérés és a csatorna irányának meg kell egyeznie.

4.2. A kontrakcióparamétereket legjobban becslő eljárás

15. Tétel.

$$\tilde{g}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) \geq \frac{1}{N} (2 - (\lambda_1^2 + \lambda_2^2)), \quad (4.9)$$

és (4.9) egyenlőséggel teljesül, ha $\tau = \vartheta = 0$.

Bizonyítás.

$$\begin{aligned} \tilde{g}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) &= \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) = \\ &= \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{12}) + \mathbf{Var}(\hat{a}_{21}) - (\mathbf{Var}(\hat{a}_{12}) + \mathbf{Var}(\hat{a}_{21})). \end{aligned} \quad (4.10)$$

A 11. lemma alapján

$$\mathbf{Var}(\hat{a}_{12}) + \mathbf{Var}(\hat{a}_{21}) = \sum_{l \in H} (c_{12,l}(\tau, \vartheta))^2 \frac{1 - x_l^2}{N} + \sum_{l \in H} (c_{21,l}(\tau, \vartheta))^2 \frac{1 - x_l^2}{N}. \quad (4.11)$$

Tehát

$$\mathbf{Var}(\hat{a}_{12}) + \mathbf{Var}(\hat{a}_{21}) \leq \sum_{l \in H} (c_{12,l}(\tau, \vartheta))^2 \frac{1}{N} + \sum_{l \in H} (c_{21,l}(\tau, \vartheta))^2 \frac{1}{N} \leq \frac{2}{N}, \quad (4.12)$$

mert a (4.3)-beli együtthatómátrix duplán sztochasztikus. (4.4) és (4.7) eredményét (4.10)-be helyettesítve azt kapjuk, hogy

$$\mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) \geq \frac{1}{N} (4 - (\lambda_1^2 + \lambda_2^2)) - \frac{2}{N}, \quad (4.13)$$

ezzel a (4.9) egyenlőtlenséget beláttuk. Könnyű ellenőrizni, hogy a $\tau = \vartheta = 0$ esetben $\mathbf{Var}(\hat{a}_{11}) = \frac{1}{N}(1 - \lambda_1^2)$, $\mathbf{Var}(\hat{a}_{22}) = \frac{1}{N}(1 - \lambda_2^2)$, így a minimum felvétetik. \square

Megjegyzés. Tegyük fel, hogy $\lambda_1 \neq 0$ vagy $\lambda_2 \neq 0$. Ebben az esetben (4.9) pontosan akkor teljesül egyenlőséggel, ha $\tau = \vartheta = 0 \pmod{\frac{\pi}{2}}$.

Bizonyítás. A 15. tétel ezen erősítését hasonló módszerrel bizonyítjuk, mint a 14. tétel kiegészítését. A (3.33) képlet alapján a $\mathbf{Var}(\hat{a}_{11})$, $\mathbf{Var}(\hat{a}_{22})$ mennyiségeket fel lehet írni τ, ϑ, X és N függvényeként, a (2.19) egyenlet szerint pedig X -et lehet $\lambda_1, \lambda_2, \tau$ és ϑ függvényeként felírni. Tehát $\tilde{g}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N)$ explicit módon kifejezhető:

$$\begin{aligned} & \tilde{g}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) = \\ &= \frac{1}{8N} \left(16 - 5(\lambda_1^2 + \lambda_2^2) - ((\lambda_1^2 + \lambda_2^2) (\cos 4\vartheta + \cos 4\tau + \cos 4\vartheta \cos 4\tau) + 2\lambda_1\lambda_2 \sin 4\vartheta \sin 4\tau) \right). \end{aligned} \quad (4.14)$$

Bővítsük a fenti kifejezést $2\lambda_1\lambda_2 \cos 4\vartheta \cos 4\tau$ -val. Legyen

$$\begin{aligned} F_1 &= (\lambda_1^2 + \lambda_2^2) (\cos 4\vartheta + \cos 4\tau), \\ F_2 &= 2\lambda_1\lambda_2 (\cos 4\vartheta \cos 4\tau + \sin 4\vartheta \sin 4\tau) = 2\lambda_1\lambda_2 \cos(4(\tau - \vartheta)), \\ F_3 &= (\lambda_1^2 + \lambda_2^2 - 2\lambda_1\lambda_2) (\cos 4\vartheta \cos 4\tau) = (\lambda_1 - \lambda_2)^2 (\cos 4\vartheta \cos 4\tau). \end{aligned}$$

Ekkor

$$\tilde{g}_2(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) = \frac{1}{8N} \left(16 - 5(\lambda_1^2 + \lambda_2^2) - (F_1 + F_2 + F_3) \right). \quad (4.15)$$

- F_1 pontosan akkor veszi fel a maximumát, ha $\tau = \vartheta = 0 \pmod{\frac{\pi}{2}}$.
- F_2 maximális, ha $\tau = \vartheta \pmod{\frac{\pi}{2}}$.
- F_3 maximális, ha $\tau = \vartheta = 0$ vagy $\tau = \vartheta = \frac{\pi}{4} \pmod{\frac{\pi}{2}}$.

A fenti három megjegyzésből és (4.15)-ből következik, hogy \tilde{g}_2 pontosan akkor minimális, ha $\tau = \vartheta = 0 \pmod{\frac{\pi}{2}}$. Könnyen ellenőrizhető, hogy (4.14) minimuma $\frac{1}{N} (2 - (\lambda_1^2 + \lambda_2^2))$. \square

Gyakorlati következmény. *Tehát nemcsak a csatornamátrix, hanem a kontrakcióparaméterek becslése is pontosan akkor optimális, ha az input állapot, a Neumann-mérés és a csatorna iránya megegyezik.*

4.3. A szögparamétert legjobban becslő eljárás

16. Tétel. *Tegyük fel, hogy $\lambda_1 > \lambda_2$, $\lambda_2 \neq 0$ és $\lambda_1 \neq -\lambda_2$.*

1. *Ha $(\lambda_1 + \lambda_2)^2 \geq 2(\lambda_1 - \lambda_2)^2$, akkor $\tilde{g}_1(\lambda_1, \lambda_2, 0, \tau, \vartheta, N)$ minimuma pontosan akkor vétetik fel, ha $\tau = \vartheta = \frac{\pi}{4} \pmod{\frac{\pi}{2}}$. A minimum:*

$$\tilde{g}_1(\lambda_1, \lambda_2, 0, \frac{\pi}{4}, \frac{\pi}{4}, N) = \frac{1}{4(\lambda_1 - \lambda_2)^2} \frac{1}{2N} \left(4 - (\lambda_1 + \lambda_2)^2 \right). \quad (4.16)$$

2. *Ha $(\lambda_1 + \lambda_2)^2 < 2(\lambda_1 - \lambda_2)^2$, akkor $\tilde{g}_1(\lambda_1, \lambda_2, 0, \tau, \vartheta, N)$ minimuma pontosan akkor vétetik fel, ha*

$$\tau = \vartheta \in \bigcup_{k \in \mathbb{Z}} \left\{ \frac{1}{4} \arccos \left(-\frac{(\lambda_1 + \lambda_2)^2}{2(\lambda_1 - \lambda_2)^2} \right) + \frac{k\pi}{2}, \frac{\pi}{2} - \frac{1}{4} \arccos \left(-\frac{(\lambda_1 + \lambda_2)^2}{2(\lambda_1 - \lambda_2)^2} \right) + \frac{k\pi}{2} \right\}. \quad (4.17)$$

A minimum:

$$\tilde{g}_1(\lambda_1, \lambda_2, 0, \tau_{opt}, \vartheta_{opt}, N) = \frac{1}{4(\lambda_1 - \lambda_2)^2} \frac{1}{2N} \left(4 - (\lambda_1^2 + \lambda_2^2) - \frac{1}{8} \frac{(\lambda_1 + \lambda_2)^4}{(\lambda_1 - \lambda_2)^2} \right). \quad (4.18)$$

Bizonyítás.

$$\begin{aligned} \tilde{g}_1(\lambda_1, \lambda_2, 0, \tau, \vartheta, N) &= \frac{1}{4(\lambda_1 - \lambda_2)^2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) = \frac{1}{4(\lambda_1 - \lambda_2)^2} \frac{1}{8N} \times \\ &\times \left(16 - 3(\lambda_1^2 + \lambda_2^2) - 2\lambda_1\lambda_2 + (\lambda_1 + \lambda_2)^2(\cos 4\tau + \cos 4\vartheta) + (\lambda_1 - \lambda_2)^2 \cos 4(\tau + \vartheta) \right). \end{aligned} \quad (4.19)$$

A (4.19)-beli kifejezést kell minimalizálni fix λ_1, λ_2, N számok mellett, ez egy kétváltozós szélsőérték-probléma. A τ és ϑ szerinti parciális deriválásokból adódik, hogy ha (τ, ϑ) lokális szélsőérték hely, akkor

$$-(\lambda_1 + \lambda_2)^2 \sin(4\vartheta) = (\lambda_1 - \lambda_2)^2 \sin(4(\vartheta + \tau)), \quad (4.20)$$

$$-(\lambda_1 + \lambda_2)^2 \sin(4\tau) = (\lambda_1 - \lambda_2)^2 \sin(4(\vartheta + \tau)). \quad (4.21)$$

$(\lambda_1 + \lambda_2)^2 \neq 0$, így (4.20) és (4.21) következménye, hogy $\sin(4\vartheta) = \sin(4\tau)$. Egy elemi trigonometriai észrevétel szerint, ha ez esetben $4\vartheta \neq 4\tau \pmod{2\pi}$, akkor $\sin(4\vartheta + 4\tau) = 0$. Tehát a (4.21), (4.21) egyenletek szerint ekkor $\tau = 0$, $\vartheta = \frac{\pi}{4} \pmod{\frac{\pi}{2}}$, vagy $\tau = \frac{\pi}{4}$, $\vartheta = 0 \pmod{\frac{\pi}{2}}$. Az viszont (4.19)-ből leolvasható, hogy $\tau = \vartheta = \frac{\pi}{4} \pmod{\frac{\pi}{2}}$ mindig jobb lesz ezeknél a helyeknél, hiszen

$$-2(\lambda_1 + \lambda_2)^2 + (\lambda_1 - \lambda_2)^2 \leq -(\lambda_1 - \lambda_2)^2. \quad (4.22)$$

(4.22) pontosan akkor teljesül egyenlőséggel, ha $\lambda_2 = 0$, azonban feltettük, hogy $\lambda_2 \neq 0$. Tehát minimumhely csak a $\tau = \vartheta$ vonalon lehet. Ez azt jelenti, hogy az optimalizálási feladat egyváltozós szélsőértékkereséssé válik: tegyük fel mostantól, hogy $\tau = \vartheta$. Ekkor (4.19)-ből leolvasható a minimalizálandó kifejezés:

$$F_{\lambda_1, \lambda_2}(\tau) := 2(\lambda_1 + \lambda_2)^2 \cos(4\tau) + (\lambda_1 - \lambda_2)^2 \cos(8\tau) \rightarrow \min. \quad (4.23)$$

(4.23) τ -ban $\frac{\pi}{2}$ -periodikus, így a számolások során feltesszük, hogy $\tau \in [0, \frac{\pi}{2})$.

1. $F'_{\lambda_1, \lambda_2}(\tau) = 0$ akkor és csak akkor, ha $(\lambda_1 + \lambda_2)^2 \sin(4\tau) = -2(\lambda_1 - \lambda_2)^2 \sin(4\tau) \cos(4\tau)$, vagyis ha

$$\tau \in \left\{ 0, \frac{\pi}{4} \right\} \text{ vagy } \cos(4\tau) = -\frac{(\lambda_1 + \lambda_2)^2}{2(\lambda_1 - \lambda_2)^2}. \quad (4.24)$$

- 2.

$$F''_{\lambda_1, \lambda_2}(\tau) = -32 \left((\lambda_1 + \lambda_2)^2 \cos(4\tau) + 2(\lambda_1 - \lambda_2)^2 \cos(8\tau) \right). \quad (4.25)$$

Ez $\tau = 0$ -ban negatív (λ_1, λ_2 -től függetlenül), így ott lokális maximum van. $\tau = \frac{\pi}{4}$ -ben pontosan akkor pozitív, ha $(\lambda_1 + \lambda_2)^2 > 2(\lambda_1 - \lambda_2)^2$.

Tehát ha $(\lambda_1 + \lambda_2)^2 > 2(\lambda_1 - \lambda_2)^2$, akkor (4.24)-ből és a koszinusz függvény korlátosságából következik, hogy a lehetséges lokális szélsőérték helyek: $\tau = 0$ és $\tau = \frac{\pi}{4}$. $(\lambda_1 + \lambda_2)^2 > 2(\lambda_1 - \lambda_2)^2$, így $\tau = \frac{\pi}{4}$ lokális minimumhely. (Korábban láttuk, hogy $\tau = 0$ lokális maximumhely). Más lokális minimumhely nincsen, így $\frac{\pi}{4}$ -ben globális minimum van.

Ha $(\lambda_1 + \lambda_2)^2 = 2(\lambda_1 - \lambda_2)^2$, akkor (4.24) ismét azt mutatja, hogy ha τ lokális szélsőérték hely, akkor $\tau \in \{0, \frac{\pi}{4}\}$. $\tau = 0$ lokális maximumhely, így az egyetlen lehetséges lokális minimumhely a $\tau = \frac{\pi}{4}$. A Weierstrass-tétel szerint folytonos, periodikus függvénynek van minimumhelye, tehát $\tau = \frac{\pi}{4}$ (globális) minimumhely.

Ha $(\lambda_1 + \lambda_2)^2 < 2(\lambda_1 - \lambda_2)^2$, akkor $\tau = 0$ és $\tau = \frac{\pi}{4}$ is lokális maximumhelyek. Tehát (4.24) szerint a lehetséges lokális minimumhelyek:

$$\left\{ \tau : \cos(4\tau) = -\frac{(\lambda_1 + \lambda_2)^2}{2(\lambda_1 - \lambda_2)^2}, \tau \in \left[0, \frac{\pi}{2}\right) \right\} = \left\{ \frac{1}{4} \arccos\left(-\frac{(\lambda_1 + \lambda_2)^2}{2(\lambda_1 - \lambda_2)^2}\right), \frac{\pi}{2} - \frac{1}{4} \arccos\left(-\frac{(\lambda_1 + \lambda_2)^2}{2(\lambda_1 - \lambda_2)^2}\right) \right\}. \quad (4.26)$$

(4.23)-ból látható, hogy $F_{\lambda_1, \lambda_2}(\tau) \equiv F_{\lambda_1, \lambda_2}(\frac{\pi}{2} - \tau)$, így a Weierstrass-tétel miatt mindkét lehetséges minimumhelyen felvétetik a globális minimum.

A minimumra vonatkozó állítások ((4.16) és (4.18)) adódnak, ha visszahelyettesítünk (4.19)-be. \square

Példa. Ha $\lambda_1 = 1, \lambda_2 = 0$, akkor \tilde{g}_1 minimuma $\tau = \vartheta = \frac{\pi}{6}$ -ban illetve $\tau = \vartheta = \frac{\pi}{3}$ -ban van.

Megjegyzés (Átfedés). A minimumhely és a minimum a kontrakcióparaméterek függvényében folytonosan változnak. Könnyen ellenőrizhető, hogy ha $(\lambda_1 + \lambda_2)^2 = 2(\lambda_1 - \lambda_2)^2$, akkor (4.17) azt adja, hogy $\tau = \vartheta = \frac{\pi}{4} \pmod{\frac{\pi}{2}}$, és ebben az esetben a (4.16) és (4.18) képletek ugyanazt az eredményt adják.

Gyakorlati következmény. A szögparaméterek optimális becsléséhez az input állapot és a mérés irányának meg kell egyeznie, azonban ez az irány általában nem azonos a csatorna irányával, és a kontrakcióparaméterektől is függ.

5. Általános qubit Pauli csatorna tomográfiája

Az általános qubit Pauli csatorna paramétereit becslő eljárás sok szempontból az egyszerűsített csatornára vonatkozó eljárás természetes általánosítása, ezért a 2.1., 2.2. alfejezetekben leírtakra is támaszkodunk a módszer ismertetésekor.

5.1. Az input qubitok és a mérési irányok kiválasztása

A tomográfias eljárás három input qubit és három Neumann-mérés segítségével becsül. Korábbi vizsgálatok azt mutatják, hogy input állapotoknak érdemes olyan tiszta állapotokat választani, melyeket páronként merőleges Bloch-vektorok reprezentálnak [1],[2]. Jelöljük az input qubitok Bloch-vektorait $\underline{\theta}_1, \underline{\theta}_2, \underline{\theta}_3$ -mal. Ekkor $[\underline{\theta}_1, \underline{\theta}_2, \underline{\theta}_3]$ ortogonális mátrix, tehát

$$[\underline{\theta}_1, \underline{\theta}_2, \underline{\theta}_3] = R_z(\vartheta_z)R_y(\vartheta_y)R_x(\vartheta_x), \quad (5.1)$$

ahol R_z, R_y, R_x ugyanazokat a forgatásokat jelölik, mint (1.18)-ban, és $0 \leq \vartheta_z, \vartheta_y < \pi$, $0 \leq \vartheta_x < \frac{\pi}{2}$. Tehát kiválasztott input állapotokat a $\vartheta_z, \vartheta_y, \vartheta_x$ szögparaméterek írják le.

Jelöljük a Neumann-méréseket $\{\{M_i, I - M_i\}\}_{i=1}^3$ -vel, a reprezentáns vektorok legyenek $\{\underline{m}_i\}_{i=1}^3$. Ugyancsak korábbi vizsgálatból tudjuk, hogy érdemes páronként merőleges mérési irányokat választani [1]. Így az $\underline{m}_1, \underline{m}_2, \underline{m}_3$ mérési irányokat

$$[\underline{m}_1, \underline{m}_2, \underline{m}_3] = R_z(\tau_z)R_y(\tau_y)R_x(\tau_x) \quad (5.2)$$

alakban írhatjuk fel, ahol $0 \leq \tau_z, \tau_y < \pi$, $0 \leq \tau_x < \frac{\pi}{2}$.

5.2. Az output qubitok Bloch-koordinátáinak becslése

Ha a három output állapot Bloch-vektorait $\underline{\theta}_1^*, \underline{\theta}_2^*, \underline{\theta}_3^*$ -val jelöljük, akkor a csatornamátrixok reprezentációjára vonatkozó 2. lemma alapján

$$[\underline{\theta}_1^*, \underline{\theta}_2^*, \underline{\theta}_3^*] = A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) [\underline{\theta}_1, \underline{\theta}_2, \underline{\theta}_3]. \quad (5.3)$$

Először az output állapotok Bloch-koordinátáit becsüljük. Ehhez a j . output qubitnak N kópiáján elvégezzük az $\{M_i, I - M_i\}$ Neumann-mérést (minden $i, j \in \{1, 2, 3\}$ esetén). Jelöljük N_{ij}^+ -val azon események számát, amikor a j . output állapotban az i . irányban mérve (vagyis az $\{M_i, I - M_i\}$ POVM-et mérve) M_i a mérés kimenetele. Legyen $x_{ij} = \underline{m}_i \cdot \underline{\theta}_j^*$. Ekkor a 6. lemma és a kópiák függetlensége miatt

$$N_{ij}^+ = \text{Binom} \left(N, \frac{1 + x_{ij}}{2} \right). \quad (5.4)$$

Tehát a 2.2. alfejezet egyik tanulsága szerint x_{ij} -t torzítatlanul, $\frac{1}{N}$ -es nagyságrendben lecsengő varianciájú valószínűségi változóval becsülhetjük:

$$\hat{x}_{ij} := \frac{2}{N} N_{ij}^+ - 1. \quad (5.5)$$

Vezessük be az

$$M = [\underline{m}_1, \underline{m}_2, \underline{m}_3], \quad \Theta^* = [\underline{\theta}_1^*, \underline{\theta}_2^*, \underline{\theta}_3^*], \quad X = \{x_{ij}\}_{i,j=1}^3, \quad \hat{X} = \{\hat{x}_{ij}\}_{i,j=1}^3 \quad (5.6)$$

jelöléseket. Ekkor az $x_{ij} = \underline{m}_i \cdot \underline{\theta}_j^* \forall i, j \in \{1, 2, 3\}$ egyenletek így írhatók fel tömören:

$$X = M^T \Theta^*, \quad (5.7)$$

vagyis $MX = \Theta^*$, így becsülni tudjuk az output qubitok koordinátáit:

$$\hat{\Theta}^* := M \hat{X}. \quad (5.8)$$

5.3. A csatornamátrix becslése

Bevezetünk néhány újabb jelölést:

$$\underline{\lambda} = [\lambda_1, \lambda_2, \lambda_3], \quad \underline{\tau} = [\tau_z, \tau_y, \tau_x], \quad \underline{\phi} = [\phi_z, \phi_y, \phi_x], \quad \underline{\vartheta} = [\vartheta_z, \vartheta_y, \vartheta_x]. \quad (5.9)$$

Ekkor az (5.3) egyenlet a következő, rövidebb alakban írható:

$$\Theta^*(\underline{\lambda}, \underline{\phi}, \underline{\vartheta}) = A(\underline{\lambda}, \underline{\phi}) \Theta(\underline{\vartheta}). \quad (5.10)$$

Θ invertálható, így $\Theta^* \Theta^{-1} = A$, ez alapján a csatornamátrixra vonatkozó becslést készíthetünk:

$$\hat{A} := \hat{\Theta}^* \Theta^{-1}. \quad (5.11)$$

5.4. A csatornaparaméterek becslése a csatornamátrixra vonatkozó becslésből

A csatornaparaméterek becslését a paraméterezés invertálásával kapjuk az (5.11)-ben definiált \hat{A} becslésből. Tehát meg kell konstruálni az (1.18)-ban leírt

$$A : \mathcal{D} \rightarrow \mathbf{M}_3(\mathbb{R}); (\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \mapsto A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \quad (5.12)$$

leképezés inverzének egy kiterjesztését. Itt $\mathcal{D} \subset \mathbb{R}^6$ egy olyan paramétertartomány, amelyen A injektív, tehát a 2. lemmában leírt paramétertartomány egy részhalmaza. Vagyis meghatározandó az a

$$T : \mathbf{M}_3(\mathbb{R}) \rightarrow \mathbb{R}^6; \hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x) \quad (5.13)$$

leképezés, melyre

$$T \circ A = \text{Id}_{\mathcal{D}} \quad (5.14)$$

teljesül. Az \hat{A} valószínűségi változó $\text{Range}(A)$ -n kívül eső értékeket is felvehet, ezért lényeges, hogy T értelmezési tartománya $\mathbf{M}_3(\mathbb{R})$ legyen.

T konstrukciója a következő. Szimmetrizáljuk a becslést: $\hat{A}_s := \frac{1}{2}(\hat{A} + \hat{A}^T)$. \hat{A}_s szimmetrikus mátrix, így ortogonális bázisban diagonalizálható. \hat{A}_s Jordan-felbontása elméleti nehézség nélkül kiszámolható, hiszen a karakterisztikus polinom harmadfokú, így a sajátértékek kiszámolhatók például a Cardano-képlet segítségével. A hozzájuk tartozó sajátvektorok is elemi lineáris algebrai módszerekkel megadhatók. Legyenek tehát adottak a $\lambda_1 \geq \lambda_2 \geq \lambda_3$ sajátértékek és a hozzájuk tartozó $\mathbf{v}_1 = (v_1^1, v_1^2, v_1^3)$, $\mathbf{v}_2 = (v_2^1, v_2^2, v_2^3)$, $\mathbf{v}_3 = (v_3^1, v_3^2, v_3^3)$ normált sajátvektorok. Feltehető, hogy $v_1^3 > 0$ vagy $v_1^3 = 0$ & $v_1^2 > 0$ vagy $v_1^1 = 1$, hiszen ha ez a feltétel nem teljesül, akkor tekintjük $-\mathbf{v}_1$ -et.

A csatorna kontrakcióparaméterei éppen a $\lambda_1 \geq \lambda_2 \geq \lambda_3$ sajátértékek. A valódi feladat a szögparaméterek kiszámolása. Az (1.18)-ben leírt csatornaparaméterezésből látható a szögparaméterek szemléletes jelentése: ϕ_z és ϕ_y a legnagyobb sajátértékhez tartozó sajátvektor polár- és azimutszöge, a \mathbf{v}_2 sajátvektor pedig a a \mathbf{v}_1 -re merőleges altérben ϕ_x szöget zár be a szóban forgó altér és az $\mathbf{e}_1, \mathbf{e}_2$ standard bázisvektorok kifeszítette sík metszéspontjával. Ezek a fogalmak pontosan akkor egyértelműek, ha $\lambda_1 > \lambda_2 > \lambda_3$. A sajátértékek nem feltétlenül különbözőek, ezért négy esetet kell végiggondolni.

1. Az általános esetben $\lambda_1 > \lambda_2 > \lambda_3$.

Ha $v_1^3 = 1$, akkor legyen $\phi_y = \frac{\pi}{2}$, $\phi_z = 0$, vagyis ha az azimutszög $\frac{\pi}{2}$, akkor a polárszög nem meghatározható, válasszuk 0-nak.

Ha $v_1^3 \neq 1$ és $v_1^2 = 0$, akkor $\phi_y := \arccos v_1^1$, $\phi_z := 0$. Azaz ha \mathbf{v}_1 a $\text{span}(\mathbf{e}_1, \mathbf{e}_3)$ síkban van, de $\mathbf{v}_1 \neq \mathbf{e}_3$, akkor a polárszög 0, az azimutszög pedig a fenti egyszerű módon számolható.

Ha $v_1^2 \neq 0$, akkor $y := \text{sgn}(v_1^2) \sqrt{(v_1^1)^2 + (v_1^2)^2}$, $z := \frac{v_1^1}{y}$, végül $\phi_y := \arccos y$, $\phi_z := \arccos z$.

Ezzel adott ϕ_z és ϕ_y , így fel tudjuk írni a \mathbf{v}_1 -re merőleges altérnek azt az ortonormált bázisát, melynek első eleme a $\mathbf{e}_1, \mathbf{e}_2$ bázisvektorok kifeszítette síkban van:

$$\mathbf{s}_1 = (-\sin \phi_z, \cos \phi_z, 0), \quad \mathbf{s}_2 = (-\sin \phi_y \cos \phi_z, -\sin \phi_y \sin \phi_z, \cos \phi_y).$$

Ez a két vektor az $R_z(\phi_z)R_y(\phi_y)$ mátrix második és harmadik oszlopvektora. Így nyilván $\mathbf{v}_1, \mathbf{s}_1, \mathbf{s}_2$ is ortonormált bázis. \mathbf{v}_2 merőleges \mathbf{v}_1 -re, így $\mathbf{v}_2 = \langle \mathbf{v}_2, \mathbf{s}_1 \rangle \mathbf{s}_1 + \langle \mathbf{v}_2, \mathbf{s}_2 \rangle \mathbf{s}_2$. A $q_1 = \langle \mathbf{v}_2, \mathbf{s}_1 \rangle, q_2 = \langle \mathbf{v}_2, \mathbf{s}_2 \rangle$ jelölést bevezetve $x := \text{sgn}(q_2)q_1$, ha $|q_1| \neq 1$, és $x := 1$, ha $|q_1| = 1$, majd $\phi_x := \arccos x$.

2. A $\lambda_1 > \lambda_2 = \lambda_3$ esetben ϕ_y, ϕ_z ugyanúgy meghatározható, mint az általános esetben, és ϕ_x -et válasszuk 0-nak. (Ha $\lambda_2 = \lambda_3$, akkor $R_x \Lambda R_x^{-1} \equiv \Lambda$.)
3. A $\lambda_1 = \lambda_2 > \lambda_3$ esetben a \mathbf{v}_3 sajátvektor határozható meg előjel erejéig egyértelműen.

Ha $|v_3^3| = 1$, akkor a $\mathbf{v}_1 \in \text{span}(\mathbf{e}_1, \mathbf{e}_2)$, így az azimutszöge 0 ($\phi_y = 0$), a polárszöge meghatározhatatlan (hiszen $\text{span}(\mathbf{e}_1, \mathbf{e}_2)$ a λ_1 -hez tartozó sajátaltér), így $\phi_z = 0$ választandó. Ebben az esetben nyilván $\phi_x = 0$.

Ha $|v_3^3| \neq 1$, akkor egyértelműen létezik (és lineáris algebrai eszközökkel kiszámolható) egy \mathbf{v}_2 $\lambda_1 = \lambda_2$ -höz tartozó sajátvektor, melyre $v_2^3 = 0, v_2^1 \leq 0$ és $v_2^2 = 0 \Rightarrow v_2^2 = 1$. (Erre a vektorra lehet úgy gondolni, mint a " λ_2 -höz tartozó sajátvektorra".) \mathbf{v}_2 -ből számolható a polárszög; $\phi_z = \arccos v_2^2$. Az azimutszöget \mathbf{v}_3 \mathbf{e}_3 -mal bezárt szöge adja: $\phi_y = \arccos v_3^3$, és $\phi_x = 0$ választandó.

4. a $\lambda_1 = \lambda_2 = \lambda_3$ esetben a $\phi_x = \phi_y = \phi_z = 0$ paraméterek választandók.

A paraméterek iménti meghatározása után adódik, hogy a Pauli csatornák (1.18) szerinti paraméterezése mely paraméterteretrományon lesz bijektív. Ez a tartomány a

$$\begin{aligned} \mathcal{D} = \{ & (\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \in \mathbb{R}^6 \mid 1 \pm \lambda_3 \geq |\lambda_1 \pm \lambda_2|, \lambda_1 \geq \lambda_2 \geq \lambda_3, \\ & \phi_z, \phi_y, \phi_x \in [0, \pi), \phi_y = \frac{\pi}{2} \Rightarrow \phi_z = 0, \lambda_1 = \lambda_2 = \lambda_3 \Rightarrow \phi_z = \phi_y = \phi_x = 0, \\ & \lambda_1 = \lambda_2 > \lambda_3 \Rightarrow (\phi_x = 0 \text{ és } \phi_y = 0 \Rightarrow \phi_z = 0), \\ & \lambda_1 > \lambda_2 = \lambda_3 \Rightarrow \phi_x = 0 \} \end{aligned} \quad (5.15)$$

halmaz.

6. A paraméterbecslés hatékonysága

A paraméterbecslés hatékonyságát a szögparaméterek illetve a kontrakcióparaméterek átlagos négyzetes hibájával, valamint a becsült paraméterekhez tartozó, és a valódi csatornamátrix átlagos négyzetes távolságával jellemezhetjük. Legyen tehát

$$f_1(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E} \left(\text{dist}(\hat{\phi}_z, \phi_z)^2 + \text{dist}(\hat{\phi}_y, \phi_y)^2 + \text{dist}(\hat{\phi}_x, \phi_x)^2 \right) \quad (6.1)$$

$$f_2(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E} \left((\hat{\lambda}_1 - \lambda_1)^2 + (\hat{\lambda}_2 - \lambda_2)^2 + (\hat{\lambda}_3 - \lambda_3)^2 \right) \quad (6.2)$$

$$f_3(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{E} \left(\|A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x) - A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x)\|^2 \right), \quad (6.3)$$

ahol $\|\cdot\|$ a Hilbert-Schmidt norma, $\text{dist}(\cdot, \cdot)$ pedig a (3.4)-ben definiált függvény.

Hasonló helyzet adódik, mint az egyszerűsített esetben: az f_1, f_2 függvények nehezen kezelhetők, mert nemlineáris becslésekkel vannak definiálva. Ezért a nemlineáris paraméterbecsléseket az elsőrendű Taylor-polinomjukkal (a bázispont az argumentum várható értéke) helyettesítjük, és az így kapott lineáris becslések hatékonyságát mérő \tilde{f}_1, \tilde{f}_2 mennyiségeket definiálunk.

6.1. A szög- és kontrakcióparaméterek becslésének linearizálása

A 10. lemma bizonyítása alapján könnyen meggondolható, hogy a csatornamátrix \hat{A} becslése az általános esetben is torzítatlan ($\mathbf{E}(\hat{A}) = A$). Ezért a $\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x : \mathbf{M}_3(\mathbb{R}) \rightarrow \mathbb{R}$ függvényeknek a következő közelítéseit tekintjük:

$$\tilde{\lambda}_i := \hat{\lambda}_i(A) + \left\langle \text{grad} \hat{\lambda}_i(A), \hat{A} - A \right\rangle \quad (\forall i \in \{1, 2, 3\}), \quad (6.4)$$

$$\tilde{\phi}_\alpha := \hat{\phi}_\alpha(A) + \left\langle \text{grad} \hat{\phi}_\alpha(A), \hat{A} - A \right\rangle \quad (\forall \alpha \in \{z, y, x\}). \quad (6.5)$$

Ezekkel a lineáris paraméterbecslésekkel definiáljuk az f_1, f_2 függvények egy közelítését.

$$\tilde{f}_1(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) := \mathbf{E} \left((\tilde{\phi}_z - \phi_z)^2 + (\tilde{\phi}_y - \phi_y)^2 + (\tilde{\phi}_x - \phi_x)^2 \right), \quad (6.6)$$

$$\tilde{f}_2(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) := \mathbf{E} \left((\tilde{\lambda}_1 - \lambda_1)^2 + (\tilde{\lambda}_2 - \lambda_2)^2 + (\tilde{\lambda}_3 - \lambda_3)^2 \right) \quad (6.7)$$

A 13. lemma bizonyítását lényegében megismételve belátható, hogy a $\tilde{\lambda}_i, \tilde{\phi}_\alpha$ becslések torzítatlanok. Ebből következik, hogy

$$\tilde{f}_1(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{Var}(\tilde{\phi}_z) + \mathbf{Var}(\tilde{\phi}_y) + \mathbf{Var}(\tilde{\phi}_x), \quad (6.8)$$

$$\tilde{f}_2(\underline{\lambda}, \underline{\phi}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{Var}(\tilde{\lambda}_1) + \mathbf{Var}(\tilde{\lambda}_2) + \mathbf{Var}(\tilde{\lambda}_3). \quad (6.9)$$

A

$$T : \mathbf{M}_3(\mathbb{R}) \rightarrow \mathbb{R}^6; \hat{A} \mapsto (\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x) \quad (6.10)$$

paraméterbecslés baloldali inverze az

$$A : \mathcal{D} \rightarrow \mathbf{M}_3(\mathbb{R}); (\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \mapsto A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) \quad (6.11)$$

csatornaparaméterezésnek (5.14), ezért

$$dT(A(\underline{\lambda}, \underline{\phi})) = (dA(\underline{\lambda}, \underline{\phi}))^{-1} \quad (6.12)$$

teljesül. A forgatásinvariancia miatt elegendő a $\phi_z = \phi_y = \phi_x = 0$ szögparaméterű csatornákkal foglalkozni. (1.18)-ba behelyettesítve látható, hogy

$$A(\lambda_1, \lambda_2, \lambda_3, \phi_z, 0, 0) = \begin{pmatrix} \lambda_1 \cos^2 \phi_z + \lambda_2 \sin^2 \phi_z & (\lambda_1 - \lambda_2) \sin \phi_z \cos \phi_z & 0 \\ (\lambda_1 - \lambda_2) \sin \phi_z \cos \phi_z & \lambda_1 \sin^2 \phi_z + \lambda_2 \cos^2 \phi_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (6.13)$$

$$A(\lambda_1, \lambda_2, \lambda_3, 0, \phi_y, 0) = \begin{pmatrix} \lambda_1 \cos^2 \phi_y + \lambda_3 \sin^2 \phi_y & 0 & (\lambda_1 - \lambda_3) \sin \phi_y \cos \phi_y \\ 0 & 1 & 0 \\ (\lambda_1 - \lambda_3) \sin \phi_y \cos \phi_y & 0 & \lambda_1 \sin^2 \phi_y + \lambda_3 \cos^2 \phi_y \end{pmatrix}, \quad (6.14)$$

$$A(\lambda_1, \lambda_2, \lambda_3, 0, 0, \phi_x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda_2 \cos^2 \phi_x + \lambda_3 \sin^2 \phi_x & (\lambda_2 - \lambda_3) \sin \phi_x \cos \phi_x \\ 0 & (\lambda_2 - \lambda_3) \sin \phi_x \cos \phi_x & \lambda_2 \sin^2 \phi_x + \lambda_3 \cos^2 \phi_x \end{pmatrix}, \quad (6.15)$$

így

$$\frac{\partial A}{\partial \phi_z}(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \begin{pmatrix} 0 & (\lambda_1 - \lambda_2) & 0 \\ (\lambda_1 - \lambda_2) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (6.16)$$

$$\frac{\partial A}{\partial \phi_y}(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \begin{pmatrix} 0 & 0 & (\lambda_1 - \lambda_3) \\ 0 & 0 & 0 \\ (\lambda_1 - \lambda_3) & 0 & 0 \end{pmatrix}, \quad (6.17)$$

$$\frac{\partial A}{\partial \phi_x}(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & (\lambda_2 - \lambda_3) \\ 0 & (\lambda_2 - \lambda_3) & 0 \end{pmatrix}. \quad (6.18)$$

Másrészt $A(\underline{\lambda}, \underline{0}) = \text{Diag}(\lambda_1, \lambda_2, \lambda_3)$, így

$$\frac{\partial A}{\partial \lambda_1}(\underline{\lambda}, \underline{0}) = \text{Diag}(1, 0, 0), \quad \frac{\partial A}{\partial \lambda_2}(\underline{\lambda}, \underline{0}) = \text{Diag}(0, 1, 0), \quad \frac{\partial A}{\partial \lambda_3}(\underline{\lambda}, \underline{0}) = \text{Diag}(0, 0, 1). \quad (6.19)$$

Írjuk fel A deriváltját a

$$dA = \begin{pmatrix} \frac{\partial a_{11}}{\partial \lambda_1} & \frac{\partial a_{11}}{\partial \lambda_2} & \frac{\partial a_{11}}{\partial \lambda_3} & \frac{\partial a_{11}}{\partial \phi_z} & \frac{\partial a_{11}}{\partial \phi_y} & \frac{\partial a_{11}}{\partial \phi_x} \\ \frac{\partial a_{22}}{\partial \lambda_1} & \frac{\partial a_{22}}{\partial \lambda_2} & \frac{\partial a_{22}}{\partial \lambda_3} & \frac{\partial a_{22}}{\partial \phi_z} & \frac{\partial a_{22}}{\partial \phi_y} & \frac{\partial a_{22}}{\partial \phi_x} \\ \frac{\partial a_{33}}{\partial \lambda_1} & \frac{\partial a_{33}}{\partial \lambda_2} & \frac{\partial a_{33}}{\partial \lambda_3} & \frac{\partial a_{33}}{\partial \phi_z} & \frac{\partial a_{33}}{\partial \phi_y} & \frac{\partial a_{33}}{\partial \phi_x} \\ \frac{\partial a_{12}}{\partial \lambda_1} & \frac{\partial a_{12}}{\partial \lambda_2} & \frac{\partial a_{12}}{\partial \lambda_3} & \frac{\partial a_{12}}{\partial \phi_z} & \frac{\partial a_{12}}{\partial \phi_y} & \frac{\partial a_{12}}{\partial \phi_x} \\ \frac{\partial a_{13}}{\partial \lambda_1} & \frac{\partial a_{13}}{\partial \lambda_2} & \frac{\partial a_{13}}{\partial \lambda_3} & \frac{\partial a_{13}}{\partial \phi_z} & \frac{\partial a_{13}}{\partial \phi_y} & \frac{\partial a_{13}}{\partial \phi_x} \\ \frac{\partial a_{23}}{\partial \lambda_1} & \frac{\partial a_{23}}{\partial \lambda_2} & \frac{\partial a_{23}}{\partial \lambda_3} & \frac{\partial a_{23}}{\partial \phi_z} & \frac{\partial a_{23}}{\partial \phi_y} & \frac{\partial a_{23}}{\partial \phi_x} \end{pmatrix} \quad (6.20)$$

alakban. Ekkor (6.16), (6.17), (6.18) és (6.19) alapján

$$dA(\lambda_1, \lambda_2, \lambda_3, 0, 0, 0) = \text{Diag}(1, 1, 1, \lambda_1 - \lambda_2, \lambda_1 - \lambda_3, \lambda_2 - \lambda_3), \quad (6.21)$$

vagyis hasonló eredményre jutottunk, mint (3.39)-ben. (6.12)-ből és (6.21)-ből következik, hogy

$$dT(A(\underline{\lambda}, \underline{0})) = \text{Diag}(1, 1, 1, \frac{1}{\lambda_1 - \lambda_2}, \frac{1}{\lambda_1 - \lambda_3}, \frac{1}{\lambda_2 - \lambda_3}). \quad (6.22)$$

T konstrukciója alapján

$$T = T|\mathbf{M}_3^s(\mathbb{R}) \circ S, \quad (6.23)$$

ahol $\mathbf{M}_3^s(\mathbb{R}) \subset \mathbf{M}_3(\mathbb{R})$ a szimmetrikus mátrixok altere, S pedig az $\hat{A} \mapsto \frac{1}{2}(\hat{A} + (\hat{A})^T)$ szimmetrizálás. Legyen

$$\hat{a}_{12,s} = \frac{1}{2}(\hat{a}_{12} + \hat{a}_{21}), \quad \hat{a}_{13,s} = \frac{1}{2}(\hat{a}_{13} + \hat{a}_{31}), \quad \hat{a}_{23,s} = \frac{1}{2}(\hat{a}_{23} + \hat{a}_{32}). \quad (6.24)$$

Ekkor (6.22) eredménye így írható:

$$\frac{\partial \hat{\lambda}_1}{\partial \hat{a}_{11}} = \frac{\partial \hat{\lambda}_2}{\partial \hat{a}_{22}} = \frac{\partial \hat{\lambda}_3}{\partial \hat{a}_{33}} = 1, \quad \frac{\partial \hat{\phi}_z}{\partial \hat{a}_{12,s}} = \frac{1}{\lambda_1 - \lambda_2}, \quad \frac{\partial \hat{\phi}_y}{\partial \hat{a}_{13,s}} = \frac{1}{\lambda_1 - \lambda_3}, \quad \frac{\partial \hat{\phi}_x}{\partial \hat{a}_{23,s}} = \frac{1}{\lambda_2 - \lambda_3} \quad (6.25)$$

az $A(\underline{\lambda}, \underline{0})$ helyen, és itt a paraméterbecslések minden más parciális deriváltja zérus. Ha alkalmazzuk a láncszabályt a (6.23) egyenletre, akkor megkapjuk a paraméterbecslések gradiensét.

17. Lemma. $A \hat{\lambda}_i, \hat{\phi}_\alpha$ ($i \in \{1, 2, 3\}, \alpha \in \{z, y, x\}$) paraméterbecslések nem zérus parciális deriváltjai az $A(\underline{\lambda}, \underline{0})$ helyen:

$$\begin{aligned} \frac{\partial \hat{\lambda}_1}{\partial \hat{a}_{11}} = \frac{\partial \hat{\lambda}_2}{\partial \hat{a}_{22}} = \frac{\partial \hat{\lambda}_3}{\partial \hat{a}_{33}} = 1, \quad \frac{\partial \hat{\phi}_z}{\partial \hat{a}_{12}} = \frac{\partial \hat{\phi}_z}{\partial \hat{a}_{21}} = \frac{1}{2(\lambda_1 - \lambda_2)}, \\ \frac{\partial \hat{\phi}_y}{\partial \hat{a}_{13}} = \frac{\partial \hat{\phi}_y}{\partial \hat{a}_{31}} = \frac{1}{2(\lambda_1 - \lambda_3)}, \quad \frac{\partial \hat{\phi}_x}{\partial \hat{a}_{23}} = \frac{\partial \hat{\phi}_x}{\partial \hat{a}_{32}} = \frac{1}{2(\lambda_2 - \lambda_3)}. \end{aligned} \quad (6.26)$$

(6.4), (6.5) és az iménti lemma alapján kiszámolható a linearizált becslések szórása:

$$\mathbf{Var}(\tilde{\lambda}_1) = \mathbf{Var}(\hat{a}_{11}), \quad \mathbf{Var}(\tilde{\lambda}_2) = \mathbf{Var}(\hat{a}_{22}), \quad \mathbf{Var}(\tilde{\lambda}_3) = \mathbf{Var}(\hat{a}_{33}) \quad (6.27)$$

és

$$\mathbf{Var}(\tilde{\phi}_z) = \frac{1}{4(\lambda_1 - \lambda_2)^2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}), \quad \mathbf{Var}(\tilde{\phi}_y) = \frac{1}{4(\lambda_1 - \lambda_3)^2} \mathbf{Var}(\hat{a}_{13} + \hat{a}_{31}), \quad (6.28)$$

$$\mathbf{Var}(\tilde{\phi}_x) = \frac{1}{4(\lambda_2 - \lambda_3)^2} \mathbf{Var}(\hat{a}_{23} + \hat{a}_{32}). \quad (6.29)$$

A valódi és a becsült csatornamátrix távolsága teljesen hasonlóan számolható, mint az egyszerűsített esetben:

$$\begin{aligned} \|A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) - A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x)\|^2 &= \|A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) - \hat{A}_s\|^2 = \\ &= (\hat{a}_{11} - a_{11})^2 + (\hat{a}_{22} - a_{22})^2 + (\hat{a}_{33} - a_{33})^2 + \\ + 2 \left(\frac{\hat{a}_{12} - a_{12}}{2} + \frac{\hat{a}_{21} - a_{21}}{2} \right)^2 &+ 2 \left(\frac{\hat{a}_{13} - a_{13}}{2} + \frac{\hat{a}_{31} - a_{31}}{2} \right)^2 + 2 \left(\frac{\hat{a}_{23} - a_{23}}{2} + \frac{\hat{a}_{32} - a_{32}}{2} \right)^2. \end{aligned} \quad (6.30)$$

Így \hat{A} torzítatlansága miatt

$$\begin{aligned} \mathbf{E} \left(\|A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x) - A(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\phi}_z, \hat{\phi}_y, \hat{\phi}_x)\|^2 \right) &= \\ = \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{33}) &+ \frac{1}{2} (\mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) + \mathbf{Var}(\hat{a}_{13} + \hat{a}_{31}) + \mathbf{Var}(\hat{a}_{23} + \hat{a}_{32})). \end{aligned} \quad (6.31)$$

Összegzés. A (6.6), (6.7), (6.3) egyenletekkel definiált $\tilde{f}_1, \tilde{f}_2, \tilde{f}_3$ minimalizálandó mennyiségek az

$$\begin{aligned} \tilde{f}_1(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) &= \\ = \frac{1}{4(\lambda_1 - \lambda_2)^2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) &+ \frac{1}{4(\lambda_1 - \lambda_3)^2} \mathbf{Var}(\hat{a}_{13} + \hat{a}_{31}) + \frac{1}{4(\lambda_2 - \lambda_3)^2} \mathbf{Var}(\hat{a}_{23} + \hat{a}_{32}), \end{aligned} \quad (6.32)$$

$$\tilde{f}_2(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) = \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{33}), \quad (6.33)$$

$$\begin{aligned} \tilde{f}_3(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) &= \\ = \mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{33}) &+ \frac{1}{2} (\mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) + \mathbf{Var}(\hat{a}_{13} + \hat{a}_{31}) + \mathbf{Var}(\hat{a}_{23} + \hat{a}_{32})) \end{aligned} \quad (6.34)$$

alakban írhatóak.

6.2. A csatornamátrix leghatékonyabb becslése

18. Tétel.

$$f_3(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) \geq \frac{1}{N} (6 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)), \quad (6.35)$$

és (6.35) egyenlőséggel teljesül, ha $\underline{\tau} = \underline{\vartheta} = \underline{0}$.

Bizonyítás. Látható, hogy

$$f_3(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) = -\frac{1}{2} \left(\mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) + \mathbf{Var}(\hat{a}_{13} - \hat{a}_{31}) + \mathbf{Var}(\hat{a}_{23} - \hat{a}_{32}) \right) + \sum_{i,j \in \{1,2,3\}} \mathbf{Var}(\hat{a}_{ij}). \quad (6.36)$$

Az mátrixokat reprezentáló oszlopvektorokra áttérve

$$\hat{A} = (R(\underline{\tau}) \otimes R(\underline{\vartheta})) \hat{X} \quad (6.37)$$

írható, ahol $R(\underline{\zeta}) = R_z(\zeta_z)R_y(\zeta_y)R_x(\zeta_x)$, ($\zeta \in \{\tau, \vartheta\}$). Tehát az \hat{x}_{ij} becslések függetlensége miatt

$$\mathbf{Var}(\hat{A}) = (R(\underline{\tau}) \otimes R(\underline{\vartheta}))^2 \mathbf{Var}(\hat{X}), \quad (6.38)$$

ahol $(\cdot)^2$ a mátrixelemenkénti négyzetemelést jelöli. A 9. lemma szerint $(R(\underline{\tau}) \otimes R(\underline{\vartheta}))^2$ duplán sztochasztikus, $\mathbf{Var}(\hat{x}_{ij}) = \frac{1-x_{ij}^2}{N}$, így

$$\sum_{i,j \in \{1,2,3\}} \mathbf{Var}(\hat{a}_{ij}) = \frac{1}{N} \left(9 - \sum_{i,j \in \{1,2,3\}} x_{ij}^2 \right). \quad (6.39)$$

A (4.5)-beli kifejtést és az azt követő gondolatmenetet szinte szó szerint ismételve belátható, hogy

$$\mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) \leq \frac{2}{N}, \quad \mathbf{Var}(\hat{a}_{13} - \hat{a}_{31}) \leq \frac{2}{N}, \quad \mathbf{Var}(\hat{a}_{23} - \hat{a}_{32}) \leq \frac{2}{N}. \quad (6.40)$$

Végül

$$\sum_{i,j \in \{1,2,3\}} x_{ij}^2 = \text{Tr}XX^T = \text{Tr} (R(\underline{\tau})^T \Lambda(\lambda_1, \lambda_2, \lambda_3) R(\underline{\vartheta}) R(\underline{\vartheta})^T \Lambda(\lambda_1, \lambda_2, \lambda_3)^T R(\underline{\tau})) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \quad (6.41)$$

ezzel a (6.35) egyenlőséget beláttuk. Könnyű ellenőrizni, hogy a $\underline{\tau} = \underline{\vartheta} = \underline{0}$ esetben

$$\mathbf{Var}(\hat{a}_{ij}) = \frac{1}{N}(1 - \delta_{ij}\lambda_i^2), \quad \mathbf{Var}(\hat{a}_{12} - \hat{a}_{21}) = \mathbf{Var}(\hat{a}_{13} - \hat{a}_{31}) = \mathbf{Var}(\hat{a}_{23} - \hat{a}_{32}) = \frac{2}{N},$$

ahol $i, j \in \{1, 2, 3\}$, δ_{ij} pedig a Kronecker-szimbólum, tehát a minimum felvétetik. \square

6.3. A kontrakcióparaméterek leghatékonyabb becslése

19. Tétel.

$$\tilde{f}_2(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) \geq \frac{1}{N} (3 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)), \quad (6.42)$$

és (6.42) egyenlőséggel teljesül, ha $\underline{\tau} = \underline{\vartheta} = \underline{0}$.

Bizonyítás.

$$\tilde{f}_2(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) = \sum_{i \in \{1,2,3\}} \mathbf{Var}(\hat{a}_{ii}) = \sum_{i,j \in \{1,2,3\}} \mathbf{Var}(\hat{a}_{ij}) - \sum_{i \neq j} \mathbf{Var}(\hat{a}_{ij}). \quad (6.43)$$

(6.38) mutatja, hogy $\mathbf{Var}(\hat{a}_{ij}) \leq \frac{1}{N}$ ($\forall i, j$), hiszen $\frac{1-x_{ij}^2}{N} \leq \frac{1}{N}$ valamennyi indexre, és $(R(\underline{\tau}) \otimes R(\underline{\vartheta}))^2$ duplán sztochasztikus. Tehát

$$\sum_{i \neq j} \mathbf{Var}(\hat{a}_{ij}) \leq \frac{6}{N}. \quad (6.44)$$

Ha (6.39) és (6.41) eredményét (6.43)-be helyettesítjük, azt kapjuk, hogy

$$\mathbf{Var}(\hat{a}_{11}) + \mathbf{Var}(\hat{a}_{22}) + \mathbf{Var}(\hat{a}_{33}) \geq \frac{1}{N} (9 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2)) - \frac{6}{N}, \quad (6.45)$$

ezzel a (6.42) egyenlőtlenséget beláttuk. Könnyen ellenőrizhető, hogy a $\underline{\tau} = \underline{\vartheta} = \underline{0}$ esetben $\mathbf{Var}(\hat{a}_{11}) = \frac{1}{N}(1 - \lambda_1^2)$, $\mathbf{Var}(\hat{a}_{22}) = \frac{1}{N}(1 - \lambda_2^2)$, $\mathbf{Var}(\hat{a}_{33}) = \frac{1}{N}(1 - \lambda_3^2)$, így a minimum felvétetik. \square

Gyakorlati következmény. *Tehát a csatornamátrixot és a kontrakcióparamétereket is akkor tudjuk optimálisan becsülni, ha az input állapot, a Neumann-mérés és a csatorna iránya megegyezik.*

Megjegyzés. *Az 5. fejezetben ismeretlen irányú csatornára kidolgozott tomográfiás eljárás alkalmazható akkor is, ha feltesszük, hogy a csatorna a $\{\sigma_i\}_{i=1}^3$ Pauli mátrixoknak megfelelő irányokban kontraktál, vagyis ismert irányú csatornára is. Ekkor szögparamétert nem becsülünk, a kontrakcióparaméterek becslései pedig egyszerűsödnek: $\hat{\lambda}_i = \hat{a}_{ii}$ ($i \in \{1, 2, 3\}$). Beláttuk, hogy $\sum_{i=1}^3 \mathbf{Var}(\hat{a}_{ii}) \geq \frac{1}{N} (3 - (\lambda_1^2 + \lambda_2^2 + \lambda_3^2))$, és a $\underline{\tau} = \underline{\vartheta} = \underline{0}$ esetben egyenlőség áll fenn. Tehát ismert irányú csatorna tomográfiája optimális, ha az input állapotoknak megfelelő irányok és a mérési irányok megegyeznek a csatornairányokkal. Vagyis a 19. tétel bizonyítása egy ismert eredmény [1] újfajta igazolását is adja.*

6.4. A szögparaméterek leghatékonyabb becslése

A (6.37) képlet alapján a $\mathbf{Var}(\hat{a}_{ij} + \hat{a}_{ji})$ ($i \neq j$) mennyiségeket fel lehet írni $\underline{\tau}, \underline{\vartheta}, X$ és N függvényeként, X -et pedig a (5.7) egyenlet szerint fel lehet írni $\underline{\lambda}, \underline{\tau}$ és $\underline{\vartheta}$ függvényeként. Tehát az

$$\begin{aligned} & \tilde{f}_1(\underline{\lambda}, \underline{0}, \underline{\tau}, \underline{\vartheta}, N) = \\ & = \frac{1}{4(\lambda_1 - \lambda_2)^2} \mathbf{Var}(\hat{a}_{12} + \hat{a}_{21}) + \frac{1}{4(\lambda_1 - \lambda_3)^2} \mathbf{Var}(\hat{a}_{13} + \hat{a}_{31}) + \frac{1}{4(\lambda_2 - \lambda_3)^2} \mathbf{Var}(\hat{a}_{23} + \hat{a}_{32}) \end{aligned} \quad (6.46)$$

mennyiség explicit módon kifejezhető, optimalizálása rögzített $\underline{\lambda}$ és N mellett egy hatváltozós szélsőérték-keresési feladat (a változók: $\tau_z, \tau_y, \tau_x, \vartheta_z, \vartheta_y, \vartheta_x$). A probléma nehézségét mutatja, hogy a szögparamétereket leghatékonyabban becslő eljárás analitikus meghatározása már az egyszerűsített esetben is technikás feladat (lásd 4.3. alfejezet). Ezért fix kontrakcióparaméterek mellett, numerikus optimalizálással keressük a legjobb becslési stratégiát.

Rögzítsük például a $\lambda_1 = 0.8, \lambda_2 = 0.65, \lambda_3 = 0.5$ kontrakcióparamétereket és az $N = 1000$ mérésszámot. Ekkor a numerikusan megkeresett $(\underline{\tau}_{opt}, \underline{\vartheta}_{opt})$ minimumhely nem mutat más szabályszerűséget, mint hogy $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$ teljesül. (Az egyszerűsített esetben bizonyítani tudtuk, hogy $\tau_{opt} = \vartheta_{opt}$, tetszőleges kontrakcióparaméterek mellett.) Azonban minimumhoz közeli érték adódik két speciális helyen: $\tilde{f}_1 = 0.03676$ a $\tau_z = \vartheta_z = \frac{\pi}{4}, \tau_y = \vartheta_y = \frac{\pi}{4}, \tau_x = \vartheta_x = 0$ és a $\tau_z = \vartheta_z = \frac{\pi}{4}, \tau_y = \vartheta_y = 0, \tau_x = \vartheta_x = \frac{\pi}{4}$ esetben is, amíg

$$\min_{\underline{\tau}, \underline{\vartheta}} \tilde{f}_1 = 0.03634. \quad (6.47)$$

Az eltérés kicsinek mondható, hiszen például $\tilde{f}_1(\underline{\tau} = \underline{0}, \underline{\vartheta} = \underline{0}) = 0.05$.

Hasonló a helyzet, ha a $\lambda_1 = 0.9, \lambda_2 = 0.67, \lambda_3 = 0.6$ kontrakcióparamétereket rögzítjük ($N = 1000$ továbbra is): $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$ ekkor is teljesül, valamint a $\underline{\tau} = \underline{\vartheta} = (\frac{\pi}{4}, \frac{\pi}{4}, 0)$, és a $\underline{\tau} = \underline{\vartheta} = (\frac{\pi}{4}, 0, \frac{\pi}{4})$ esetben $\tilde{f}_1 = 0.01675$, amíg

$$\min_{\underline{\tau}, \underline{\vartheta}} \tilde{f}_1 = 0.01659 \quad (6.48)$$

Összehasonlításképpen: $\tilde{f}_1(\underline{\tau} = \underline{0}, \underline{\vartheta} = \underline{0}) = 0.02446$.

A Függelékben empirikus szimulációkon alapuló vizsgálatok néhány eredménye látható, amelyek az elméleti eredmények illusztrációja és ellenőrzése mellett sejtések megfogalmazására is ösztönöznek. A Függelék 4. és 5. ábráján az \tilde{f}_1 függvényt ábrázoljuk a $\lambda_1 = 0.8, \lambda_2 = 0.65, \lambda_3 = 0.5, N = 1000, \underline{\tau} = \underline{\vartheta}$ feltételek mellett, a $\vartheta_x = \tau_x = 0$, illetve a $\vartheta_x = \tau_x = \frac{\pi}{4}$ helyeken.

Tehát az optimális mérési eljárást leíró szögparamétereket nem tudjuk meghatározni, azonban konkrét Pauli csatornák numerikus vizsgálata alapján valószínűsíthető, hogy

1. $\underline{\tau}_{opt} = \underline{\vartheta}_{opt}$,
2. a $\underline{\tau} = \underline{\vartheta} = (\frac{\pi}{4}, \frac{\pi}{4}, 0)$, és a $\underline{\tau} = \underline{\vartheta} = (\frac{\pi}{4}, 0, \frac{\pi}{4})$ szögparaméterekkel leírható becslési eljárások *közeli* optimálisak

tetszőleges $\underline{\lambda}$, N esetén.

7. Általános irányú Pauli csatorna prímmhatvány-szintű rendszere

A qubit Pauli csatorna általánosításaként *Petz* és *Ohno* bevezették az *általánosított* Pauli csatorna fogalmát tetszőleges véges dimenziós kvantum rendszerre [6]. Az idézett mű nem definiálja a csatornairány fogalmát, ahogyan a kapcsolódó irodalom semelyik más dolgozata sem. E fejezet célja a *csatornairány* és a *szögparaméter* fogalmának kiterjesztése prímmhatvány-szintű rendszerre.

Az általánosított Pauli csatornák egy bizonyos osztálya bijekcióban áll a megfelelő Hilbert-tér *kölcsönösen torzítatlan bázisrendszereivel*; ezt a megfeleltetést és az unitér leképezések egy paraméterezését használva Pauli csatornákhoz általánosított szögparamétereket tudunk rendelni.

7.1. Kölcsönösen torzítatlan bázisrendszerek

4. Definíció (MUB). Legyenek $\mathcal{F}_1 = \{f_1^1, \dots, f_1^n\}, \dots, \mathcal{F}_r = \{f_r^1, \dots, f_r^n\}$ a \mathbb{C}^n tér ortonormált bázisai. Ha $1 \leq k \neq l \leq r$ esetén $\forall i, j \in \{1, \dots, n\}$ -re

$$|\langle f_k^i, f_l^j \rangle|^2 = \frac{1}{n} \quad (7.1)$$

teljesül, akkor az $\mathcal{F}_1, \dots, \mathcal{F}_r$ bázisrendszert kölcsönösen torzítatlan bázisrendszernek (röviden MUB-nak) hívjuk.

Ha $n = p^M$, ahol p prím, $M \in \mathbb{Z}^+$, akkor $n + 1$ elemű MUB konstruálható [7]. Azonosítsuk a \mathbb{C}^n standard bázisának elemeit az $(F(n)$ -nel jelölt) n elemű véges test elemeivel. Vezessük be a $\chi(\theta) = \exp\left(\frac{2\pi i}{p}(\theta + \theta^p + \theta^{p^2} + \dots + \theta^{p^{M-1}})\right)$ függvényt. Tekintsük az

$$X_q := \sum_{s \in F(n)} |s + q\rangle \langle s|, \quad Z_r := \sum_{s \in F(n)} \chi(rs) |s\rangle \langle s| \quad (7.2)$$

mátrixokat. Ekkor a $\{Z_r : r \in F(n)\}$ halmaz és az $\{\{X_q Z_{qr} : q \in F(n)\} : r \in F(n)\}$ halmaz elemei együtt egy $n + 1$ halmazból álló rendszert alkotnak. A rendszer valamennyi eleme n mátrixot tartalmaz, az azonos halmazba tartozó mátrixok kommutálnak. Ha a halmazrendszer minden eleméhez hozzárendeljük az elemeinek közös ortonormált sajátbázisát, egy $n + 1$ elemű kölcsönösen torzítatlan bázisrendszert kapunk. Az így kapott $\mathcal{F}_1, \dots, \mathcal{F}_{n+1}$ MUB-ot a \mathbb{C}^n tér *kitüntetett* MUB-jának hívjuk.

Példa. Legyen $p = 2, M = 1$. Ekkor $F(n) = \mathbb{Z}_2 = \{0, 1\}$, és (7.2) alapján

$$X_0 = |0\rangle \langle 0| + |1\rangle \langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X_1 = |1\rangle \langle 0| + |0\rangle \langle 1| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (7.3)$$

$$Z_0 = 1 \cdot |0\rangle \langle 0| + 1 \cdot |1\rangle \langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$Z_1 = 1 \cdot |0\rangle \langle 0| + \exp\left(\frac{2\pi i}{2} \cdot 1\right) |1\rangle \langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (7.4)$$

Tehát

$$\begin{aligned} \{X_0 Z_0, X_1 Z_0\} &= \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}, \quad \{X_0 Z_0, X_1 Z_1\} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}, \\ \{Z_0, Z_1\} &= \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\}. \end{aligned} \quad (7.5)$$

A fenti három halmazhoz tartozó közös ortonormált sajátbázisok rendre

$$\mathcal{F}_1 = \left\{ \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \right\},$$

$$\mathcal{F}_2 = \left\{ \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle) \right\}, \quad \mathcal{F}_3 = \{|0\rangle, |1\rangle\}. \quad (7.6)$$

Tehát ez \mathbb{C}^2 kitüntetett MUB-ja.

MUB-ok unitér transzformációja. Világos, hogy ha $\mathcal{F}_1, \dots, \mathcal{F}_{n+1}$ MUB, X unitér mátrix, akkor az $X\mathcal{F}_i := \{Xf_i^1, \dots, Xf_i^n\}$ definícióval adott $X\mathcal{F}_1, \dots, X\mathcal{F}_n$ bázisrendszer is MUB. Ha $X \in SU(n)$, akkor léteznek $\alpha_1, \dots, \alpha_{n-1}, \theta_1, \dots, \theta_{\frac{n(n-1)}{2}}, \beta_1, \dots, \beta_{\frac{n(n-1)}{2}}$ valós paraméterek, hogy

$$X = D(\alpha_1, \dots, \alpha_{n-1}) \prod_{1 \leq j < k \leq n} U_{j,k}(\theta_{jk}, \beta_{jk}), \quad (7.7)$$

ahol

$$D(\alpha_1, \dots, \alpha_{n-1}) = \text{Diag}\left(e^{i\alpha_1}, e^{i\alpha_2}, \dots, e^{i\alpha_{n-1}}, e^{-i\sum_{l=1}^{n-1} \alpha_l}\right), \quad (7.8)$$

$U_{j,k}$ pedig a $|j-1\rangle$ és $|k-1\rangle$ standard bázisvektorok kifeszítette altér forgatása: a hatását az említett altéren az

$$\tilde{U}(\theta, \beta) = \begin{pmatrix} \cos \theta & -\sin \theta e^{-i\beta} \\ \sin \theta e^{i\beta} & \cos \theta \end{pmatrix} \quad (7.9)$$

mátrix írja le, a komplementer altéren pedig megegyezik az identitással (lásd [8]).

7.2. Kvázi-merőleges részalgebrák kapcsolata kölcsönösen torzítatlan bázisrendszerekkel

5. Definíció. Az $\mathcal{A}_1, \mathcal{A}_2 \subset \mathbf{M}_n(\mathbb{C})$ (egységelemes, adjungálásra zárt) részalgebrákat kvázi-merőlegesnek (vagy komplementárisnak) mondjuk, ha a zérus nyomú altereik ($\mathcal{A}_1 \ominus \mathbb{C}I$ és $\mathcal{A}_2 \ominus \mathbb{C}I$) merőlegesek a Hilbert-Schmidt skaláris szorzásra nézve.

20. Lemma. Legyen $\mathcal{F}_1, \dots, \mathcal{F}_{n+1} \subset \mathbb{C}^n$ MUB. Legyen \mathcal{A}_i az \mathcal{F}_i bázisban diagonális mátrixok részalgebrája ($i \in \{1, \dots, n+1\}$). Ekkor $\{\mathcal{A}_i\}_{i=1}^{n+1}$ páronként komplementáris, maximális kommutatív részalgebrák, melyek lineárisan kifeszítik $\mathbf{M}_n(\mathbb{C})$ -t.

Bizonyítás. A definíció alapján világos, hogy \mathcal{A}_i maximális kommutatív részalgebra (röviden: M-részalgebra). Bizonyítandó a kvázi-merőlegesség. Legyen $1 \leq k \neq l \leq n+1$, $X \in \mathcal{A}_k \ominus \mathbb{C}I$, $Y \in \mathcal{A}_l \ominus \mathbb{C}I$, vagyis

$$X = (f_k^1, \dots, f_k^n) \text{Diag}(x_1, \dots, x_n) \begin{pmatrix} (f_k^1)^* \\ \vdots \\ (f_k^n)^* \end{pmatrix}, \quad Y = (f_l^1, \dots, f_l^n) \text{Diag}(y_1, \dots, y_n) \begin{pmatrix} (f_l^1)^* \\ \vdots \\ (f_l^n)^* \end{pmatrix}, \quad (7.10)$$

ahol $x_i, y_i \in \mathbb{C}$, $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i = 0$. Ekkor

$$\begin{aligned} \text{Tr}XY^* &= \text{Tr} \left((f_k^1, \dots, f_k^n) \text{Diag}(x_1, \dots, x_n) \begin{pmatrix} (f_k^1)^* \\ \vdots \\ (f_k^n)^* \end{pmatrix} (f_l^1, \dots, f_l^n) \text{Diag}(\bar{y}_1, \dots, \bar{y}_n) \begin{pmatrix} (f_l^1)^* \\ \vdots \\ (f_l^n)^* \end{pmatrix} \right) = \\ &= \text{Tr} \left(\begin{pmatrix} x_1 (f_k^1)^* \\ \vdots \\ x_n (f_k^n)^* \end{pmatrix} (\bar{y}_1 f_l^1, \dots, \bar{y}_n f_l^n) \begin{pmatrix} (f_l^1)^* \\ \vdots \\ (f_l^n)^* \end{pmatrix} (f_k^1, \dots, f_k^n) \right) = \\ &= \sum_{d=1}^n \sum_{c=1}^n x_d \bar{y}_c \underbrace{\langle f_k^d, f_l^c \rangle \langle f_l^c, f_k^d \rangle}_{\frac{1}{n}} = \frac{1}{n} \sum_{d=1}^n \sum_{c=1}^n x_d \bar{y}_c = \frac{1}{n} \left(\sum_{d=1}^n x_d \right) \left(\sum_{c=1}^n \bar{y}_c \right) = 0. \quad (7.11) \end{aligned}$$

Végül $\dim(\mathcal{A}_i) = n$, így $\dim(\mathcal{A}_i \ominus \mathbb{C}I) = n - 1$. Az imént beláttuk, hogy $\mathbb{C}I, \mathcal{A}_1 \ominus \mathbb{C}I, \dots, \mathcal{A}_{n+1} \ominus \mathbb{C}I$ merőleges alterek, tehát

$$\begin{aligned} \dim(\text{span}(\mathcal{A}_1, \dots, \mathcal{A}_{n+1})) &= \dim(\mathbb{C}I \oplus (\mathcal{A}_1 \ominus \mathbb{C}I) \oplus \dots \oplus (\mathcal{A}_{n+1} \ominus \mathbb{C}I)) = \\ &= 1 + (n+1)(n-1) = n^2 = \dim(\mathbf{M}_n(\mathbb{C})), \text{ így } \text{span}(\mathcal{A}_1, \dots, \mathcal{A}_{n+1}) = \mathbf{M}_n(\mathbb{C}). \end{aligned} \quad (7.12)$$

□

Tehát ha $\mathcal{F}_1, \dots, \mathcal{F}_{n+1} \subset \mathbb{C}^n$ MUB, \mathcal{A}_i az \mathcal{F}_i bázisban diagonális mátrixok részalgebrája, $E_i : \mathbf{M}_n(\mathbb{C}) \rightarrow \mathbf{M}_n(\mathbb{C})$ az \mathcal{A}_i részalgebrára (mint lineáris altérre) való merőleges vetítés, akkor Petz és Ohno definíciója szerint [6] az

$$\mathcal{E} : \mathbf{M}_n(\mathbb{C}) \rightarrow \mathbf{M}_n(\mathbb{C}); A \mapsto \mathcal{E}(A) := \left(1 - \sum_{i=1}^{n+1} \lambda_i\right) \frac{\text{Tr}A}{n} I + \sum_{i=1}^{n+1} \lambda_i E_i(A) \quad (7.13)$$

leképezés általánosított Pauli csatorna, amenyiben a $\lambda_i \in \mathbb{R}$ számokat úgy választjuk meg, hogy \mathcal{E} teljesen pozitív legyen. Mivel $\mathcal{A}_1, \dots, \mathcal{A}_{n+1}$ páronként komplementáris M -részalgebrák, így az (7.13)-ban definiált \mathcal{E} leképezés akkor és csak akkor teljesen pozitív, ha

$$1 + n\lambda_i \geq \sum_{j=1}^{n+1} \lambda_j \geq -\frac{1}{n-1} \quad \forall i \in \{1, \dots, n+1\}. \quad (7.14)$$

(Lásd [6].)

6. Definíció. A (7.13) egyenlettel adott általánosított Pauli csatorna csatornairányainak a

$$\mathcal{D}_i = \{\sigma \in \mathcal{A}_i : \sigma^* = \sigma, \text{Tr}\sigma = 1\} \subset \mathbf{M}_n(\mathbb{C}) \quad (7.15)$$

affin altereket tekintjük, ahol $\mathcal{A}_1, \dots, \mathcal{A}_{n+1}$ a megfelelő páronként komplementáris részalgebrák.

Megjegyzés. A fenti definíció a qubit esetben definiált csatornairány természetes általánosítása (lásd 1. definíció), és tetszőleges általánosított Pauli csatorna iránya definiálható ily módon (nyilván ekkor a részalgebrák száma nem feltétlenül $n+1$).

7.3. Pauli csatorna általánosított szögparaméterrel, prímszintű rendszere

A 7.1., 7.2. alfejezetek eredményeire támaszkodva definiálni tudjuk M -részalgebrákkal adott Pauli csatornák általánosított szögparamétereit.

7. Definíció. Legyen $n = p^M$ (p prím, $M \in \mathbb{Z}^+$), $\mathcal{F}_1, \dots, \mathcal{F}_{n+1}$ a \mathbb{C}^n tér kitüntetett kölcsönösen torzítatlan bázisrendszere, X az $\alpha_1, \dots, \alpha_{n-1}, \theta_1, \dots, \theta_{\frac{n(n-1)}{2}}, \beta_1, \dots, \beta_{\frac{n(n-1)}{2}}$ paraméterekkel leírható unitér mátrix. Legyen \mathcal{A}_i az $X\mathcal{F}_i$ bázisban diagonális mátrixok M -részalgebrája, E_i az \mathcal{A}_i -re való merőleges vetítés. Legyenek $\{\lambda_i\}_{i=1}^{n+1}$ az (7.14) feltételt kielégítő számok. Ekkor az

$$\mathcal{E} : \mathbf{M}_n(\mathbb{C}) \rightarrow \mathbf{M}_n(\mathbb{C}); A \mapsto \mathcal{E}(A) := \left(1 - \sum_{i=1}^{n+1} \lambda_i\right) \frac{\text{Tr}A}{n} I + \sum_{i=1}^{n+1} \lambda_i E_i(A) \quad (7.16)$$

leképezést az $(\alpha_1, \dots, \alpha_{n-1}, \theta_1, \dots, \theta_{\frac{n(n-1)}{2}}, \beta_1, \dots, \beta_{\frac{n(n-1)}{2}})$ általánosított szögparaméterű és $(\lambda_1, \dots, \lambda_{n+1})$ kontrakciós paraméterű Pauli csatornának nevezzük.

Példa. Vizsgáljuk a qubit rendszert. Egy korábbi példában kiszámoltuk az $\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3$ kitüntetett MUB-ot (7.6). Könnyen látható, hogy a kitüntetett MUB-hoz tartozó $\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ részalgebrák a következők:

$$\mathcal{A}_1 = \text{span}\{I, \sigma_1\}, \quad \mathcal{A}_2 = \text{span}\{I, \sigma_2\}, \quad \mathcal{A}_3 = \text{span}\{I, \sigma_3\}. \quad (7.17)$$

Legyen \mathcal{E} az (α, θ, β) általánosított szögparaméterű Pauli csatorna. Ekkor az \mathcal{E} -hoz tartozó $\mathcal{A}'_1, \mathcal{A}'_2, \mathcal{A}'_3$ részalgebrák az $X\mathcal{F}_1, X\mathcal{F}_2, X\mathcal{F}_3$ bázisokban diagonális mátrixok részalgebrái, ahol

$$X = X(\alpha, \theta, \beta) = \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta e^{-i\beta} \\ \sin \theta e^{i\beta} & \cos \theta \end{pmatrix}. \quad (7.18)$$

Tehát (X argumentumát elhagyva)

$$\mathcal{A}'_1 = \text{span}\{I, X\sigma_1X^*\}, \quad \mathcal{A}'_2 = \text{span}\{I, X\sigma_2X^*\}, \quad \mathcal{A}'_3 = \text{span}\{I, X\sigma_3X^*\}. \quad (7.19)$$

Ebből az alakból látszik, hogy az $A = \rho^{-1} \circ \mathcal{E} \circ \rho$ egyenlettel definiált (ρ a Bloch-paraméterezés) csatornamátrix a következő alakot ölti:

$$A = \tilde{R}(\alpha, \theta, \beta) \text{Diag}(\lambda_1, \lambda_2, \lambda_3) \tilde{R}^{-1}(\alpha, \theta, \beta), \quad (7.20)$$

ahol

$$\tilde{R}(\alpha, \theta, \beta) = \left\{ \left\langle \frac{1}{\sqrt{2}}\sigma_i, X(\alpha, \theta, \beta) \frac{1}{\sqrt{2}}\sigma_j X^*(\alpha, \theta, \beta) \right\rangle \right\}_{i,j=1}^3. \quad (7.21)$$

Az általánosított szögparaméterek általában nem egyeznek meg a 2. lemmában bevezetett szögparaméterekkel, hiszen direkt számolás mutatja, hogy

$$\tilde{R}(\alpha, \theta, \beta) = \begin{pmatrix} \cos 2\alpha \cos^2 \theta - \cos 2(\alpha - \beta) \sin^2 \theta & \sin 2\alpha \cos^2 \theta + \sin 2(\alpha - \beta) \sin^2 \theta & \cos(2\alpha - \beta) \sin 2\theta \\ -\sin 2\alpha \cos^2 \theta + \sin 2(\alpha - \beta) \sin^2 \theta & \cos 2\alpha \cos^2 \theta + \cos 2(\alpha - \beta) \sin^2 \theta & -\sin(2\alpha - \beta) \sin 2\theta \\ -\cos \beta \sin 2\theta & -\sin \beta \sin 2\theta & \cos 2\theta \end{pmatrix}, \quad (7.22)$$

ugyanakkor

$$R_z(\phi_z)R_y(\phi_y)R_x(\phi_x) = \begin{pmatrix} \cos \phi_z \cos \phi_y & -\cos \phi_z \sin \phi_y \sin \phi_x - \sin \phi_z \cos \phi_x & -\cos \phi_z \sin \phi_y \cos \phi_x + \sin \phi_z \sin \phi_x \\ \sin \phi_z \cos \phi_y & -\sin \phi_z \sin \phi_y \sin \phi_x + \cos \phi_z \cos \phi_x & -\sin \phi_z \sin \phi_y \cos \phi_x - \cos \phi_z \sin \phi_x \\ \sin \phi_y & \cos \phi_y \sin \phi_x & \cos \phi_y \cos \phi_x \end{pmatrix}, \quad (7.23)$$

ahol R_z, R_y, R_x a 2. lemmában definiált forgatásmátrixok. Azonban néhány speciális esetben látványos az analógia. Például

$$\tilde{R}(\alpha, 0, 0) = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha & 0 \\ -\sin 2\alpha & \cos 2\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv R_z(-2\alpha), \quad (7.24)$$

valamint

$$\tilde{R}(0, \theta, 0) = \begin{pmatrix} \cos 2\theta & 0 & \sin 2\theta \\ 0 & 1 & 0 \\ -\sin 2\theta & 0 & \cos 2\theta \end{pmatrix} \equiv R_y(-2\theta). \quad (7.25)$$

8. Összefoglalás

Jelen dolgozatban a kvantum rendszerek állapotát megváltoztató Pauli csatornák tomográfiájával foglalkoztunk, a kvantum bit rendszerre koncentrálnak.

A qubit esetben precízen definiáltuk a *csatornairány* fogalmát, és bevezettük a Pauli csatorna hatását leíró *csatornamátrixot*. Az $\mathbf{O}(3, \mathbb{R})$ forgáscsoport egy természetes faktorának paraméterezésével a csatornairányt leíró szögparamétereket vezettünk be.

Olyan tomográfiás sémára adtunk konstrukciót, amely az ismeretlen irányú Pauli csatorna csatornamátrixának, kontrakcióparamétereinek és szögparamétereinek becslésére szolgál. Egy becslési eljárás hatékonyságát sokféle mennyiség jellemzi, ebben a dolgozatban a *valószínűségi* jellegű mennyiségekkel foglalkoztunk: a csatornamátrix, a kontrakcióparaméterek és a szögparaméterek átlagos négyzetes hibájával. A kísérlettervezési feladat az említett három célfüggvény szerinti optimális becslési eljárások meghatározása.

Tekintettük a csatornabecslési feladat egy egyszerűsített változatát: a Pauli csatorna némely paramétereiről feltettük, hogy ismert. Az egyszerűsített esetben analitikus eszközökkel meghatároztuk az optimális csatornabecslési eljárást valamennyi célfüggvény esetében, az eredményeket három tételben foglaltuk össze.

Visszatértünk a teljes csatornabecslési feladathoz; meghatároztuk csatornamátrixot illetve a kontrakcióparamétereket leghatékonyabban becselő eljárást, ugyancsak analitikus úton. *Ismeretlen* irányú Pauli csatorna esetében az analitikus megközelítés újdonság (lásd [1, 2]). A harmadik célfüggvény (a szögparaméterek bizonytalansága) esetében numerikus optimalizálás és empirikus szimulációk alapján tudtunk sejtéseket megfogalmazni. Az általános eset vizsgálatának eredményeit két tételben és egy sejtésben foglaltuk össze.

A qubit eset általánosításaként prímmhatvány-szintű rendszerekre is definiáltuk a *csatornairány* fogalmát, valamint a maximális kommutatív részalgebrákkal adott Pauli csatornák osztályára kiterjesztettük a *szögparaméterezést*. Az említett osztály viszonylag általános: a qubit esetben minden Pauli csatorna maximális kommutatív részalgebrákkal adott.

Ez utóbbi munka a legérdekesebbnek tűnő továbblépi lehetőség irányába mutat, ami nem más, mint a qubit Pauli csatornára kidolgozott csatornabecslési eljárások és hatékonyságvizsgálati módszerek általánosítása n -szintű rendszerekre.

Hivatkozások

- [1] L. Ruppert, D. Virostek and K. M. Hangos. Optimal parameter estimation of Pauli channels. *J. Phys. A: Math. Theor.*, 45:265305, 2012.
- [2] G. Balló, K. M. Hangos and D. Petz. Convex Optimization-Based Parameter Estimation and Experiment Design for Pauli Channels. *IEEE Trans. on Automatic Control*, 99, 2012.
- [3] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [4] D. Petz. *Quantum Information Theory and Quantum Statistics*. Theoretical and Mathematical Physics. Springer-Verlag, 2008.

- [5] M. Mohseni, A. T. Rezakhani, and D. A. Lidar. Quantum process tomography: Resource analysis of different strategies. *Physical Review A*, 77:032322, 2008.
- [6] D. Petz and H. Ohno. Generalizations of Pauli channels. *Acta Math. Hungar.*, 124:165-177, 2009.
- [7] I. Bengtsson. Three ways to look at mutually unbiased bases. <http://arxiv.org/abs/quant-ph/0610216>
- [8] D. D'Alessandro. *Introduction to Quantum Control and Dynamics*. Applied Mathematics and Nonlinear Science. Chapman and Hall/CRC, 2008.
- [9] Wolfram Research, Inc., Champaign, Illinois. *Mathematica Edition: Version 8.0*. 2010.
- [10] D. Virosztek, L. Ruppert and K. M. Hangos. Pauli channels with unknown channel directions. *Publikáció előkészületben*.

Függelék

Az empirikus hatékonyságvizsgálat célja, kivitelezése

Bár az egyszerűsített esetben az optimális paraméterbecslési eljárást minden célfüggvény esetében sikerült analitikusan meghatározni, az általános esetben a szögparaméterek átlagos négyzetes hibáját nem tudtuk analitikusan minimalizálni (lásd 6.4. alfejezet). Ezért nemcsak az analitikus eredmények numerikus ellenőrzéséhez, hanem sejtések megfogalmazásához is hasznos eszköz az empirikus szimuláció.

A 2.1., 2.2., 2.3. alfejezetekben illetve az 5. fejezetben leírt paraméterbecslési eljárások számítógéppel modellezhetők, hiszen

1. a Neumann-mérések kimeneteleit számláló N_{ij}^+ valószínűségi változók (lásd (2.13), (5.4)) előállíthatók a matematikai programcsomagok beépített, véletelen változókat generáló függvényeivel,
2. a paraméterbecslések meghatározása N_{ij}^+ ismeretében olyan számítási feladat, amely automatizálható.

A paraméterbecslési eljárást M -szer szimulálva az egyszerűsített esetben a

$$\hat{\lambda}_1^j, \hat{\lambda}_2^j, \hat{\phi}^j, \quad j \in \{1 \dots M\} \quad (8.1)$$

paraméterbecsléseket, az általános esetben a

$$\hat{\lambda}_1^j, \hat{\lambda}_2^j, \hat{\lambda}_3^j, \hat{\phi}_z^j, \hat{\phi}_y^j, \hat{\phi}_x^j, \quad j \in \{1 \dots M\} \quad (8.2)$$

realizációkat nyerjük. Ekkor definiálhatjuk a

$$\hat{g}_1 = \frac{1}{M} \sum_{j=1}^M \text{dist}(\hat{\phi}^j, \phi)^2, \quad (8.3)$$

$$\hat{g}_2 = \frac{1}{M} \sum_{j=1}^M (\hat{\lambda}_1^j - \lambda_1)^2 + (\hat{\lambda}_2^j - \lambda_2)^2, \quad (8.4)$$

$$\hat{g}_3 = \frac{1}{M} \sum_{j=1}^M \|A(\hat{\lambda}_1^j, \hat{\lambda}_2^j, \hat{\phi}^j) - A(\lambda_1, \lambda_2, \phi)\|^2 \quad (8.5)$$

mennyiségeket, illetve analóg módon az

$$\hat{f}_1 = \frac{1}{M} \sum_{j=1}^M \text{dist}(\hat{\phi}_z^j, \phi_z)^2 + \text{dist}(\hat{\phi}_y^j, \phi_y)^2 + \text{dist}(\hat{\phi}_x^j, \phi_x)^2, \quad (8.6)$$

$$\hat{f}_2 = \frac{1}{M} \sum_{j=1}^M (\hat{\lambda}_1^j - \lambda_1)^2 + (\hat{\lambda}_2^j - \lambda_2)^2 + (\hat{\lambda}_3^j - \lambda_3)^2, \quad (8.7)$$

$$\hat{f}_3 = \frac{1}{M} \sum_{j=1}^M \|A(\hat{\lambda}_1^j, \hat{\lambda}_2^j, \hat{\lambda}_3^j, \hat{\phi}_z^j, \hat{\phi}_y^j, \hat{\phi}_x^j) - A(\lambda_1, \lambda_2, \lambda_3, \phi_z, \phi_y, \phi_x)\|^2. \quad (8.8)$$

függvényeket. Tehát a paraméterbecslés hatékonyságát jellemző várható értékek (lásd (3.1), (3.2),(3.3), illetve (6.1), (6.2),(6.3)) helyett tekinthetjük az *empirikus* várható értéket is.

A mintaátlag tulajdonságaiból adódik, hogy

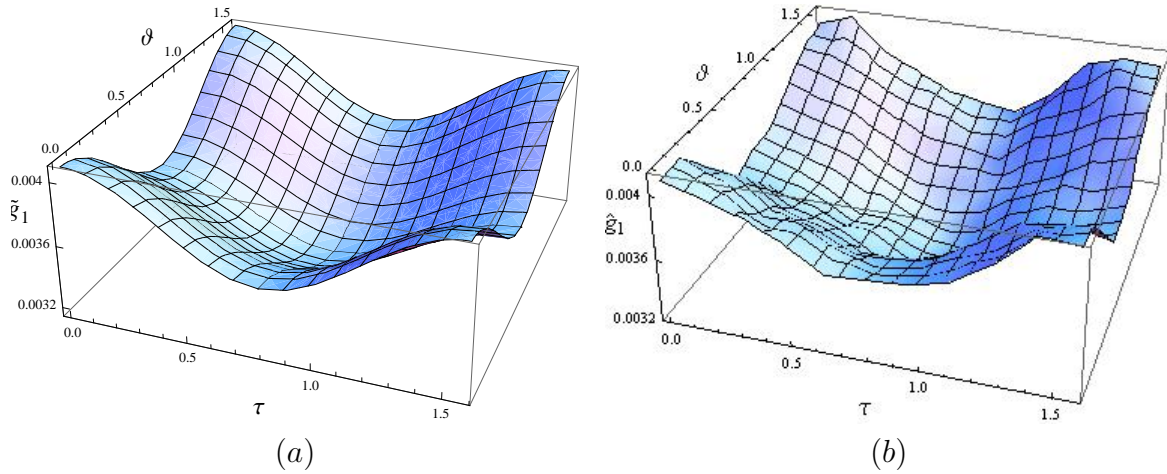
$$\mathbf{E}\hat{g}_i = g_i, \quad \mathbf{E}\hat{f}_i = f_i \quad (\forall i \in \{1, 2, 3\}), \quad \text{és} \quad \mathbf{Var}(\hat{g}_i) = \mathbf{Var}(\hat{f}_i) = \underline{\underline{O\left(\frac{1}{M}\right)}}. \quad (8.9)$$

Tehát nagy M esetén \hat{g}_i (\hat{f}_i) jól közelíti g_i -t (f_i -t).

Az empirikus hatékonyságvizsgálatot *Mathematica 8* környezetben valósítottam meg [9]. A következő oldalakon néhány eredményt mutatok be, melyek a paraméterbecslési eljárás szimulálásából adódtak.

Empirikus hatékonyságvizsgálat az egyszerűsített esetben

1. Ábrázoljuk a $\tilde{g}_1(\lambda_1 = 0.6, \lambda_2 = 0.35, \phi = 0, \tau, \vartheta, N = 1000)$ függvényt, illetve a $\hat{g}_1(0.6, 0.35, 0, \tau, \vartheta, 1000)$ függvényt $M = 25000$ mintaelemszám mellett.

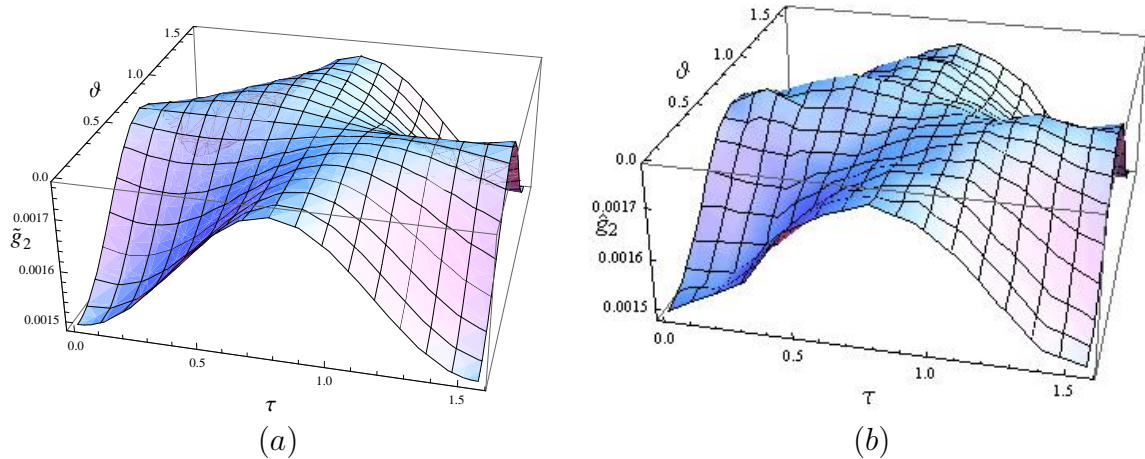


1. ábra. (a) \tilde{g}_1 , illetve (b) \hat{g}_1 , mint ϑ és τ függvénye ($0 \leq \vartheta, \tau \leq \frac{\pi}{2}$).

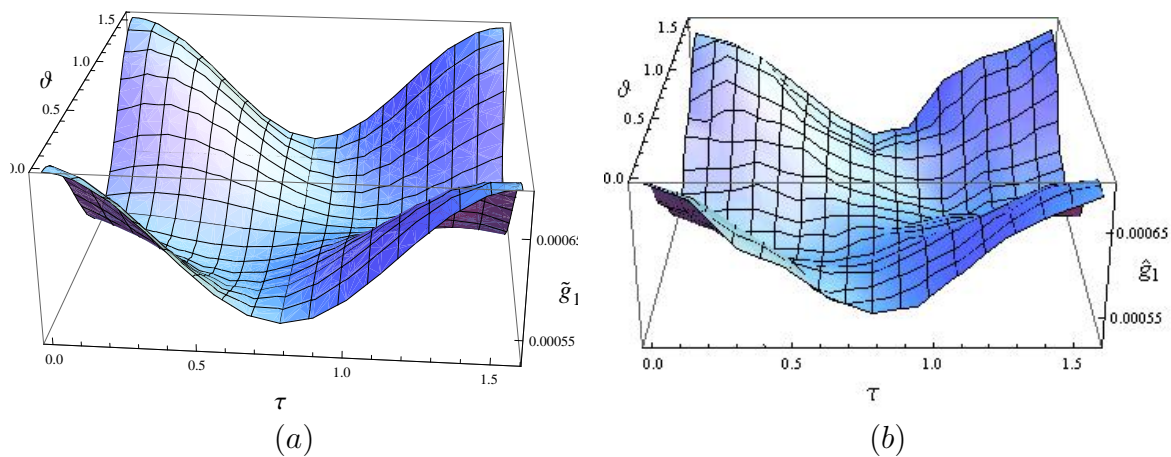
2. Ábrázoljuk a $\tilde{g}_2(\lambda_1 = 0.6, \lambda_2 = 0.35, \phi = 0, \tau, \vartheta, N = 1000)$ függvényt, valamint a $\hat{g}_2(0.6, 0.35, 0, \tau, \vartheta, 1000)$ függvényt $M = 25000$ mintaelemszám mellett (2. ábra).
3. Tekintsünk egy példát, amikor $2(\lambda_1 - \lambda_2)^2 > (\lambda_1 + \lambda_2)^2$. Ekkor a 16. tétel szerint \tilde{g}_1 minimuma nem a $\tau = \vartheta = \frac{\pi}{4}$ -ben vétetik fel. Az alábbi numerikus ellenőrzés mutatja, hogy ez a jelenség nem a $\hat{\phi}$ becslés linearizálásának hibájából származik, hanem a becslési eljárás egy érdekes tulajdonsága. Ábrázoljuk $\tilde{g}_1(0.9, 0.05, 0, \tau, \vartheta, 1000)$ -t illetve $\hat{g}_1(0.9, 0.05, 0, \tau, \vartheta, 1000)$ -t $M = 10000$ mintaelemszám mellett (3. ábra).

Empirikus hatékonyságvizsgálat az általános esetben

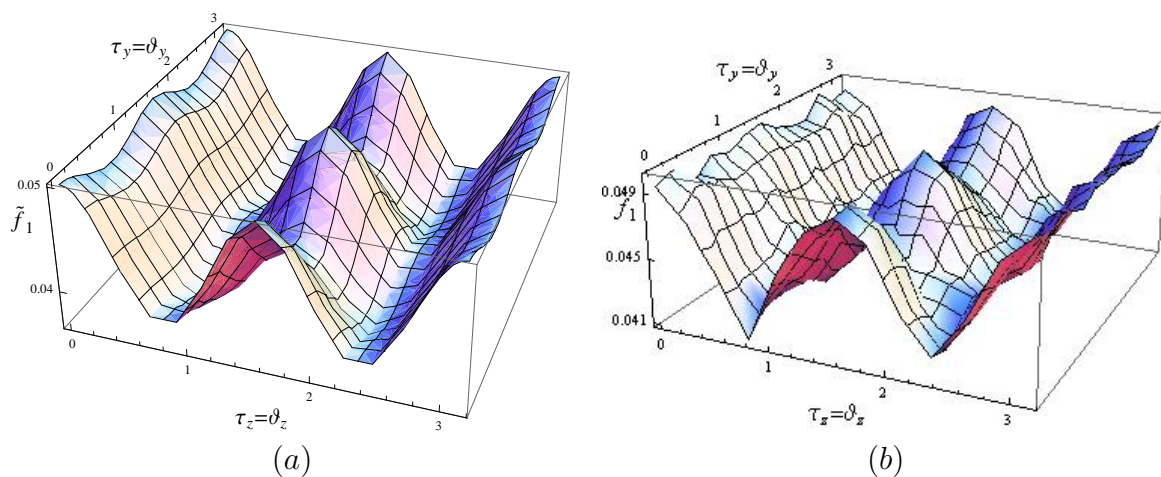
Legyen $\lambda_1 = 0.8, \lambda_2 = 0.65, \lambda_3 = 0.5, N = 1000$. Rögzítsük, hogy $\tau = \vartheta$. Ekkor a \tilde{f}_1 függvényt, és f_1 numerikus becslését ábrázolhatjuk fix $\vartheta_x = \tau_x = 0$, illetve $\vartheta_x = \tau_x = \frac{\pi}{4}$ mellett (4. és 5. ábra). Az \tilde{f}_2, \hat{f}_2 függvényeket is ábrázolhatjuk fix $\vartheta_x = \tau_x = 0$, illetve $\vartheta_x = \tau_x = \frac{\pi}{4}$ mellett (6. és 7. ábra).



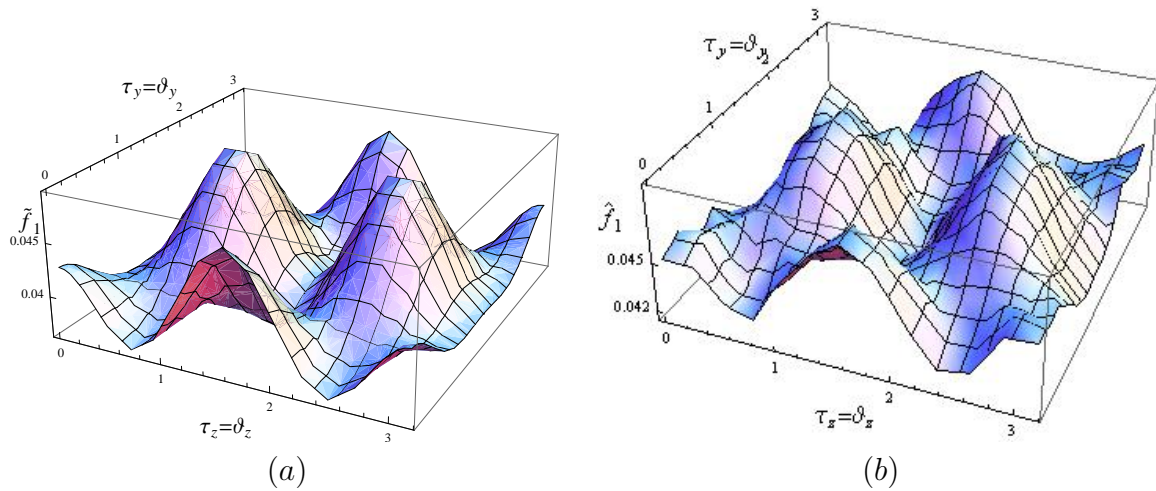
2. ábra. (a) \tilde{g}_2 , illetve (b) \hat{g}_2 , mint ϑ és τ függvénye ($0 \leq \vartheta, \tau \leq \frac{\pi}{2}$).



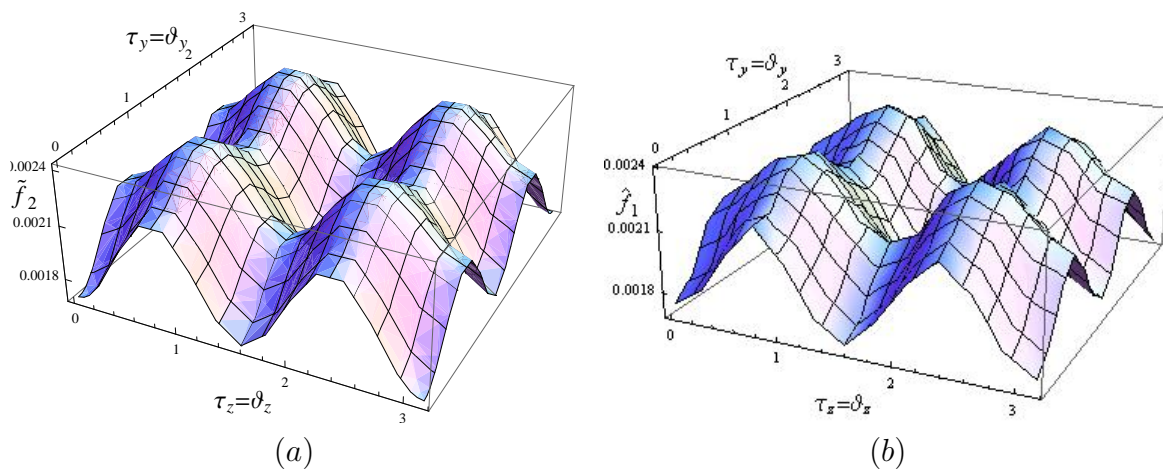
3. ábra. (a) \tilde{g}_1 , illetve (b) \hat{g}_1 , mint ϑ és τ függvénye ($0 \leq \vartheta, \tau \leq \frac{\pi}{2}$).



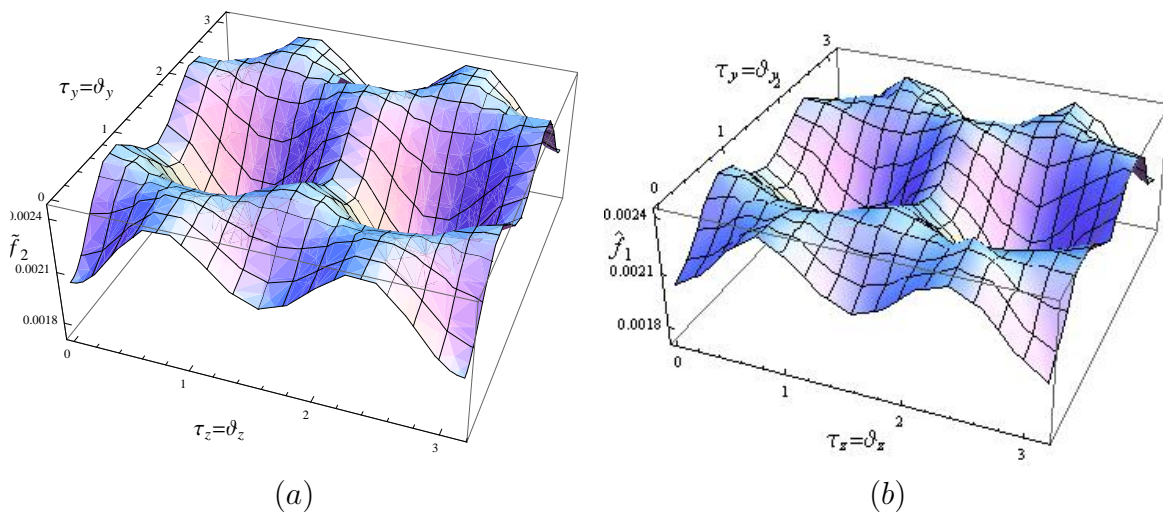
4. ábra. (a) \tilde{f}_1 és (b) \hat{f}_1 . $\vartheta_x = \tau_x = 0$, $0 \leq \vartheta_z = \tau_z, \vartheta_y = \tau_y \leq \pi$



5. ábra. (a) \tilde{f}_1 és (b) \hat{f}_1 . $\vartheta_x = \tau_x = \frac{\pi}{4}$, $0 \leq \vartheta_z = \tau_z, \vartheta_y = \tau_y \leq \pi$.



6. ábra. (a) \tilde{f}_2 és (b) \hat{f}_1 . $\vartheta_x = \tau_x = 0$, $0 \leq \vartheta_z = \tau_z, \vartheta_y = \tau_y \leq \pi$



7. ábra. (a) \tilde{f}_2 és (b) \hat{f}_2 . $\vartheta_x = \tau_x = \frac{\pi}{4}$, $0 \leq \vartheta_z = \tau_z, \vartheta_y = \tau_y \leq \pi$.