Uniwersytet im. Adama Mickiewicza w Poznaniu
Wydział Fizyki

Rozprawa doktorska

Wybrane własności korelacji w mechanice kwantowej i ogólnych teoriach probabilistycznych

Waldemar Kłobus

Promotor:
prof. UAM dr hab. Andrzej Grudka
Wydział Fizyki UAM

Promotor pomocniczy:
dr Karol Horodecki
Instytut Informatyki UG

Poznań, 16 kwietnia 2014

OŚWIADCZENIE

Ja, niżej podpisany

Waldemar Kłobus,
słuchacz studiów doktoranckich na Wydziale Fizyki
Uniwersytetu im. Adama Mickiewicza w Poznaniu

oświadczam, że przedkładaną rozprawę doktorską pt.:

Wybrane własności korelacji w mechanice kwantowej i ogólnych teoriach probabilistycznych

napisałem samodzielnie. Oznacza to, że przy pisaniu pracy, poza niezbędnymi konsultacjami, nie korzystałem z pomocy innych osób, a w szczególności nie zlecałem opracowania rozprawy lub jej części innym osobom, ani nie odpisywałem tej rozprawy lub jej części od innych osób.

Oświadczam również, że egzemplarz rozprawy doktorskiej w formie wydruku komputerowego jest zgodny z egzemplarzem rozprawy doktorskiej w formie elektronicznej.

Jednocześnie przyjmuję do wiadomości, że gdyby powyższe oświadczenie okazało się nieprawdziwe, decyzja o wydaniu mi dyplomu zostanie cofnięta.

.....
Waldemar Kłobus

Składam serdeczne podziękowania za granty, które przyczyniły się do powstania niniejszej pracy:

- European Research Council Advanced Grant: QOLAPS (jednostka kierująca: Uniwersytet Gdański, partner: Uniwersytet im. Adama Mickiewicza w Poznaniu);
- grant Narodowego Centrum Nauki: Maestro DEC-2011/02/A/ST2/00305 (jednostka kierująca: Uniwersytet Gdański);
- grant Ministerstwa Nauki i Szkolnictwa Wyższego: IdP2011 000361 (jednostka kierująca: Uniwersytet Gdański).

Jednocześnie pragnę złożyć podziękowania za przyznane mi Stypendium Fundacji UAM na rok akademicki 2013/14.

Serdecznie dziękuję Panu Prof. UAM dr. hab. Andrzejowi Grudce za wszelakie merytoryczne i pozamerytoryczne wsparcie, którego mogłem doświadczyć przez ostatnie cztery lata, a czego nie sposób szczegółowo wymienić na tejże skromnych rozmiarów stronnicy, by nie pominąć niczego, co należałoby przy tej sposobności uczynić.

Serdeczne podziękowania kieruję również pod adresem dra Karola Horodeckiego, w szczególności zaś za jego niestrudzoną chęć pomocy przy wyjaśnianiu nawet najdrobniejszych zawiłości podczas pracy naukowej.

Pragnę również podziękować wszystkim, dzięki którym możliwe było napisanie tejże pracy oraz tym, którzy nawet bezwiednie udzielali mi wsparcia, nie wyłączając rodziny, przyjaciół i osób jedynie mnie samemu wiadomych.

Siostrzenicy i bratankowi.

Streszczenie

Rozprawa doktorska jest poświęcona analizie pewnych własności korelacji w mechanice kwantowej oraz ogólnych teoriach probabilistycznych.

Pierwsza część rozprawy dotyczy aktywacji łamania nierówności CHSH za pomocą wymiany splątania. Pokazano, że wykorzystując wiele odpowiednio dobranych stanów niełamiących nierówności CHSH można, w wyniku wykonywania wielokrotnej wymiany splątania, otrzymać stan kwantowy łamiący tę nierówność.

W drugiej części rozprawy wprowadzono miary kontekstualności: *wzajemną informację kontekstualności* oraz *względną entropię kontekstualności*. Pokazano, że obie miary są sobie równoważne. Ponadto wyznaczono wartości względnej entropii kontekstualności między innymi dla następujących układów: kwadratu Peresa-Mermina, gwiazdy Mermina oraz układu Popescu-Rohrlicha.

Kolejna część rozprawy dotyczy ograniczeń na ilość korelacji między wynikami pomiarów wykonywanych przez dwie osoby na współdzielonym stanie kwantowym. Zaproponowano i udowodniono dla szerokich klas stanów kwantowych nowe zasady wykluczania informacji. Pierwsza z nich ogranicza sumę wzajemnych informacji w przypadku, gdy jedna osoba mierzy jedną obserwablę, a druga osoba mierzy jedną z dwóch obserwabli. Druga zasada ogranicza sumę wzajemnych informacji w przypadku, gdy obie osoby mierzą po jednej z dwóch obserwabli.

W ostatniej części rozprawy zbadano relacje między kodami swobodnego dostępu a układami (nie-)sygnalizującymi. Zdefiniowano *układ związany z kodem swobodnego dostępu* (*uKSD*), który wraz z wykorzystaniem dodatkowego bitu informacji funkcjonuje jak kod swobodnego dostępu. Wykazano, że *niesygnalizujący uKSD* jest równoważny układowi Popescu-Rohrlicha. Dodatkowo sformułowano nierówności wiążące możliwości wykorzystania tych zasobów oraz pokazano, że w niektórych przypadkach zasoby sygnalizujące mogą być mniej użyteczne niż niesygnalizujące.

Abstract

The matter of the dissertation covers the topic of correlations in quantum mechanics and in generalized probabilistic theories.

The first section of the dissertation concerns the activation of violation of CHSH inequality using entanglement swapping. It was shown that by performing multiple entanglement swappings on a number of states which do not violate CHSH inequality it is possible to obtain a state which violates this inequality.

In the second section two measures of contextuality are introduced: *mutual information of contextuality* and *relative entropy of contextuality*. It was shown that the two measures are equal. The relative entropy of contextuality was calculated for some particular boxes including among others Peres-Mermin square, Mermin star and Popescu-Rohrlich box.

The next section analyzes bounds on correlations between outputs of measurements performed by two parties on a shared quantum state. The new information exclusion relations are postulated and proved for a vast class of quantum states. The first relation bounds the sum of two mutual informations when one party measures a single observable and the other party measures one of two observables. The second relation bounds the sum of two mutual informations when each party measures one of two observables.

In the last section the relation between random access codes and (non-)signalling boxes is analyzed. A new box called *racbox* is defined, which if supplied with one bit of communication offers random access code. It was shown that *nonsignalling racbox* is equivalent to Popescu-Rohrlich box. Additionally, the resource inequality is formulated and it was shown that in some cases signalling resources can be less efficient than nonsignalling ones.

Spis treści

Spis rysunków

Rozdział 1
Wstęp

Niniejsza rozprawa jest poświęcona zbadaniu wybranych własności korelacji w mechanice kwantowej i ogólnych teoriach probabilistycznych w różnych ich aspektach. Struktura pracy przedstawia się następująco:

W Rozdziale 2 omówimy podstawowe pojęcia wykorzystane w dalszej części pracy. Zdefiniujemy tu pojęcie korelacji lokalnych, wyprowadzimy nierówność CHSH [1] oraz pokażemy w jaki sposób jest ona łamana przez pewne szczególne układy kwantowe. Dalej omówimy warunki niesygnalizowania oraz zdefiniujemy układ Popescu-Rohrlicha [2, 3] łamiący maksymalnie nierówność CHSH. Omówimy kody swobodnego dostępu, po czym rozpatrzymy możliwości jego symulowania wykorzystując układy kwantowe lub układ Popescu-Rohrlicha. W dalszej części przypomnimy znane w literaturze zasady nieoznaczoności ze szczególnym uwzględnieniem entropowych zasad nieoznaczoności. Następnie omówimy relację wykluczania informacji Halla [4], która wynika bezpośrednio z entropowej zasady nieoznaczoności Maassena-Uffinka [5]. W ostatniej części wprowadzimy pojęcie kontekstualności. Przytoczymy tutaj również przykłady układów kontekstualnych realizowanych w ramach mechaniki kwantowej: kwadrat Peresa-Mermina, gwiazda Mermina [6–8] oraz układ KCBS [9].

W Rozdziale 3 omówimy aktywację łamania nierówności CHSH z wykorzystaniem wielokrotnej wymiany splątania kwantowego dokonaną w łańcuchu stanów dwuqubitowych [10–13]. W pierwszej kolejności scharakteryzujemy wykorzystywane stany kwantowe oraz określimy przedziały parametrów, dla których nie łamią one nierówności CHSH. Pokażemy, że w przypadku otrzymywania pewnych określonych wyników pomiaru Bella po dostatecznie dużej liczbie wymian splątania, z początkowych stanów niełamiących nierówności CHSH można uzyskać stan końcowy łamiący tę nierówność. Następnie rozważymy przypadki, które dla określonych parametrów początkowych stanów kwantowych nie prowadzą do aktywacji łamania nierówności CHSH.

Rozdział 4 poświęcony jest miarom korelacji kontekstualnych [14]. W szczególności wprowadzimy pojęcie dwóch miar: pierwsza z nich (*wzajemna informacja kontekstualności*) określa ilość korelacji między zmienną losową określającą kontekst pomiarowy a wynikami pomiarów dokonanych na danym układzie; druga miara (*względnej entropii kontekstualności*) określa minimalną sumę odległości rodzin rozkładów prawdopodobieństwa opisujących poszczególne konteksty pomiarowe od rodzin rozkładów prawdopodobieństwa pochodzących z łącznego niekontekstualnego rozkładu prawdopodobieństwa. W dalszej kolejności wykażemy równość obu miar kontekstualności.Następnie przystąpimy do wyznaczenia wartości względnej entropii kontekstualności dla następujących układów: *kwadratu Peresa-Mermina*, *gwiazdy Mermina*

[6–8], *układu Popescu-Rohrlicha* [2, 3] oraz najprostszego układu kontekstualnego przedstawionego w pracy [9]. W tym celu skorzystamy z odpowiednich technik symetryzacji [15–17] dla pewnych szczególnych rodzin rozkładów prawdopodobieństwa określanych dalej układami izotropowymi. Umożliwi to określenie wartości miary kontekstualności dla dowolnych układów izotropowych, których szczególnymi reprezentantami są wyżej wymienione układy. Należy wspomnieć, że wcześniej użyto w analogiczny sposób względnej entropii do opisu układów nielokalnych [18]: dla rozważanych tam układów względna entropia kontekstualności redukuje się do wprowadzonej tamże miary *statystycznej siły dowodów nielokalności*. Wielkości tej w [18] nadano operacyjną interpretację związaną z rozróżnianiem układów lokalnych i nielokalnych. W Rozdziale 4 wykażemy, że możliwa jest również inna operacyjna interpretacja - oparta na symulacji układu kontekstualnego za pomocą układów niekontekstualnych.

W Rozdziale 5 wprowadzimy relacje wykluczania informacji [4] ograniczające sumę dwóch wzajemnych informacji dla odpowiednich par obserwabli mierzonych przez dwie odległe osoby [19]. Pierwsza relacja dotyczy przypadku dwóch wzajemnych informacji, gdzie na pierwszym podukładzie mierzona jest tylko jedna obserwabla, a na drugim podukładzie mierzona jest jedna z dwóch obserwabli. Pokażemy, że relacja ta w pewnych przypadkach jest znacznie silniejsza niż relacja wykluczania informacji Halla [4]. Druga relacja dotyczy przypadku dwóch wzajemnych informacji, gdzie na każdym podukładzie może być mierzona jedna z dwóch obserwabli. Obie relacje udowodnimy dla pewnych klas stanów kwantowych. W rozdziale tym przywołamy również dowód pierwszej z nich dla *dowolnych* stanów przedstawiony jako jeden z wyników w pracy autorstwa Colesa i Pianiego [20].

W Rozdziale 6 zajmiemy się porównaniem dwóch zasobów: układu Popescu-Rohrlicha z *kodami swobodnego dostępu* [21]. W tym celu zdefiniujemy dodatkowy układ ($uKSD$) po czym określimy jego własności. Pokażemy, że *niesygnalizujący uKSD* jest równoważny układowi Popescu-Rohrlicha. W dalszej części sformułujemy nierówność wiążącą te zasoby. W szczególności wykażemy, że dysponując *dowolnym uKSD* (sygnalizującym bądź niesygnalizującym) możemy, z wykorzystaniem jednego bitu informacji oraz jednego bitu współdzielonej losowości, zasymulować układ Popescu-Rohrlicha a dodatkowo otrzymać kanał wymazujący. Podamy również przykład *sygnalizującego uKSD* nasycającego tę nierówność.

W Rozdziale 7 streścimy najważniejsze wyniki przedstawione w tej pracy.

Rozdział 2

Podstawowe wiadomości teoretyczne

2.1. Nielokalność

Jedną z najistotniejszych własności korelacji w mechanice kwantowej, fundamentalnie odróżniających ją od klasycznych korelacji, jest ich nielokalny charakter. W tym podrozdziale przytoczymy podstawowe pojęcia, za pomocą których możliwe jest ścisłe zdefiniowane teorii lokalnych. Teorie te muszą spełniać pewne ograniczenia zwane nierównościami Bella [22]. Pokażemy również, że korelacje wynikające z mechaniki kwantowej, bądź ogólnych niesygnalizujących teorii probabilistycznych mogą nie spełniać tych nierówności. Świadczy to o nielokalnym charakterze korelacji kwantowych i niesygnalizujących [23].

2.1.1 Korelacje lokalne

Rozważmy następujący eksperyment („eksperyment Bella"): dwa przestrzennie odseparowane układy, które w przeszłości mogły oddziaływać ze sobą, poddane są pewnym pomiarom przez dwoje użytkowników, Alicję i Bolka (oznaczanych odpowiednio A i B). Osoby te wykonują pomiary na różnych układach. Alicja i Bolek mogą w ogólności dokonać jednego z wielu różnych pomiarów. Wybrany przez Alicję pomiar oznaczymy przez x. Analogicznie, przez y oznaczymy pomiar wybrany przez Bolka. Po dokonanym pomiarze, zarówno Alicja jak i Bolek otrzymują pewne wyniki pomiarów. Przez a oznaczymy wynik pomiaru x otrzymany przez Alicję, a przez b wynik pomiaru y otrzymany przez Bolka. Zwróćmy uwagę, że x, y, a, b mogą przyjmować dowolne wartości, ponieważ jest to całkowicie arbitralny zestaw liczb oznaczający pomiary i wyniki tych pomiarów.

Zauważmy, że wyniki a i b mogą się różnić w różnych seriach pomiarowych, nawet jeśli Alicja i Bolek wykonywali te same pomiary x i y. Innymi słowy, wyniki pomiarów zależą w ogólności od łącznego rozkładu prawdopodobieństwa dla wyników a i b jeśli wykonywane były pomiary x oraz y: $p(a, b|x, y)$. Dany rozkład prawdopodobieństwa $p(a, b|x, y)$ można z kolei szacować powtarzając odpowiednio dużą liczbę razy pomiary x oraz y, a następnie zliczać otrzymane wyniki a oraz b. Zauważmy, że w ogólności wyniki pomiarów nie są całkowicie niezależne od siebie:

$$p(a, b|x, y) \neq p(a|x)p(b|y), \tag{2.1}$$

tzn. wyniki a oraz b mogą być w pewien sposób *skorelowane* ze sobą.

W celu zdefiniowania pojęcia *lokalności* załóżmy, że istnieje pewien zbiór czynników mających wpływ na otrzymywane wyniki pomiarów. Oznaczmy ten zbiór przez λ. Zaznaczmy, że w ogólności w tym zbiorze mogą istnieć pewne zmienne, których wartości nie jesteśmy w stanie poznać na drodze jakichkolwiek eksperymentów (tzw. „zmienne ukryte"). Zbiór λ oraz wybrane pomiary x i y determinują możliwe wyniki pomiarów a i b zadając pewien łączny rozkład prawdopodobieństwa $p(a,b|x,y,\lambda)$. Jednak po odseparowaniu dwóch układów od siebie *lokalne* pomiary wykonywane na jednym układzie nie mogą mieć wpływu na drugi układ i *vice versa*, tj. statystyczny rozkład wyników a i b jest produktowy:

$$p(a,b|x,y,\lambda) = p(a|x,\lambda)p(b|y,\lambda). \tag{2.2}$$

Sam czynnik λ może jednak zmieniać się w różnych seriach pomiarowych (z pewnym rozkładem prawdopodobieństwa $p(\lambda)$). W celu oszacowania łącznego rozkładu prawdopodobieństwa $p(a,b|x,y)$ po wielu seriach pomiarowych musimy uwzględnić losowy charakter czynnika λ:

$$p(a,b|x,y) = \sum_{\Lambda} p(\lambda)p(a|x,\lambda)p(b|y,\lambda), \tag{2.3}$$

jeśli zmienna λ przyjmowała wartości dyskretne (gdzie Λ jest tutaj przestrzenią, na którą zmienna λ jest zdefiniowana), lub

$$p(a,b|x,y) = \int_{\Lambda} d\lambda\, p(\lambda)p(a|x,\lambda)p(b|y,\lambda), \tag{2.4}$$

jeśli zmienna λ przyjmowała wartości ciągłe. Wszystkie możliwe korelacje wyników opisane łącznym rozkładem prawdopodobieństwa $p(a,b|x,y)$, które można rozpisać w formie (2.3) lub (2.4) nazywamy *lokalnymi*. W przeciwnym przypadku będziemy mówić o korelacjach *nielokalnych*, albo krótko: *nielokalności*.

2.1.2 Nierówność CHSH

Poniżej wyprowadzimy nierówność CHSH (Clausera-Horna-Shimony'ego-Holta) [1], która jest najprostszą nierównością Bella. W ogólności, nierówności Bella są relacjami wyrażonymi w postaci nierówności, które muszą być spełnione przez *wszystkie* lokalne korelacje (wszystkie lokalne rozkłady prawdopodobieństwa $p(a,b|x,y)$). Załóżmy, że dwoje użytkowników, Alicja i Bolek, dokonują jednego z dwóch pomiarów $x,y \in \{0,1\}$ takich, że każdy pomiar może dać dwa różne wyniki $a,b \in \{-1,+1\}$. Oznaczmy przez $\langle a_x b_y \rangle$ wartość oczekiwaną iloczynu ab przy pomiarach x i y:

$$\langle a_x b_y \rangle = \sum_{a,b} ab\, p(a,b|x,y). \tag{2.5}$$

Przez $\langle a_x \rangle|_\lambda$ oznaczmy wartość oczekiwaną wyniku a warunkowanego przez λ:

$$\langle a_x \rangle|_\lambda = \sum_a a\, p(a|x,\lambda), \tag{2.6}$$

i analogicznie

$$\langle b_y \rangle|_\lambda = \sum_b b\, p(b|y,\lambda). \tag{2.7}$$

Załóżmy, że korelacje zadane przez łączny rozkład prawdopodobieństwa $p(a,b|x,y)$ są lokalne. Mamy wtedy:

$$\begin{aligned}
\langle a_x b_y \rangle &= \sum_{a,b} a\, b\, p(a,b|x,y) \\
&= \sum_{a,b} a\, b \int_\Lambda d\lambda\, p(\lambda) p(a|x,\lambda) p(b|y,\lambda) \\
&= \int_\Lambda d\lambda\, p(\lambda) \sum_a a\, p(a|x,\lambda) \sum_a a\, p(b|y,\lambda) \\
&= \int_\Lambda d\lambda\, p(\lambda) \langle a_x \rangle|_\lambda \langle b_y \rangle|_\lambda .
\end{aligned} \tag{2.8}$$

Rozważmy sumę $\langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle$. Mamy

$$\begin{aligned}
&\langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle \\
&= \int_\Lambda d\lambda \left(\langle a_0 \rangle|_\lambda \langle b_0 \rangle|_\lambda + \langle a_0 \rangle|_\lambda \langle b_1 \rangle|_\lambda + \langle a_1 \rangle|_\lambda \langle b_0 \rangle|_\lambda - \langle a_1 \rangle|_\lambda \langle b_1 \rangle|_\lambda \right) \\
&= \int_\Lambda d\lambda \left(\langle a_0 \rangle|_\lambda (\langle b_0 \rangle|_\lambda + \langle b_1 \rangle|_\lambda) + \langle a_1 \rangle|_\lambda (\langle b_0 \rangle|_\lambda - \langle b_1 \rangle|_\lambda) \right) \\
&\leq \int_\Lambda d\lambda \left(\left| \langle b_0 \rangle|_\lambda + \langle b_1 \rangle|_\lambda \right| + \left| \langle b_0 \rangle|_\lambda - \langle b_1 \rangle|_\lambda \right| \right),
\end{aligned} \tag{2.9}$$

gdzie w ostatniej linijce skorzystaliśmy z faktu, że wartości $\langle a_0 \rangle|_\lambda, \langle a_1 \rangle|_\lambda$ mieszczą się w przedziale liczbowym $[-1,1]$. Bez straty ogólności możemy dalej założyć, że

$$0 \leq \langle b_1 \rangle|_\lambda \leq \langle b_0 \rangle|_\lambda, \tag{2.10}$$

co w konsekwencji daje

$$\left| \langle b_0 \rangle|_\lambda + \langle b_1 \rangle|_\lambda \right| + \left| \langle b_0 \rangle|_\lambda - \langle b_1 \rangle|_\lambda \right| \leq 2 \langle b_0 \rangle|_\lambda \leq 2, \tag{2.11}$$

gdzie skorzystaliśmy z faktu, że również wartości $\langle b_0 \rangle|_\lambda, \langle b_1 \rangle|_\lambda$ mieszczą się w przedziale liczbowym $[-1,1]$. Jeśli powyższą nierówność scałkujemy po całej przestrzeni

zmiennych Λ, to po wstawieniu do (2.9) otrzymamy nierówność CHSH:

$$\langle a_0 b_0\rangle + \langle a_0 b_1\rangle + \langle a_1 b_0\rangle - \langle a_1 b_1\rangle \leq 2, \tag{2.12}$$

która musi być spełniona przez wszystkie lokalne korelacje postaci (2.4) (w przypadku dyskretnych wartości zmiennych λ analogicznie wyprowadzamy nierówność CHSH, która musi być spełniona przez wszystkie lokalne korelacje postaci (2.3)).

2.1.3 Korelacje kwantowe i łamanie nierówności CHSH

Zgodnie z postulatami mechaniki kwantowej korelacje opisywane łącznym rozkładem prawdopodobieństwa $p(a,b|x,y)$ dane są wyrażeniem

$$p(a,b|x,y) = \mathrm{Tr}\left(M_{a|x} \otimes M_{b|y}\rho_{AB}\right), \tag{2.13}$$

gdzie ρ_{AB} to macierz gęstości opisująca stan kwantowy współdzielony przez dwoje użytkowników, Alicję i Bolka, którzy dokonują pomiarów opisywanych operatorami pomiarowymi $\{M_{a|x}\}$ oraz $\{M_{b|y}\}$. W przypadku, gdy Alicja i Bolek dokonują kompletnych pomiarów von Neumanna $\{\Pi_{a|x}\}$ oraz $\{\Pi_{b|y}\}$ na stanie czystym $|\Psi\rangle$ łączny rozkład prawdopodobieństwa (2.13) można zapisać równoważnie jako

$$p(a,b|x,y) = \langle\Psi|\Pi_{a|x} \otimes \Pi_{b|y}|\Psi\rangle. \tag{2.14}$$

Niech teraz zbiory wartości wyników pomiarów $\{a_x\}$ oraz $\{b_y\}$ reprezentują wartości własne obserwabli odpowiednio $\hat{A}$ oraz $\hat{B}$:

$$\hat{A} = \sum a_x M_{a|x}, \tag{2.15}$$

$$\hat{B} = \sum b_y M_{b|y}. \tag{2.16}$$

Wartości oczekiwane $\langle a_x\rangle$, $\langle b_y\rangle$ i $\langle a_x b_y\rangle$ wyrażają się odpowiednio przez:

$$\langle a_x\rangle = \mathrm{Tr}\left(\hat{A} \otimes id_{d\times d}\rho_{AB}\right), \tag{2.17}$$

$$\langle b_y\rangle = \mathrm{Tr}\left(id_{d\times d} \otimes \hat{B}\rho_{AB}\right), \tag{2.18}$$

$$\langle a_x b_y\rangle = \mathrm{Tr}\left(\hat{A} \otimes \hat{B}\rho_{AB}\right), \tag{2.19}$$

(gdzie d jest wymiarem podprzestrzeni Hilberta podukładu). Analogicznie dla stanów czystych mamy

$$\langle a_x\rangle = \langle\Psi|\hat{A} \otimes id_{d\times d}|\Psi\rangle, \tag{2.20}$$

$$\langle b_y\rangle = \langle\Psi|id_{d\times d} \otimes \hat{B}|\Psi\rangle, \tag{2.21}$$

$$\langle a_x b_y\rangle = \langle\Psi|\hat{A} \otimes \hat{B}|\Psi\rangle. \tag{2.22}$$

Przykład. Załóżmy, że Alicja i Bolek współdzielą stan maksymalnie splątany

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \tag{2.23}$$

na którym dokonują pomiarów odpowiednio $\vec{x}\cdot\vec{\sigma}$ oraz $\vec{y}\cdot\vec{\sigma}$, gdzie $\vec{x}$, $\vec{y}$ to wektory jednostkowe, natomiast $\vec{\sigma} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ to wektor macierzy Pauliego. Wartości wyników pomiarów w tym przypadku należą do zbioru $\{-1, +1\}$, natomiast wartość oczekiwana iloczynu dana jest wyrażeniem

$$\begin{aligned}\langle a_x b_y\rangle &\equiv \langle\Psi^-|\vec{x}\cdot\vec{\sigma}\otimes\vec{y}\cdot\vec{\sigma}|\Psi^-\rangle \\ &= -\vec{x}\cdot\vec{y}.\end{aligned} \tag{2.24}$$

Niech teraz pomiary Alicji $x \in \{0,1\}$ odpowiadają pomiarom operatorów $\vec{x}\cdot\vec{\sigma}$ wzdłuż dwóch dowolnych ortogonalnych kierunków $\hat{e}_0$ i $\hat{e}_1$, natomiast pomiary Bolka $y \in \{0,1\}$ odpowiadają pomiarom operatorów $\vec{y}\cdot\vec{\sigma}$ wzdłuż kierunków wyznaczanych przez $\frac{-(\hat{e}_0+\hat{e}_1)}{\sqrt{2}}$ oraz $\frac{-\hat{e}_0+\hat{e}_1}{\sqrt{2}}$. W tym przypadku mamy

$$\langle a_0 b_0\rangle = \langle a_0 b_1\rangle = \langle a_1 b_0\rangle = \frac{1}{\sqrt{2}} \tag{2.25}$$

$$\langle a_1 b_1\rangle = -\frac{1}{\sqrt{2}}. \tag{2.26}$$

Zauważmy, że zachodzi

$$\langle a_0 b_0\rangle + \langle a_0 b_1\rangle + \langle a_1 b_0\rangle - \langle a_1 b_1\rangle = 2\sqrt{2}, \tag{2.27}$$

co łamie nierówność CHSH (2.12) prawdziwą dla wszystkich korelacji lokalnych postaci (2.4). Tym samym wnioskujemy, że w ogólności korelacje kwantowe (2.13) nie są lokalne. Pozwala to mówić o nielokalnych własnościach korelacji w mechanice kwantowej.

W powyższym przykładzie łamanie nierówności CHSH było możliwe przy wykorzystaniu odpowiednio dobranych pomiarów, które wykonywane są na stanie kwantowym, będącym stanem maksymalnie splątanym. Zauważmy, że wykorzystanie stanu splątanego jest warunkiem koniecznym do otrzymania korelacji nielokalnych [15]. Rzeczywiście, jeśli pomiary byłyby wykonywane na stanie separowalnym, który w ogólności możemy zapisać w postaci

$$\rho_{AB} = \sum_\ell p(\ell)\rho_A^\ell \otimes \rho_B^\ell, \tag{2.28}$$

gdzie stany ρ_A^ℓ (ρ_B^ℓ) są stanami warunkowymi generowanymi w zależności od klasycznej zmiennej losowej ℓ, to dokonując pomiarów opisywanych operatorami pomiaro-

wymi $\{M_{a|x}\}$ oraz $\{M_{b|y}\}$, otrzymywane korelacje $p(a,b|x,y)$ będą dane wyrażeniem

$$\begin{aligned} p(a,b|x,y) &= \mathrm{Tr}\left(M_{a|x}\otimes M_{b|y}\sum_{\ell} p(\ell)\rho_A^{\ell}\otimes\rho_B^{\ell}\right) \\ &= \sum_{\ell} p(\ell)\,\mathrm{Tr}(M_{a|x}\rho_A^{\ell})\,\mathrm{Tr}(M_{b|y}\rho_B^{\ell}) \\ &= \sum_{\ell} p(\ell)\,p(a|x,\ell)\,p(b|y,\ell). \end{aligned} \tag{2.29}$$

Ponieważ wyrażenie to ma postać (2.3) wnioskujemy, że korelacje mają wyłącznie charakter lokalny.

Chociaż wykorzystanie kwantowych stanów splątanych jest konieczne do uzyskania korelacji nielokalnych, w ogólności nie jest łatwe stwierdzenie, czy wykorzystując pewien określony stan splątany można uzyskać nielokalne korelacje (innymi słowy: czy dany stan kwantowy jest nielokalny). Istnieje natomiast kryterium jednoznacznie stwierdzające, czy określony dwuqubitowy stan kwantowy łamie nierówność CHSH [24]. Rozpatrzmy przypadek, w którym Alicja i Bolek wykonują pomiary na stanie ρ_{AB}, którego podukłady A i B mają wymiar $d=2$ („qubity"). Dwuqubitowy stan ρ_{AB} łamie nierówność CHSH wtedy i tylko wtedy, gdy [24]:

$$\sqrt{\lambda_i+\lambda_j}>1, \tag{2.30}$$

gdzie λ_i oraz λ_j są odpowiednio największą oraz drugą co do wielkości wartością własną macierzy R^TR, zaś macierz R jest zdefiniowana poprzez następujące elementy macierzowe:

$$R_{ij}=\mathrm{Tr}\left((\hat{\sigma}_i\otimes\hat{\sigma}_j)\rho_{AB}\right), \tag{2.31}$$

gdzie $\hat{\sigma}_i$ to macierze Pauliego. Zauważmy jednak, że powyższe kryterium dotyczy wyłącznie nielokalności stwierdzanej poprzez łamanie nierówności CHSH. Jak wspomnimy później istnieją również stany nielokalne (łamiące pewną nierówność Bella), które nie łamią nierówności CHSH [25].

Rozważmy raz jeszcze wyrażenie $\langle a_0b_0\rangle+\langle a_0b_1\rangle+\langle a_1b_0\rangle-\langle a_1b_1\rangle$. W mechanice kwantowej jest ono równoważne wartości oczekiwanej operatora $\hat{\mathscr{B}}_{\mathrm{CHSH}}$, zdefiniowanego przez

$$\hat{\mathscr{B}}_{\mathrm{CHSH}}=\hat{A}_1(\hat{B}_1+\hat{B}_2)+\hat{A}_2(\hat{B}_1-\hat{B}_2), \tag{2.32}$$

gdzie $\hat{A}_x$ ($\hat{B}_y$) są operatorami o wartościach własnych $a_x, b_y=\pm1$. W przedstawionym wcześniej przykładzie pokazaliśmy, że dla pewnych określonych pomiarów otrzymujemy $\langle\hat{\mathscr{B}}_{\mathrm{CHSH}}\rangle=2\sqrt{2}$. Powstaje naturalne pytanie, czy dla operatora $\hat{\mathscr{B}}_{\mathrm{CHSH}}$ można uzyskać większą wartość oczekiwaną, a tym samym osiągnąć jeszcze większe łamanie nierówności CHSH. Pytanie to jest w istocie równoważne z pytaniem o największą

wartość własną operatora $\hat{\mathscr{B}}_{\text{CHSH}}$. Można pokazać [26, 27], że największa wartość własna operatora

$$\hat{\mathscr{B}}^2_{\text{CHSH}} = 4 + [\hat{A}_1, \hat{A}_2][\hat{B}_1, \hat{B}_2] \tag{2.33}$$

wynosi 8, a zatem wartości własne operatora $\hat{\mathscr{B}}_{\text{CHSH}}$ są ograniczone przez $2\sqrt{2}$. Tym samym uzyskujemy górne ograniczenie na łamanie nierówności CHSH przez korelacje w mechanice kwantowej. Ograniczenie to nazywamy *ograniczeniem Tsirelsona*.

2.1.4 Inne nierówności Bella

W poprzednim podrozdziale wyprowadziliśmy nierówność CHSH, która jest najprostszą nierównością Bella, opartą na założeniu, że mamy dwoje obserwatorów, z których każdy może wykonać jeden z dwóch pomiarów dających dwa możliwe wyniki. W ogólności jednak można skonstruować również inne nierówności Bella [23], w zależności od ilości obserwatorów, ilości różnych pomiarów oraz ilości wyników pomiarów [28–47]. Z reguły jest to zadanie wysoce nietrywialne. Wiele nierówności Bella (ograniczamy się do dwóch obserwatorów) ma postać odpowiedniej kombinacji liniowej prawdopodobieństw dla łącznego rozkładu prawdopodobieństwa $p(a, b|x, y)$:

$$\sum_{a,b,x,y} M^{ab}_{xy} p(a, b|x, y) \leq \mathscr{C} \tag{2.34}$$

gdzie M^{ab}_{xy} to rzeczywiste współczynniki, natomiast $\mathscr{C}$ jest nietrywialnym ograniczeniem dla wszystkich lokalnych rozkładów prawdopodobieństwa $p(a, b|x, y)$. W przypadku nierówności CHSH mamy przykładowo:

$$M^{ab}_{xy} = (-1)^{a+b+xy}, \tag{2.35}$$

$$\mathscr{C} = 2, \tag{2.36}$$

gdzie $a, b, x, y \in \{0, 1\}$.

Dla pewnych współczynników M^{ab}_{xy} nierówności Bella mogą być zapisane w formie nierówności, które muszą być spełniane przez odpowiednią sumę wartości oczekiwanych iloczynów wyników pomiarów (korelatorów), jak np. w postaci (2.12).

Poniżej przedstawimy przykład innej nierówności Bella, która w odróżnieniu od nierówności CHSH zakłada większą ilość pomiarów dla każdego z dwóch obserwatorów.

Rozważmy przypadek, gdzie każdy z dwojga użytkowników może wykonać N pomiarów, których binarne wyniki oznaczymy przez a_i dla pomiarów Alicji i b_j dla pomiarów Bolka ($i, j \in \{0, 1, ..., N-1\}$). Dla dowolnych lokalnych rozkładów prawdopodobieństwa zachodzi

$$\langle a_0 b_0 \rangle + \langle a_1 b_0 \rangle + \langle a_1 b_1 \rangle + \langle a_2 b_1 \rangle + ... + \langle a_{N-1} b_{N-1} \rangle - \langle a_0 b_{N-1} \rangle \leq 2N - 2. \tag{2.37}$$

Nierówność tę nazywamy nierównością Braunsteina-Cavesa [48]. Zauważmy, że nierówność CHSH jest szczególnym przypadkiem powyższej nierówności dla $N = 2$.

2.1.5 Korelacje w teoriach niesygnalizujących

Mechanika kwantowa spełnia zasadę niesygnalizowania. Zasada ta mówi, że statystyka wyników pomiaru u pewnego obserwatora (Alicji) nie może w żaden sposób zależeć od wyboru pomiaru obserwatora będącego przestrzennie od niego odseparowanego (Bolka). Możemy ją zapisać w formie równań

$$\forall_{a,x,y,x'} \sum_b p(a,b|x,y) = \sum_b p(a,b|x',y), \tag{2.38}$$

$$\forall_{b,x,y,y'} \sum_a p(a,b|x,y) = \sum_a p(a,b|x,y'), \tag{2.39}$$

lub krócej:

$$p(a|x) = p(a|x,y) = \sum_b p(a,b|x,y), \tag{2.40}$$

$$p(b|y) = p(b|x,y) = \sum_a p(a,b|x,y). \tag{2.41}$$

Jeśli któryś z warunków (2.38) nie byłby spełniony, wtedy odpowiednie korelacje $p(a,b|x,y)$ byłyby w konflikcie ze szczególną teorią względności, a mianowicie dysponując wieloma kopiami układów jeden użytkownik mógłby poprzez swój wybór obserwabli *natychmiastowo* zmieniać obserwowalną statystykę wyników u drugiego użytkownika, co dawałoby możliwość natychmiastowej komunikacji. Dany układ określilibyśmy wtedy jako *sygnalizujący*. Okazuje się jednak, że zbiór teorii probabilistycznych spełniających zasadę niesygnalizowania jest szerszy niż mechanika kwantowa [2,3,49–51]. Na poniższym przykładzie pokażemy, że zasada niesygnalizowania dopuszcza nielokalne korelacje, których nie dopuszcza mechanika kwantowa.

Przykład. Załóżmy, że Alicja i Bolek dysponują pewnym układem, na którym każdy z nich może wykonywać jeden z dwóch pomiarów $x \in \{0,1\}$ przez Alicję i dwóch pomiarów $y \in \{0,1\}$ przez Bolka. Jeśli wyniki pomiarów uzyskiwanych przez Alicję $a \in \{0,1\}$ i Bolka $b \in \{0,1\}$ mają rozkład prawdopodobieństwa taki, że

$$p(a,b|x,y) = \begin{cases} \frac{1}{2} & \text{dla } a \oplus b = xy, \\ 0 & \text{w przeciwnym wypadku,} \end{cases} \tag{2.42}$$

to taką rodzinę rozkładów prawdopodobieństwa będziemy nazywać *układem Popescu-Rohrlicha* [2,3] i oznaczać skrótem *PR* (więcej o układzie *PR* p. Rozdz. 4 oraz 6). Wypisując jawnie wszystkie 4 łączne rozkłady prawdopodobieństwa łatwo pokazać, że układ *PR* spełnia warunki (2.38), a zatem jest niesygnalizujący.

Zauważmy, że warunek

$$a \oplus b = xy \tag{2.43}$$

dla wyników $a, b \in \{0, 1\}$ jest równoważny warunkowi

$$\frac{1}{2}(1 - ab) = xy \tag{2.44}$$

dla wyników $a, b \in \{-1, +1\}$ (0 dla wyników takich samych, 1 dla wyników przeciwnych). Tak dobrane przenumerowanie wyników nie zmieni oczywiście rozkładu ich statystyk. Wyznaczmy teraz wartość oczekiwaną iloczynu $\langle a_x b_y \rangle = \sum_{a,b} abp(a, b|x, y)$ wykorzystując statystyki układu PR i relację (2.44):

$$\langle a_x b_y \rangle = \begin{cases} p(+1,+1|x,y) + p(-1,-1|x,y) = 1 & \text{dla } xy = 0, \\ -p(-1,+1|x,y) - p(+1,-1|x,y) = -1 & \text{dla } xy = 1. \end{cases} \tag{2.45}$$

Dla układu PR otrzymujemy zatem

$$\langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle = 4, \tag{2.46}$$

co równocześnie jest maksymalną algebraiczną wartością dla tego wyrażenia. Widzimy, że łamiąc ograniczenie Tsirelsona $2\sqrt{2}$ układ PR nie może być zrealizowany z wykorzystaniem korelacji w ramach mechaniki kwantowej, a tym bardziej z wykorzystaniem korelacji lokalnych postaci (2.4).

2.1.6 Nielokalność a kody swobodnego dostępu

Rozpatrzmy teraz sytuację, w której zadanie pewnego użytkownika (Bolka) polega na odgadnięciu bitów informacji $x_1, ..., x_n$ posiadanych przez innego użytkownika (Alicję). *Kodem swobodnego dostępu* (n, p) nazywamy funkcjonalność, dzięki której Bolek może poznać wartość dowolnego z n bitów Alicji z prawdopodobieństwem równym przynajmniej p. W niniejszej pracy będziemy zajmować sie wyłącznie kodami $(2, p)$, w szczególności zaś skrótem KSD oznaczać będziemy kod $(2, 1)$, tj. funkcjonalność, dzięki której Bolek może z pewnością poznać wartość jednego, ale dowolnego, z dwóch posiadanych przez nią bitów.

Jak pokażemy poniżej, istnieje protokół [52–54] pozwalający na otrzymanie KSD $(2, 1)$ przy wykorzystaniu układu PR i przesłaniu jednego bitu informacji. Załóżmy, że Alicja dysponuje dwoma bitami x_0 oraz x_1, z kolei zadaniem Bolka jest wygenerowanie za pomocą zasobów jemu dostępnych wartości x_y, gdzie indeks $y = 0, 1$ sygnuje wybór bitu Alicji, który chce poznać. Zauważmy, że sam układ PR działa jak funkcjonalność, której jednobitowe wejścia x, y i wyjścia a, b spełniają warunek $a \oplus b = xy$.

Niech wejściem Alicji będzie suma binarna x_0 i x_1:

$$x = x_0 \oplus x_1. \tag{2.47}$$

Wtedy mamy oczywiście

$$a \oplus b = (x_0 \oplus x_1)y. \tag{2.48}$$

Jeśli do powyższego równania obustronnie dodamy x_0, to otrzymujemy wartość równą wartości x_y:

$$b \oplus a \oplus x_0 = x_0 \oplus y(x_0 \oplus x_1) \equiv x_y. \tag{2.49}$$

Widzimy zatem, że Bolek chcąc poznać wartość dowolnego bitu x_y, musi jedynie do swojego wyniku z układu PR b dodać binarnie wartość $m = a \oplus x_0$, która jest wartością bitu przesłanego mu przez Alicję. Tym samym KSD można symulować za pomocą układu PR przy dodatkowym przesłaniu jednego bitu informacji.

W powyższym protokole założyliśmy, że Alicja i Bolek dysponują układem PR, którego nie da się zrealizować w ramach mechaniki kwantowej. Naturalnym jest zatem pytanie, czy można zasymulować KSD korzystając wyłącznie z zasobów kwantowych (plus przesłanie jednego bitu informacji). Jak pokażemy poniżej, optymalny protokół [55] wykorzystujący odpowiednio dobrane pomiary wykonywane na stanie maksymalnie splątanym pozwala osiągnąć kod swobodnego dostępu $(2, p_C)$, gdzie

$$p_C = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right) \approx 0,8536. \tag{2.50}$$

Niech Alicja i Bolek współdzielą stan maksymalnie splątany (2.23), na którym dokonują pomiarów odpowiednio $\vec{x}\cdot\vec{\sigma}$ oraz $\vec{y}\cdot\vec{\sigma}$, gdzie pomiary Alicji (Bolka) sygnowane $x \in \{0,1\}$ ($y \in \{0,1\}$) odpowiadają pomiarom operatorów zdefiniowanych identycznie jak w przykładzie z Podrozdz. 2.1.3. Teraz, z definicji wartości oczekiwanej $\langle a_x b_y \rangle$ oraz (2.24), mamy:

$$p(a = b|x,y) = \frac{1 - \vec{x}\cdot\vec{y}}{2}. \tag{2.51}$$

Jeśli dodatkowo oznaczymy wyniki pomiarów $-1, +1$ przez odpowiednio $0, 1$ (por. warunek (2.44)), to dla wybranych ustawień pomiarów wynik (2.51) możemy zapisać jako

$$p(a \oplus b = xy|x,y) = \frac{1}{2}\left(1 + \frac{1}{\sqrt{2}}\right), \tag{2.52}$$

(dla układu PR $p(a\oplus b = xy|x,y) = 1$). Wykorzystując następnie identyczny protokół kodowania wejść x, y oraz przesyłanej informacji m jak w przypadku wykorzystania

układu *PR*, widzimy, że warunek

$$x_y = b \oplus m \tag{2.53}$$

zachodzi z prawdopodobieństwem równym $\frac{1}{2}\left(1+\frac{1}{\sqrt{2}}\right)$. Tym samym, prawdopodobieństwo odgadnięcia przez Bolka jednego ale dowolnego z dwóch bitów Alicji przy wykorzystaniu zasobów kwantowych jest równe w przybliżeniu $0,8536$.

2.2. Zasady nieoznaczoności

Jedną z fundamentalnych cech kwantowomechanicznego opisu świata jest tzw. zasada nieoznaczoności, która mówi, że w ogólności nie jest możliwe dokładne poznanie dwóch różnych wielkości opisujących dany układ fizyczny. W tym podrozdziale przytoczymy relacje nieoznaczoności, zarówno w powszechnie znanej formie danej przez Robertsona [56], jak również w formie tzw. nierówności entropowych [57]. Wskażemy też na związek zasad nieoznaczoności z tzw. zasadami wykluczania informacji [4].

2.2.1 Pierwotne zasady nieoznaczoności

Już przy samych narodzinach mechaniki kwantowej zdawano sobie sprawę na fundamentalne różnice między kwantowym i klasycznym opisem świata. Jednym z pierwszych przykładów była słynna zasada nieoznaczoności Heisenberga [58] dla dwóch (niekomutujących ze sobą) obserwabli położenia i pędu, $\hat{x}$ i $\hat{p}$, wyrażona w formie

$$\Delta\hat{x}\Delta\hat{p} \geq \frac{1}{2}\hbar, \tag{2.54}$$

gdzie $\Delta\hat{O}$ jest odchyleniem standardowym obserwabli O otrzymywanym przy pomiarze na stanie kwantowym $|\Psi\rangle$:

$$\Delta\hat{O} = \sqrt{\langle\Psi|O^2|\Psi\rangle - \langle\Psi|O|\Psi\rangle^2}. \tag{2.55}$$

Tak ujęta zasada nieoznaczoności została następnie uogólniona przez Robertsona [56] na przypadek dwóch dowolnych obserwabli $\hat{X}$ i $\hat{Y}$:

$$\Delta\hat{X}\Delta\hat{Y} \geq \frac{1}{2}|\langle\Psi|[\hat{X},\hat{Y}]|\Psi\rangle|. \tag{2.56}$$

Powyższa nierówność mówi, że nie jest możliwe jednoczesne określenie dwóch wielkości fizycznych z dowolną dokładnością, jeśli obserwable opisujące te wielkości nie komutują ze sobą. Im dokładniej znamy jedną wartość, tym większa niepewność przypisana jest drugiej wielkości i *vice versa*. Należy jednak zauważyć, że powyższa nie-

równość posiada pewne cechy, świadczące o tym, że tak ujęta relacja nieoznaczoności nie jest zadowalająca. Po pierwsze, zależy ona bezpośrednio od postaci stanu kwantowego, na którym wykonywane są pomiary. Po drugie, dla pewnych obserwabli można uzyskać nierówność trywialną (wartość oczekiwana komutatora równa zeru), choć odchylenia standardowe obu obserwabli są niezerowe.

2.2.2 Entropowe zasady nieoznaczoności

Rozważmy entropię Shannona [59] rozkładu prawdopodobieństwa dla pewnej obserwabli X, zdefiniowaną jako

$$H(X)_\rho = -\sum_{i=1}^{d} p_i(X,\rho)\log_2 p_i(X,\rho), \tag{2.57}$$

gdzie $p_i(X,\rho)$ to prawdopodobieństwo otrzymania i-tego wyniku wyniku pomiaru obserwabli X jeśli pomiar wykonujemy na stanie kwantowym ρ o wymiarze d. W dalszej części, o ile nie zostanie zaznaczone inaczej, przez wielkość $H(\mathscr{A})_\rho$ (gdzie $\mathscr{A}$ oznacza pomiar w bazie $\{|a_i\rangle\}$ z odpowiadającymi im jej operatorami rzutowymi $P_i = |a_i\rangle\langle a_i|$) rozumieć będziemy entropię dla rozkładu prawdopodobieństwa $\{p_i\}$, gdzie

$$p_i = \mathrm{Tr}(P_i\rho). \tag{2.58}$$

Określiwszy miarę niepewności dla dowolnej obserwabli, powstaje naturalne pytanie, czy istnieje zasada nieoznaczoności dla dwóch obserwabli w postaci

$$H(\mathscr{A})_\rho + H(\mathscr{B})_\rho \geq C(\mathscr{A},\mathscr{B}), \tag{2.59}$$

gdzie $C(\mathscr{A},\mathscr{B})$ to w ogólności pewne nietrywialne ograniczenie zależne tylko od pomiarów $\mathscr{A}$ i $\mathscr{B}$. Istotnie, pierwsza entropowa zasada nieoznaczoności dla operatorów położenia i pędu została zaproponowana w [60] i ulepszona następnie w [61] oraz [62, 63], gdzie pokazano, że zasada nieoznaczoności Heisenberga wynika z entropowej zasady nieoznaczoności. Tym samym pokazano, że nierówności entropowe stanowią ogólniejszą formę zasad nieoznaczoności.

Ważnym wkładem do entropowych zasad nieoznaczoności była relacja nieoznaczoności dla dowolnych pomiarów $\mathscr{A}$ i $\mathscr{B}$ zaproponowana w pracy Deutscha [64]:

$$H(\mathscr{A})_\rho + H(\mathscr{B})_\rho \geq -2\log_2\left(\frac{1+\sqrt{c}}{2}\right), \tag{2.60}$$

gdzie współczynnik c jest definiowany jako

$$c = \max_{i,j} |\langle a_i|b_j\rangle|^2, \tag{2.61}$$

przy czym

$$\frac{1}{d} \leq c \leq 1. \tag{2.62}$$

Kraus [65] zaproponował ulepszoną zasadę nieoznaczoności, udowodnioną później przez Maassena i Uffinka [5], w formie nierówności

$$H(\mathscr{A})_\rho + H(\mathscr{B})_\rho \geq -\log_2 c. \tag{2.63}$$

Zwróćmy jednak uwagę, że w przypadku, gdy bazy pomiarów $\mathscr{A}$ i $\mathscr{B}$ współdzielą choćby jeden wspólny wektor, natychmiast dostajemy $c = 1$, co jest zgodne z tym, że istnieje stan, dla którego możemy przewidzieć wyniki pomiarów obserwabli $\mathscr{A}$ i $\mathscr{B}$.

Ważnym rozszerzeniem entropowych zasad nieoznaczoności są zasady nieoznaczoności z pamięcią kwantową [66–70]. Nim je przedstawimy, rozważmy najpierw następującą grę między dwoma użytkownikami Alicją i Bolkiem. Niech Alicja przygotuje pewien stan kwantowy, po czym wyśle go lub jego część do Bolka. Bolek wykonuje jeden z dwóch pomiarów $\mathscr{B}^{(1)}$ lub $\mathscr{B}^{(2)}$, po czym informuje Alicję o wybranym przez siebie pomiarze. Jej zadaniem jest przewidzieć z najmniejszą możliwą niepewnością wynik pomiaru Bolka. Zauważmy, że jeśli Alicja prześle cały układ do Bolka, jej niepewność jest zawsze ograniczona, gdyż

$$H(\mathscr{B}^{(1)}) + H(\mathscr{B}^{(2)}) \geq -\log_2 c. \tag{2.64}$$

Okazuje się jednak, że Alicja może pokonać to ograniczenie w przypadku, gdy posiada pamięć kwantową w postaci podukładu, który jest splątany z układem przekazywanym Bolkowi. W tym przypadku, relacja nieoznaczoności z pamięcią kwantową przyjmuje postać

$$S(\mathscr{B}^{(1)}|A) + S(\mathscr{B}^{(2)}|A) \geq -\log_2 c + S(B|A), \tag{2.65}$$

gdzie A i B oznaczają podukłady odpowiednio Alicji i Bolka, natomiast $S(\mathscr{B}^{(s)}|A)$ i $S(B|A)$ to warunkowe entropie von Neumanna. Zauważmy, że w przypadku gdy podukłady Alicji i Bolka są splątane, to $S(B|A) < 0$ i tym samym niepewność wyniku pomiaru w wyżej wspomnianej grze może mieć wartość mniejszą niż $-\log_2 c$.

Przytoczymy przy tym trzy interesujące przypadki zasady nieoznaczoności z pamięcią kwantową [66].

- Jeśli podukład B jest maksymalnie splątany z podukładem A, to $S(B|A) = -\log_2 d$ i wtedy ze względu na fakt, że $-\log_2 c \leq -\log_2 d$ otrzymujemy trywialne ograniczenie

$$S(\mathscr{B}^{(1)}|A) + S(\mathscr{B}^{(2)}|A) \geq 0. \tag{2.66}$$

  Wobec tego Alicja ma możliwość dokładnego przewidzenia wyniku pomiaru

Bolka.

- Jeśli układ B nie jest splątany z A, to $S(B|A) \geq 0$. Jeśli dodatkowo weźmiemy pod uwagę, że warunkowanie zmniejsza entropię $H(\mathscr{B}^{(s)}|A) \leq S(\mathscr{B}^{(s)})$, to z relacji nieoznaczoności (2.65) otrzymujemy relację Maassena-Uffinka (2.63).

- W przypadku braku pamięci kwantowej relacja (2.65) redukuje się do

$$H(\mathscr{B}^{(1)}) + H(\mathscr{B}^{(2)}) \geq -\log_2 c + S(B). \tag{2.67}$$

  Jeśli teraz stan układu B jest czysty, to powyższa relacja ponownie redukuje się do relacji Maassena-Uffinka (2.63). Jeśli natomiast stan układu B jest stanem mieszanym, wtedy $S(B) > 0$ i powyższa relacja jest silniejszym ograniczeniem niż (2.63). Widzimy jednak, że mówimy tutaj o relacji zależnej od stanu, natomiast relacja (2.63) jest zasadą nieoznaczoności, która nie zależy od stanu układu.

2.2.3 Zasady wykluczania informacji

We wcześniejszych rozważaniach mówiliśmy o nieoznaczoności wyrażonej jako suma niepewności wyników pomiarów różnych wielkości fizycznych. W tym kontekście należy zauważyć, że jeżeli suma niepewności jest zawsze ograniczona z dołu, to w przypadku, kiedy mielibyśmy pełną informację o wyniku jednego pomiaru, nie moglibyśmy jednocześnie posiadać pełnej informacji o wyniku innego pomiaru. Owo stwierdzenie wyrazić można w postaci zasady wykluczania informacji [4, 71], która ogranicza z góry sumę dostępnych informacji o wynikach różnych pomiarów. W dalszej części przytoczymy wyprowadzenie relacji wykluczania informacji przedstawionej w pracy [4] w postaci:

$$I(\mathscr{B}^{(1)}|\mathscr{E}) + I(\mathscr{B}^{(2)}|\mathscr{E}) \leq 2\log_2 d + \log_2 c, \tag{2.68}$$

gdzie $I(\mathscr{B}|\mathscr{E})$ jest *dostępną informacją* uzyskiwaną przy pomiarze $\mathscr{B}$

$$I(\mathscr{B}|\mathscr{E}) = H(\mathscr{B})_{\rho_\mathscr{E}} - \sum_i p_i H(\mathscr{B})_{\rho_i} \tag{2.69}$$

na zespole stanów $\mathscr{E}$ danym przez $\mathscr{E} = \{\rho_i, p_i\}$, gdzie

$$\rho_\mathscr{E} = \sum_i p_i \rho_i. \tag{2.70}$$

Zauważmy przy tym, że przygotowanie zespołu stanów powyższej postaci w naturalny sposób odpowiada sytuacji, gdy zostaje wykonany pomiar na jednym układzie dwuukładowego stanu kwantowego ρ_{AB}. Istotnie, jeśli przykładowo Alicja dokona po-

miaru $\mathscr{A}$ w bazie $\{P_i\}$, to stan całego układu po pomiarze ma postać

$$\rho'_{AB} = \sum_i p_i \rho^i_{AB}, \tag{2.71}$$

gdzie

$$p_i = \mathrm{Tr}(P_i \otimes id_{d\times d}\rho_{AB}) \tag{2.72}$$

to prawdopodobieństwo otrzymania i-tego wyniku pomiaru, natomiast ρ^i_{AB} to stany postaci

$$\rho^i_{AB} = \frac{P_i \otimes id_{d\times d}\rho_{AB} P_i \otimes id_{d\times d}}{\mathrm{Tr}(P_i \otimes id_{d\times d}\rho_{AB})}. \tag{2.73}$$

Zgodnie z definicją entropii warunkowej mamy

$$H(\mathscr{B}|\mathscr{A})_{\rho_{AB}} = \sum_i p_i H(\mathscr{B})_{\rho^i_B}, \tag{2.74}$$

gdzie $H(\mathscr{B})_{\rho^i_B}$ liczone są na stanach warunkowych podukładu Bolka $\rho^i_B = \mathrm{Tr}_A \rho^i_{AB}$. Tym samym wiedząc, że wzajemna informacja pomiarów $\mathscr{A}$ i $\mathscr{B}$

$$I(\mathscr{A}:\mathscr{B})_{\rho_{AB}} = H(\mathscr{B})_{\rho_B} - H(\mathscr{B}|\mathscr{A})_{\rho_{AB}}, \tag{2.75}$$

możemy zasadę wykluczania informacji Halla (2.68) przepisać w postaci

$$I(\mathscr{A}:\mathscr{B}^{(1)})_{\rho_{AB}} + I(\mathscr{A}:\mathscr{B}^{(2)})_{\rho_{AB}} \leq 2\log_2 d + \log_2 c, \tag{2.76}$$

która w tej formie bezpośrednio nabiera znaczenia jako nierówność ograniczająca korelacje wyników pomiarów uzyskiwanych przez dwóch obserwatorów dokonujących określone pomiary.

Dowód relacji (2.76) wynika bezpośrednio z zasady nieoznaczoności Maassena-Uffinka [5], co pokażemy poniżej. Zgodnie z tym co zostało powiedziane powyżej, suma dwóch wzajemnych informacji jest równa

$$\begin{aligned} &I(\mathscr{A}:\mathscr{B}^{(1)})_{\rho_{AB}} + I(\mathscr{A}:\mathscr{B}^{(2)})_{\rho_{AB}} \\ &= H(\mathscr{B}^{(1)})_{\rho_B} + H(\mathscr{B}^{(2)})_{\rho_B} - H(\mathscr{B}^{(1)}|\mathscr{A})_{\rho_{AB}} - H(\mathscr{B}^{(2)}|\mathscr{A})_{\rho_{AB}} \\ &= H(\mathscr{B}^{(1)})_{\rho_B} + H(\mathscr{B}^{(2)})_{\rho_B} - \sum_i p_i \left(H(\mathscr{B}^{(1)})_{\rho^i_B} + H(\mathscr{B}^{(2)})_{\rho^i_B} \right), \end{aligned} \tag{2.77}$$

(w dalszej części, jeśli będzie wynikać to z kontekstu, będziemy pomijać oznaczenie stanu, na którym liczone są wartości entropii i wzajemnych informacji). Zwróćmy uwagę, że dla każdego i zgodnie z zasadą nieoznaczoności Maassena-Uffinka mamy:

$$H(\mathscr{B}^{(1)})_{\rho^i_B} + H(\mathscr{B}^{(2)})_{\rho^i_B} \geq -\log_2 c, \tag{2.78}$$

przy czym współczynnik c zależy wyłącznie od wektorów własnych obserwabli $\mathscr{B}^{(1)}$ i $\mathscr{B}^{(2)}$. Tym samym ograniczenie na sumę dwóch wzajemnych informacji wyraża się przez

$$\begin{aligned} I(\mathscr{A}:\mathscr{B}^{(1)})+I(\mathscr{A}:\mathscr{B}^{(2)}) &\leq H(\mathscr{B}^{(1)})+H(\mathscr{B}^{(2)})+\sum_i p_i \log_2 c \\ &\leq 2\log_2 d+\log_2 c, \end{aligned} \tag{2.79}$$

gdzie w ostatniej linii wykorzystaliśmy fakt, że

$$H(\mathscr{B}) \leq \log_2 d. \tag{2.80}$$

2.3. Kontekstualność

Rozważmy raz jeszcze zasadę nieoznaczoności w postaci podanej przez Robertsona (2.56). Załóżmy, że chcemy dokonać pomiaru dwóch niekomutujących ze sobą obserwabli na pewnym stanie kwantowym. Zasada nieoznaczoności mówi, że w najlepszym wypadku nie można wygenerować zespołu statystycznego opisanego pewnym stanem kwantowym, który z prawdopodobieństwem równym 1 (a więc z odchyleniem standardowym równym 0) zwracałby zawsze te same wyniki pomiarów dla obu obserwabli jednocześnie. O ile takie rozumowanie można jeszcze przeprowadzić rozpatrując stan kwantowy w kategoriach zespołu statystycznego, to nie można jeszcze na tej podstawie wnioskować, czy przykładowo pojedynczy obiekt może być opisany w kategoriach *zmiennych ukrytych*, które determinują wyniki pomiarów, którym jest poddany. Zamiast jednak rozpatrywać dwie niewspółmierzalne obserwable, zastanówmy się do jakich wniosków można dojść rozpatrując większą ilość obserwabli, z których przynajmniej niektóre są współmierzalne.

Rozważmy pewien dowolny stan kwantowy, na którym możemy wykonywać pomiary obserwabli ze zbioru $V = \{A_1, A_2, A_3, ...\}$, przy czym przez c_k będziemy oznaczać podzbiory obserwabli współmierzalnych, które będziemy dalej nazywać *kontekstami*. **Załóżmy**, że każdej obserwabli A_i w sposób jednoznaczny (niezależny od wyboru kontekstu) możemy przypisać pewną określoną wartość wyniku pomiaru a_i tak, aby przy każdorazowym pomiarze obserwabli A_i z określonego kontekstu c_k wynikiem pomiaru była ta konkretna wartość (w przypadku mechaniki kwantowej a_i może być jedną ze zbioru wartości własnych obserwabli A_i). Rozpatrzmy teraz pewien dowolny związek między obserwablami

$$f_k(A_1, A_2, ..., A_n) = 0, \tag{2.81}$$

gdzie obserwable $A_i \in c_k$ ($i \in \{1, ..., n\}$). Zauważmy, że jeśli wszystkie powyższe obserwable są współmierzalne (komutują ze sobą), to również wartości własne danych

obserwabli będą spełniały powyższy związek [7,8], tzn.

$$f_k(a_1, a_2, ..., a_n) = 0. \tag{2.82}$$

Rozważanymi obserwablami w szczególnym przypadku mogą być również operatory rzutowe rzędu 1 Π_i, którym w sposób jednoznaczny przypisujemy wartość π_i ($\pi_i \in \{1, 0\}$). Tym samym, jeśli przez d oznaczymy wymiar układu, to dla dowolnego kontekstu zawierającego d operatorów rzutowych musimy mieć następujący związek między obserwablami

$$\sum_{i=1}^{d} \Pi_i = id_{d\times d}, \tag{2.83}$$

a tym samym, zgodnie z przypisaniem (2.82) spełniony będzie analogiczny związek dla ich wartości własnych

$$\sum_{i=1}^{d} \pi_i = 1. \tag{2.84}$$

Okazuje się jednak, zgodnie z rezultatem Kochena i Speckera [72] oraz Bella [73], że dla $d \geq 3$ istnieją takie zbiory obserwabli Π_i, że nie istnieje takie przypisanie $\Pi_i \mapsto \pi_i$, dla którego spełnione będą relacje (2.84) (zbiory obserwabli Π_i mające tę własność nazywamy *zbiorami KS*). Tym samym, nie istnieje jednoznaczne przypisanie wyników wszystkim obserwablom, które mogłyby być w zgodzie ze statystycznymi przewidywaniami mechaniki kwantowej, bądź też każde takie przypisanie musiałoby koniecznie zależeć od wyboru kontekstu, w ramach którego wykonujemy pomiary, co nazywamy *kontekstualnością*. Oryginalny dowód Kochena i Speckera wymagał użycia [72] 117 obserwabli Π_i w przestrzeni $d = 3$, z czasem jednak znaleziono znacznie mniejsze zbiory KS [74–78]. Dodajmy jednak, że kontekstualność może się ujawniać również poprzez łamanie pewnych nierówności [9, 79–83]. W tym kontekście warto również wspomnieć o entropowych testach kontekstualności [84–88].

Poniżej przedstawimy przykłady układów kontekstualnych, które można zrealizować na gruncie mechaniki kwantowej.

Przykład 1, „kwadrat Peresa-Mermina".

Rozważmy dowolny stan układu dwóch spinów $\frac{1}{2}$ ($d = 4$), oraz następujący zbiór

$$\begin{array}{ccc} A_1 & A_2 & A_3 \\ A_4 & A_5 & A_6 \\ A_7 & A_8 & A_9 \end{array}$$

Rys. 2.1: 6 podzbiorów współmierzalnych obserwabli A_i (6 kontekstów) tworzą 3 rzędy oraz 3 kolumny.

obserwabli [6–8]:

$$\begin{aligned} A_1 &= \hat{\sigma}_x \otimes id_{2\times 2}, \\ A_2 &= id_{2\times 2} \otimes \hat{\sigma}_x, \\ A_3 &= \hat{\sigma}_x \otimes \hat{\sigma}_x, \\ A_4 &= id_{2\times 2} \otimes \hat{\sigma}_y, \\ A_5 &= \hat{\sigma}_y \otimes id_{2\times 2}, \\ A_6 &= \hat{\sigma}_y \otimes \hat{\sigma}_y, \\ A_7 &= \hat{\sigma}_x \otimes \hat{\sigma}_y, \\ A_8 &= \hat{\sigma}_y \otimes \hat{\sigma}_x, \\ A_9 &= \hat{\sigma}_z \otimes \hat{\sigma}_z, \end{aligned} \tag{2.85}$$

z których każda ma dwie wartości własne $+1$ oraz -1. Jeśli teraz rozmieścimy powyższe obserwable zgodnie ze schematem przedstawionym na Rys. 2.1, to 6 podzbiorów współmierzalnych obserwabli (6 kontekstów) tworzą 3 rzędy oraz 3 kolumny. Zauważmy, że dla wszystkich kontekstów mamy

$$A_i A_j A_k = id_{4\times 4}, \tag{2.86}$$

z wyjątkiem

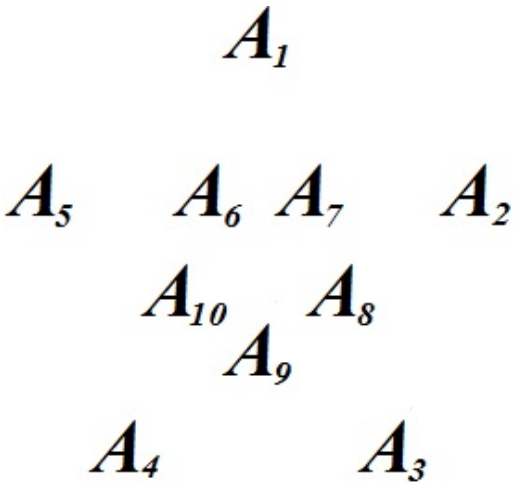
$$A_3 A_6 A_9 = -id_{4\times 4}. \tag{2.87}$$

Okazuje się, że nie jest możliwe jednoznaczne przypisanie każdej obserwabli A_i ustalonego wyniku a_i, w taki sposób, aby spełnione były relacje:

$$a_i a_j a_k = 1, \tag{2.88}$$

dla wszystkich kontekstów, z wyjątkiem

$$a_3 a_6 a_9 = -1. \tag{2.89}$$



Rys. 2.2: 5 podzbiorów współmierzalnych obserwabli (5 kontekstów) tworzonych jest przez 5 grup współliniowych obserwabli A_i.

Rzeczywiście, zgodnie z powyższymi relacjami, iloczyn wszystkich 9 wartości a_i wynosi 1 jeśli mnożylibyśmy wartości otrzymane dla trzech rzędów, bądź też -1 jeśli mnożylibyśmy wartości otrzymane dla trzech kolumn. Tym samym dowolne przyporządkowanie wartości $a_i = \pm 1$ obserwablom A_i zachowujące reguły (2.88) oraz (2.89) musiałoby być kontekstualne.

Przykład 2, „gwiazda Mermina".

Rozważmy dowolny stan układu trzech spinów $\frac{1}{2}$ ($d = 8$), oraz następujący zbiór obserwabli [7, 8]:

$$\begin{aligned}
A_1 &= \hat{\sigma}_y \otimes id_{2\times 2} \otimes id_{2\times 2},\\
A_2 &= \hat{\sigma}_x \otimes \hat{\sigma}_y \otimes \hat{\sigma}_y,\\
A_3 &= id_{2\times 2} \otimes \hat{\sigma}_x \otimes id_{2\times 2},\\
A_4 &= id_{2\times 2} \otimes \hat{\sigma}_y \otimes id_{2\times 2},\\
A_5 &= \hat{\sigma}_x \otimes \hat{\sigma}_x \otimes \hat{\sigma}_x,\\
A_6 &= \hat{\sigma}_y \otimes \hat{\sigma}_y \otimes \hat{\sigma}_x,\\
A_7 &= \hat{\sigma}_y \otimes \hat{\sigma}_x \otimes \hat{\sigma}_y,\\
A_8 &= id_{2\times 2} \otimes id_{2\times 2} \otimes \hat{\sigma}_y,\\
A_9 &= \hat{\sigma}_x \otimes id_{2\times 2} \otimes id_{2\times 2},\\
A_{10} &= id_{2\times 2} \otimes id_{2\times 2} \otimes \hat{\sigma}_x,
\end{aligned} \tag{2.90}$$

z których każda ma dwie wartości własne $+1$ oraz -1. Jeśli teraz rozmieścimy powyższe obserwable zgodnie ze schematem przedstawionym na Rys. 2.2, to 5 podzbiorów współmierzalnych obserwabli (5 kontekstów) tworzonych jest przez 5 grup współli-

niowych obserwabli. Zauważmy, że dla wszystkich kontekstów mamy

$$A_i A_j A_k A_l = i d_{8\times 8}, \tag{2.91}$$

z wyjątkiem

$$A_2 A_5 A_6 A_8 = -i d_{8\times 8}. \tag{2.92}$$

Okazuje się, że nie jest możliwe jednoznaczne przypisanie każdej obserwabli A_i ustalonego wyniku a_i, w taki sposób, aby spełnione były relacje:

$$a_i a_j a_k a_l = 1, \tag{2.93}$$

dla wszystkich kontekstów, z wyjątkiem

$$a_2 a_5 a_6 a_8 = -1. \tag{2.94}$$

Rzeczywiście, zgodnie z powyższymi relacjami, mnożąc wartości otrzymane dla wszystkich 5 krawędzi otrzymalibyśmy wartość -1, co jednak musiałoby odpowiadać pomnożeniu kwadratów wszystkich 10 wyników przez siebie (każda obserwabla występuje w dwóch różnych kontekstach) dając tym samym wynik 1. Wobec tego dowolne przyporządkowanie wartości $a_i = \pm 1$ obserwablom A_i zachowujące reguły (2.93) oraz (2.94) musiałoby być kontekstualne.

Warto zauważyć, że w przeciwieństwie do kwadratu Peresa-Mermina, można znaleźć przypisanie wartości a_i obserwablom A_i, które spełniają przeciwne reguły w stosunku do reguł (2.93) oraz (2.94):

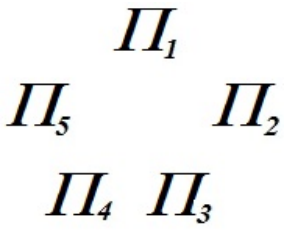
$$a_i a_j a_k a_l = -1, \tag{2.95}$$

dla wszystkich kontekstów, z wyjątkiem

$$a_2 a_5 a_6 a_8 = 1. \tag{2.96}$$

Reguły te będą zachowane, jeśli tylko $a_3 = a_4 = -1$, z kolei pozostałe $a_i = 1$.

W powyższych przykładach mogliśmy zauważyć, że wiele z relacji współmierzalności obserwabli A_i wynikała z prostego faktu, że dwa operatory, np. $\hat{\sigma}_x$ oraz $\hat{\sigma}_y$ działały na różnych podprzestrzeniach przestrzeni Hilberta ($\hat{\sigma}_x \otimes id_{2\times 2}$ oraz $id_{2\times 2} \otimes \hat{\sigma}_y$). Z tym samym mieliśmy do czynienia rozważając dwie pary niewspółmierzalnych obserwabli A_0, A_1 oraz B_0, B_1 w nierównościach CHSH. W tym kontekście można powiedzieć, że *lokalne zmienne ukryte* uzasadniające nierówności Bella można uważać za *niekontekstualne zmienne ukryte* przy uwzględnieniu kontekstów wynikających z przestrzennego odseparowania dwóch podukładów, na którym wykonywane są pomiary. Z



Rys. 2.3: 5 podzbiorów ortogonalnych operatorów rzutowych (5 kontekstów) tworzonych jest przez 5 par $\Pi_i, \Pi_{i+1\text{mod}5}$.

tego względu wszystkie nierówności Bella można uznać za szczególnego rodzaju nierówności kontekstualności, których łamanie świadczy o kontekstualnym charakterze określonego układu. O ile jednak warunkiem koniecznym na złamanie nierówności Bella (w odniesieniu do mechaniki kwantowej) jest wykonanie pomiarów na stanie splątanym, to kontekstualność układu można w pewnych przypadkach wykazać dla dowolnych stanów kwantowych (tzw. kontekstualność niezależna od stanu), czego przykładem jest kwadrat Peresa-Mermina, lub gwiazda Mermina. Jak pokażemy na poniższym przykładzie, w odróżnieniu od nierówności Bella, można znaleźć nierówności kontekstualności również dla układów niepodzielnych, czego przykładem jest pojedynczy układ o wymiarze przestrzeni Hilberta $d = 3$, np. cząstka o spinie 1. Należy przy tym powiedzieć, że jest to najprostszy układ kontekstualny, gdyż nie można wykazać kontekstualności dla układu o wymiarze przestrzeni Hilberta $d = 2$ [89].

Przykład 3, układ KCBS.

Rozważmy 5 jednorzędowych operatorów rzutowych Π_i, których pary $\Pi_i, \Pi_{i+1\text{mod}5}$ są ortogonalne względem siebie (a tym samym współmierzalne) [9]. Zauważmy, że żadne z tak wybranych operatorów rzutowych nie tworzą pełnej bazy. Z kolei ortogonalność par $\Pi_i, \Pi_{i+1\text{mod}5}$ narzuca warunek wykluczania:

$$\Pi_i \Pi_{i+1\text{mod}5} = 0. \tag{2.97}$$

Jeśli teraz rozmieścimy je w formie pięciokąta przedstawionego na Rys. 2.3, to łatwo zauważyć, że chcąc operatorom Π_i przypisać w sposób jednoznaczny wartości $\pi_i = 1, 0$ spełniające warunek

$$\pi_i \pi_{i+1\text{mod}5} = 0, \tag{2.98}$$

suma wszystkich wartości π_i nie może być większa niż 2, a zatem musi być spełniona następująca nierówność kontekstualności:

$$\sum_{i=1}^{5} \langle \Pi_i \rangle \leq 2. \tag{2.99}$$

Dla powyższego układu można sformułować jednocześnie nierówność kontekstualności w formie ograniczenia na korelacje między wynikami współmierzalnych obserwabli. W tym celu w miejsce operatorów rzutowych wstawmy obserwable zdefiniowane poprzez $A_i = 2\Pi_i - id_{3\times 3}$ o wartościach własnych $a_i = \pm 1$. Łatwo pokazać, że przy wykorzystaniu warunku wykluczania (2.97) oraz nierówności (2.99) musi być spełniona nierówność:

$$\sum_{i=1}^{5} \langle A_i A_{i+1 \bmod 5} \rangle + 3 \geq 0. \tag{2.100}$$

Rozważmy teraz stan kwantowy $|\Psi\rangle = (0,0,1)$, z operatorami rzutowymi $\Pi_i = |v_i\rangle\langle v_i|$ takimi, że

$$\begin{aligned}
|v_1\rangle &= \mathcal{N}_1 \left(1, 0, \sqrt{\cos(\pi/5)}\right), \\
|v_2\rangle &= \mathcal{N}_2 \left(\cos(4\pi/5), \sin(4\pi/5), \sqrt{\cos(\pi/5)}\right), \\
|v_3\rangle &= \mathcal{N}_3 \left(\cos(2\pi/5), -\sin(2\pi/5), \sqrt{\cos(\pi/5)}\right), \\
|v_4\rangle &= \mathcal{N}_4 \left(\cos(2\pi/5), \sin(2\pi/5), \sqrt{\cos(\pi/5)}\right), \\
|v_5\rangle &= \mathcal{N}_5 \left(\cos(4\pi/5), -\sin(4\pi/5), \sqrt{\cos(\pi/5)}\right),
\end{aligned} \tag{2.101}$$

gdzie $\mathcal{N}_i$ są odpowiednimi czynnikami normalizacyjnymi. Dla tak wybranych operatorów, mamy:

$$\Pi_i \Pi_{i+1 \bmod 5} = 0, \tag{2.102}$$

$$\langle \Psi | \Pi_i | \Psi \rangle = \frac{1}{\sqrt{5}}, \tag{2.103}$$

$$\langle \Psi | A_i A_{i+1 \bmod 5} | \Psi \rangle = 1 - \frac{4}{\sqrt{5}}, \tag{2.104}$$

a tym samym

$$\sum_{i=1}^{5} \langle \Psi | \Pi_i | \Psi \rangle = \sqrt{5} \approx 2,2361, \tag{2.105}$$

$$\sum_{i=1}^{5} \langle \Psi | A_i A_{i+1 \bmod 5} | \Psi \rangle + 3 = 8 - 4\sqrt{5} \approx -0,9443, \tag{2.106}$$

co łamie nierówności (2.99) oraz (2.100). Widzimy zatem, że jakiekolwiek niekontekstualne przyporządkowanie wartości $\pi_i = 1,0$ operatorom rzutowym Π_i zachowujące regułę (2.98) prowadzi do wniosków sprzecznych z wnioskami płynącymi z mechaniki kwantowej.

Rozdział 3

Aktywacja nielokalnych korelacji kwantowych

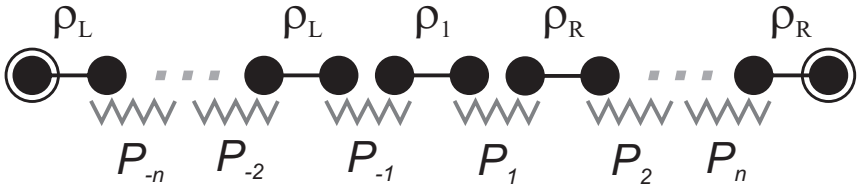
3.1. Wstęp

W bieżącym rozdziale omówimy wielokrotną wymianę splątania kwantowego dokonaną w łańcuchu stanów dwuqubitowych [10–13]. Pokażemy, że w przypadku otrzymywania pewnych określonych wyników pomiaru Bella po dostatecznie dużej liczbie wymian splątania, z początkowych stanów niełamiących nierówności CHSH uzyskać można stan końcowy łamiący tę nierówność, co dalej będziemy określać mianem aktywacji łamania nierówności CHSH.

3.2. Łańcuch stanów

Rozważymy przypadek, kiedy początkowe stany kwantowe zostały rozdystrybuowane pomiędzy pary użytkowników tworzących łańcuch (por Rys. 3.1). Każda para użytkowników w łańcuchu oznaczona przez P_i, P_{i+1} $(1 \leqslant i \leqslant N-1)$ współdzieli po jednej kopii stanu ρ_R, natomiast każda para użytkowników P_i, P_{i+1} $(-N \leqslant i \leqslant -2)$ współdzieli po jednej kopii stanu ρ_L, z kolei „centralna para" użytkowników P_{-1}, P_1 współdzieli stan ρ_1. Poszczególne stany kwantowe mają następujące postaci:

$$\rho_R = p|\Psi_R\rangle\langle\Psi_R| + (1-p)|00\rangle\langle 00|, \tag{3.1}$$



Rys. 3.1: Łańcuch stanów wykorzystywanych w niniejszej procedurze. Każda para użytkowników P_{-i-1}, P_{-i} współdzieli stan ρ_L, podczas gdy pary użytkowników P_i, P_{i+1} $(1 \leqslant i \leqslant N-1)$ współdzielą stan ρ_R. Centralny stan ρ_1 jest współdzielony przez użytkowników P_{-1} oraz P_1. Każdy użytkownik dokonuje pomiaru w bazie Bella na qubitach pochodzących z dwóch sąsiednich stanów. Źródło: [13].

gdzie

$$|\Psi_R\rangle = \sin\alpha|01\rangle + \cos\alpha|10\rangle, \tag{3.2}$$

podobnie

$$\rho_L = p|\Psi_L\rangle\langle\Psi_L| + (1-p)|00\rangle\langle 00|, \tag{3.3}$$

gdzie

$$|\Psi_L\rangle = \cos\alpha|01\rangle + \sin\alpha|10\rangle, \tag{3.4}$$

podczas gdy

$$\rho_1 = p_1|\Psi^+\rangle\langle\Psi^+| + (1-p_1)|00\rangle\langle 00|, \tag{3.5}$$

przy ogólnej notacji

$$|\Psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle). \tag{3.6}$$

Ponieważ badamy możliwość aktywacji łamania nierówności CHSH, jesteśmy zainteresowani wyłącznie takimi stanami początkowymi, dla których nie jest możliwe złamanie wspomnianej nierówności przez dwoje sąsiednich użytkowników dokonujących pomiarów na qubitach ze współdzielonego stanu. Tym samym musimy znaleźć zakres parametrów p, α oraz p_1, dla których wszystkie stany ρ_R, ρ_L oraz ρ_1 będą stanami lokalnymi. W tym celu posłużymy się następującym kryterium pozwalającym jednoznacznie stwierdzić, które stany mają tę własność: dwuqubitowy stan ρ nie łamie nierówności CHSH wtedy i tylko wtedy, gdy [24]:

$$\sqrt{\lambda_i + \lambda_j} \leq 1, \tag{3.7}$$

gdzie λ_i oraz λ_j są odpowiednio największą oraz drugą co do wielkości wartością własną macierzy R^TR. Macierz R jest zdefiniowana poprzez następujące elementy macierzowe:

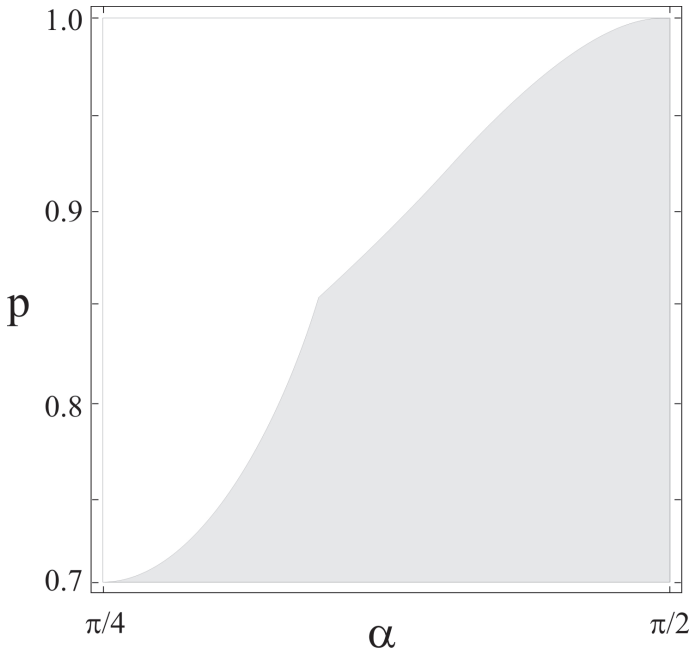
$$R_{ij} = \mathrm{Tr}[(\hat{\sigma}_i \otimes \hat{\sigma}_j)\rho], \tag{3.8}$$

gdzie $\hat{\sigma}_i$ to macierze Pauliego. Dla stanów (3.1) i (3.3) macierz R^TR przyjmuje postać:

$$R^TR = \begin{pmatrix} p^2\sin^2 2\alpha & 0 & 0 \\ 0 & p^2\sin^2 2\alpha & 0 \\ 0 & 0 & (1-2p)^2 \end{pmatrix}, \tag{3.9}$$

z kolei macierz R^TR dla stanu (3.5) otrzymamy z powyższej kładąc $\alpha = \frac{\pi}{4}$.

Zastosowanie kryterium (3.7) do stanów (3.1), (3.3) i (3.5) daje następujące ogra-



Rys. 3.2: Parametry stanów ρ_R i ρ_L nie pozwalające na złamanie nierówności CHSH (zacieniony obszar). Źródło: [13].

niczenia na początkowe parametry, dla których stany ρ_R i ρ_L nie łamią nierówności CHSH:

$$\max_{p,\alpha}\left[2p^2\sin 2\alpha,\ 1-4p+\frac{p^2}{2}(9-\cos 4\alpha)\right]\leq 1, \tag{3.10}$$

natomiast stan ρ_1 nie łamie nierówności CHSH dla:

$$p_1 \leq \frac{1}{\sqrt{2}} \approx 0.707. \tag{3.11}$$

Otrzymany zakres parametrów p oraz α (przy arbitralnie wybranym ograniczeniu $\alpha \in [\frac{\pi}{4}, \frac{\pi}{2}]$), dla których nie można złamać nierówności CHSH w przypadku stanów ρ_R oraz ρ_L przedstawia Rys. 3.2. W tym miejscu warto wspomnieć, iż w przypadku $\alpha = \frac{\pi}{2}$ stany ρ_R oraz ρ_L są separowalne, natomiast w przypadku $\alpha = \frac{\pi}{4}$ tworzą one stan mieszany powstały w wyniku mieszania stanu maksymalnie splątanego i stanu produktowego.

3.3. Wymiana splątania w łańcuchu stanów

Procedura wymiany splątania w łańcuchu stanów przebiega w następujący sposób. W pierwszej kolejności wymiany splątania dokonują użytkownicy centralni, P_{-1} oraz P_1. W tym celu mierzą oni swoje qubity w bazie Bella ($\{|\Psi^+\rangle, |\Psi^-\rangle, |\Phi^+\rangle, |\Phi^-\rangle\}$, gdzie $|\Phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)$): użytkownik P_1 mierzy qubity pochodzące ze stanów ρ_1 oraz

ρ_R, natomiast użytkownik P_{-1} mierzy qubity pochodzące ze stanów ρ_1 oraz ρ_L. Po takim pomiarze dokonanym na stanie $\rho_L \otimes \rho_1 \otimes \rho_R$ i zakomunikowaniu wyników $|\Psi^+\rangle$ lub $|\Psi^-\rangle$ użytkownicy P_{-2} oraz P_2 współdzielą pewien stan ρ_2. Następnie użytkownicy P_{-2} oraz P_2 dokonują wymiany splątania mierząc qubity pochodzące ze stanów ρ_2 i ρ_L oraz odpowiednio ρ_2 i ρ_R. W konsekwencji tego ze stanu $\rho_L \otimes \rho_2 \otimes \rho_R$ otrzymują stan ρ_3, który będzie współdzielony przez kolejnych użytkowników P_{-3} oraz P_3. Procedura przebiega analogicznie do momentu, kiedy dwoje ostatnich użytkowników – P_{-n} oraz P_n – współdzielą stan ρ_n.

Postaci poszczególnych stanów $\rho_2, \ldots, \rho_n$ zależą od wyników pomiarów uzyskanych podczas każdego pomiaru. Załóżmy najpierw, że wyniki wszystkich pomiarów wynoszą $|\Psi^+\rangle$. W tym przypadku, po pomiarze dokonanym przez parę użytkowników P_{-1} i P_1, użytkownicy P_{-2} oraz P_2 będą współdzielić stan:

$$\begin{aligned}\rho_2 &= \frac{\mathrm{Tr}_{P_{-1},P_1}[(id_{2\times 2} \otimes |\Psi^+\rangle\langle\Psi^+| \otimes |\Psi^+\rangle\langle\Psi^+| \otimes id_{2\times 2})\rho_L \otimes \rho_1 \otimes \rho_R]}{\mathrm{Tr}[(id_{2\times 2} \otimes |\Psi^+\rangle\langle\Psi^+| \otimes |\Psi^+\rangle\langle\Psi^+| \otimes id_{2\times 2})\rho_L \otimes \rho_1 \otimes \rho_R]} \\ &= p_2|\Psi^+\rangle\langle\Psi^+| + (1-p_2)|00\rangle\langle 00|, \end{aligned} \tag{3.12}$$

gdzie Tr_{P_{-1},P_1} oznacza śladowanie po qubitach należących do użytkowników P_{-1} i P_1, natomiast

$$p_2 = p_1 p \left(p\,\mathrm{ctg}^2\alpha + p_1 \frac{1-p-p\cos 2\alpha}{\sin^2\alpha} \right)^{-1}. \tag{3.13}$$

Łatwo zauważyć, że stan ρ_2 jest tej samej postaci co początkowy stan ρ_1, a zmianie ulega jedynie stosunek wagi stanu maksymalnie splątanego $|\Psi^+\rangle$ do wagi stanu produktowego $|00\rangle$. Kolejne operacje wymiany splątania działają na identycznej zasadzie, a zatem po $k-1$ krokach tej procedury, użytkownicy P_{-k} i P_k będą współdzielić stan

$$\rho_k = p_k|\Psi^+\rangle\langle\Psi^+| + (1-p_k)|00\rangle\langle 00|, \tag{3.14}$$

z pewnym współczynnikiem p_k, który wyznaczamy poniżej. Wiemy już, że w każdym kolejnym kroku procedury (po dwóch kolejnych pomiarach) zmianie ulega waga stanu maksymalnie splątanego $|\Psi^+\rangle$:

$$p_{k+1} = p_k p \left(p\,\mathrm{ctg}^2\alpha + p_k \frac{1-p-p\cos 2\alpha}{\sin^2\alpha} \right)^{-1}. \tag{3.15}$$

Powyższe równanie rekurencyjne ma rozwiązanie dane przez:

$$p_k = \left[\frac{\mathrm{ctg}^{2(k-1)}\alpha}{p_1} + (1-\mathrm{ctg}^{2(k-1)}\alpha)\left(\frac{p-1}{p\cos 2\alpha} + 1 \right) \right]^{-1}. \tag{3.16}$$

W omówionej procedurze przyjęliśmy, że wszyscy użytkownicy otrzymują w wyniku pomiaru Bella stan $|\Psi^+\rangle$. Naturalne jest zatem pytanie, jaką postać będzie miał stan końcowy w przypadku, gdy różni użytkownicy otrzymają w wyniku pomiaru

również pozostałe stany Bella: $|\Psi^-\rangle$, $|\Phi^+\rangle$ oraz $|\Phi^-\rangle$. Ogólny przypadek zostanie omówiony szerzej w podrozdziale 3.5.1, tutaj natomiast wspomnimy o sytuacji, gdy jedynymi wynikami pomiarów Bella są stany $|\Psi^\pm\rangle$. Jeśli w pierwszym kroku wymiany splątania użytkownicy P_{-1} oraz P_1 otrzymają różne wyniki pomiarów ($|\Psi^+\rangle$ i $|\Psi^-\rangle$), to stan współdzielony przez kolejnych użytkowników będzie postaci

$$\rho_2^- = p_2|\Psi^-\rangle\langle\Psi^-| + (1-p_2)|00\rangle\langle 00|. \tag{3.17}$$

Analogicznie, jeśli po $k-1$ krokach procedury wymiany splątania nieparzysta ilość użytkowników otrzyma wynik $|\Psi^-\rangle$, to stan współdzielony przez użytkowników P_{-k} oraz P_k będzie postaci

$$\rho_k^- = p_k|\Psi^-\rangle\langle\Psi^-| + (1-p_k)|00\rangle\langle 00|. \tag{3.18}$$

W tym przypadku dany stan można bezpośrednio sprowadzić do postaci (3.14) poprzez poprawę czynnika fazowego, wykonując w tym celu lokalnie (tylko przez jednego użytkownika) operację $\hat{\sigma}_z$. Z tego też względu, procedurę wymiany splątania będziemy dalej nazywać udaną, jeśli każdy użytkownik otrzyma w wyniku pomiaru Bella jeden ze stanów $|\Psi^+\rangle$ bądź $|\Psi^-\rangle$. Na zakończenie zwróćmy uwagę, iż prawdopodobieństwo dokonania m kroków wymiany splątania z wynikami pomiarów tylko $|\Psi^\pm\rangle$ maleje eksponencjalnie z m.

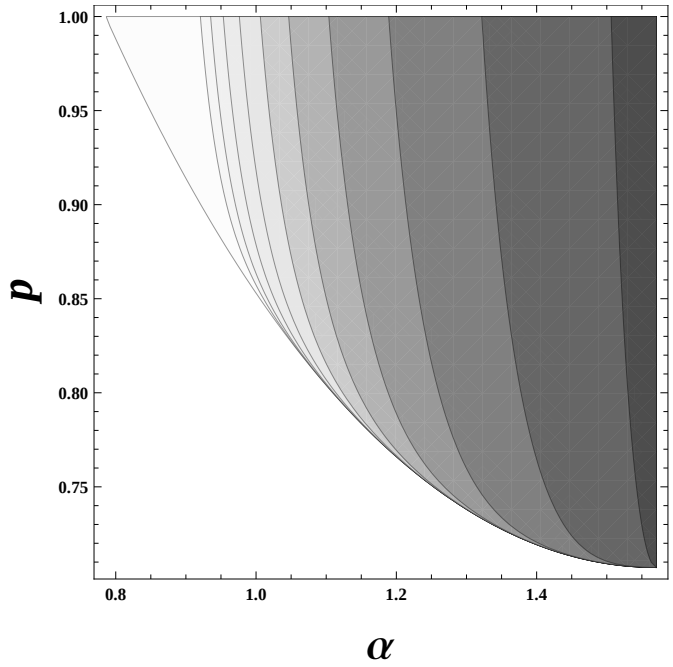
3.4. Aktywacja nielokalności w łańcuchu stanów

W przypadku klasy stanów postaci (3.14) warunkiem koniecznym i wystarczającym na łamanie nierówności CHSH jest

$$p_k > \frac{1}{\sqrt{2}}, \tag{3.19}$$

natomiast z relacji (3.16) wynika, że dla pewnych wartości p i α czynnik p_k wzrasta wraz z liczbą kolejnych kroków k (pamiętajmy, że w każdym kroku wykonywane są 2 pomiary, a zatem ściśle rzecz ujmując k kroków odpowiada $2n$ liczbie wymian splątania). Z tego względu wymiana splątania pozwala na otrzymanie z początkowych stanów ρ_L, ρ_1 oraz ρ_R (dla pewnych parametrów p i α), które nie pozwalają na łamanie nierówności CHSH, stanu ρ_k, który łamie tę nierówność po wykonaniu dostatecznej liczby wymian splątania. W tym sensie nielokalność może zostać aktywowana.

Rys. 3.3 przedstawia zakres parametrów stanów ρ_L i ρ_R, dla których można aktywować nielokalność po dokonaniu odpowiedniej liczby wymian splątania (stan początkowy ρ_1 użyty w tej konfiguracji ma parametr $p_1 = 0,01$). Przy porównaniu z Rys. 3.2 widać wyraźnie, że dla pewnych stanów, które pierwotnie nie pozwalały na złamanie nierówności CHSH, pewna liczba wymian splątania pozwala na uzyskanie stanu, dla którego można złamać tę nierówność. Podobnie, istnieją stany, dla któ-



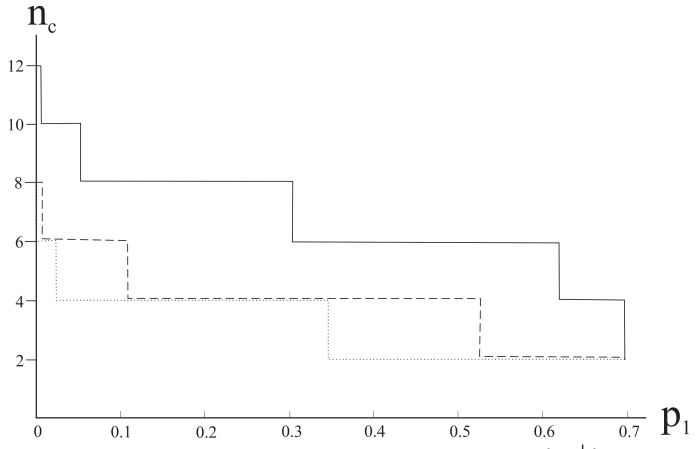
Rys. 3.3: Stany ρ_L oraz ρ_R, które pozwalają na złamanie nierówności CHSH po dokonaniu n wymian splątania, w przypadku gdy wszystkie pomiary Bella dają wynik $|\Psi^{\pm}\rangle$ dla $n = 2, 4, ..., 20$ i $n \to \infty$ (odpowiednie zacienione obszary w kierunku malejących wartości α) przy $p_1 = 0,01$.

rych nie można aktywować nielokalności po dokonaniu pewnej liczby wymian splątania, jednak nielokalność zostaje aktywowana po dokonaniu kolejnego kroku wymiany splątania.

W dalszej kolejności pokażemy, iż liczba wykonywanych pomiarów Bella z wynikami $|\Psi^{\pm}\rangle$ ma kluczowe znaczenie jeśli chodzi o aktywację nielokalności przy użyciu początkowych stanów ρ_1, ρ_L oraz ρ_R. W tym celu określamy krytyczną liczbę wymian splątania, n_c, niezbędnych do aktywacji łamania nierówności CHSH. Rys. 3.4 przedstawia wartości n_c dla kilku klas stanów początkowych ρ_1, ρ_L i ρ_R. Warto przy tym zauważyć, że dla pewnych stanów ρ_L i ρ_R możliwa jest aktywacja nielokalności dla wszystkich stanów postaci (3.14) nawet przy dowolnej wadze stanu $|00\rangle\langle 00|$ z przedziału $[0, 1)$.

3.5. Przypadki nieprowadzące do aktywacji nielokalności

Jak dotąd rozpatrywaliśmy przypadki aktywacji nielokalnych korelacji w łańcuchu stanów kwantowych (dla odpowiednich parametrów stanów początkowych) z wykorzystaniem wymiany splątania, gdzie kluczową rolę odgrywały trzy warunki. Po pierwsze, każdy z użytkowników otrzymywał jako wynik pomiaru Bella stan $|\Psi^{+}\rangle$



Rys. 3.4: Krytyczna liczba pomiarów Bella n_c z wynikami $|\Psi^{\pm}\rangle$ konieczna do aktywacji nielokalności dla wybranych stanów początkowych ρ_L i ρ_R (p=0,75) z $\alpha = \frac{20}{25}\pi$ (linia ciągła), $\alpha = \frac{21}{25}\pi$ (linia przerywana) oraz $\alpha = \frac{22}{25}\pi$ (linia kropkowana). Źródło: [13].

bądź $|\Psi^{-}\rangle$. Po drugie, aktywacja nielokalności wymagała dokonania przynajmniej n_c pomiarów, w których otrzymano opisane wyniki pomiarów. Po trzecie, stan końcowy otrzymywany był w przypadku, kiedy wymiana splątania dokonywana była przez taką samą liczbę użytkowników po prawej jak i po lewej stronie łańcucha. W takiej sytuacji powstaje naturalne pytanie, czy można dokonać aktywacji nielokalności w przypadku, gdy liczba pomiarów Bella jest mniejsza niż n_c przy założeniu, że użytkownicy mogą otrzymywać inne wyniki swoich pomiarów. Ponadto zbadamy, czy aktywację nielokalności można otrzymać w przypadku, kiedy liczba użytkowników wykonujących wymianę splątania jest różna po obu stronach łańcucha.

W ogólnym przypadku trudne jest znalezienie prostej formuły opisującej stan końcowy po pomiarach dokonanych przez dowolny ciąg użytkowników, które dodatkowo mogły dać dowolne wyniki $|\Psi^{\pm}\rangle$ bądź $|\Phi^{\pm}\rangle$. Częściowej odpowiedzi na tak postawiony problem dostarczyć może jedynie analiza numeryczna, która dokonana została w ograniczonej liczbie przypadków. Mianowicie, sprawdzono wszystkie możliwe wyniki pomiarów Bella, które otrzymane zostały przez użytkowników: (i) P_{-1}, (ii) P_1, (iii) P_{-2}, (iv) P_2, (v) P_{-1} i P_1, (vi) P_1 i P_2, etc., uwzględniając wszystkie możliwe ciągi użytkowników aż do przypadku $P_{-3}, ..., P_3$ (ogółem 236 możliwości). Okazuje się między innymi, że jeśli stany początkowe wykorzystywane w tej metodzie nie łamią nierówności CHSH, to aktywacja nielokalnych korelacji nie jest możliwa w przypadku, gdy choćby jeden użytkownik otrzyma w wyniku swojego pomiaru rezultat $|\Phi^{\pm}\rangle$. Bezpośrednia numeryczna analiza problemu dla $i > 3$ staje się jednak nieefektywna zważywszy na fakt, że liczba możliwych konfiguracji wzrasta eksponencjalnie wraz ze zwiększaniem liczby użytkowników dokonujących pomiary. Okazuje się jednak, że przynajmniej dla pewnej ograniczonej klasy lokalnych stanów ρ_R i ρ_L możemy otrzymać ścisłe rezultaty pokazujące, iż jedynie przy wynikach pomiarów $|\Psi^{\pm}\rangle$ można dokonać aktywacji nielokalności. Ponadto, w dalszej kolejności pokażemy, że aktywa-

cji nielokalności nie można dokonać dysponując tylko jedną klasą stanów (ρ_R lub ρ_L) oraz w przypadku gdy wymiany splątania dokona różna liczba użytkowników po prawej oraz po lewej stronie łańcucha (licząc od centralnego stanu ρ_1).

3.5.1 Stany $|\Phi^{\pm}\rangle$ jako wyniki pomiarów Bella

Rozpatrzmy teraz tylko jedną (prawą) część łańcucha, w której wszyscy użytkownicy $P_1, ..., P_k$ $(k > 1)$ współdzielą stany ρ_R. Przypuśćmy dalej, że pewien użytkownik P_i $(1 < i < k)$ otrzymał jako wynik pomiaru $|\Phi^{\pm}\rangle$. Wobec tego, sąsiedni użytkownicy P_{i-1} oraz P_{i+1} będą współdzielić stan

$$\rho_{RR}^{\Phi^{\pm}} = \frac{\mathrm{Tr}_{P_i}[(id_{2\times 2} \otimes |\Phi^{\pm}\rangle\langle\Phi^{\pm}| \otimes id_{2\times 2})\rho_R \otimes \rho_R]}{\mathrm{Tr}[(id_{2\times 2} \otimes |\Phi^{\pm}\rangle\langle\Phi^{\pm}| \otimes id_{2\times 2})\rho_R \otimes \rho_R]}. \tag{3.20}$$

Korzystając z kryterium separowalności Peresa-Horodeckich [90, 91] dla stanów dwuqubitowych znajdziemy zakres parametrów p oraz α, dla których stan $\rho_{RR}^{\Phi^{\pm}}$ będzie stanem separowalnym. Zakres ten został przedstawiony na Rys. 3.5. Zauważmy teraz, że jeżeli w łańcuchu stanów pojawi się stan separowalny i pozostali użytkownicy wykonają protokół wymiany splątania, to stan końcowy również będzie separowalny, a zatem również lokalny.

Z podobną sytuacją możemy się spotkać, jeśli użytkownik P_1 dokona pomiaru na qubitach pochodzących ze stanów ρ_1 oraz ρ_R. Jeśli w wyniku jego pomiaru otrzyma stan $|\Phi^{\pm}\rangle$, to sąsiedni użytkownicy P_{-1} oraz P_2 współdzielić będą stan

$$\rho_{1R}^{\Phi^{\pm}} = \frac{\mathrm{Tr}_{P_i}[(id_{2\times 2} \otimes |\Phi^{\pm}\rangle\langle\Phi^{\pm}| \otimes id_{2\times 2})\rho_1 \otimes \rho_R]}{\mathrm{Tr}[(id_{2\times 2} \otimes |\Phi^{\pm}\rangle\langle\Phi^{\pm}| \otimes id_{2\times 2})\rho_1 \otimes \rho_R]}, \tag{3.21}$$

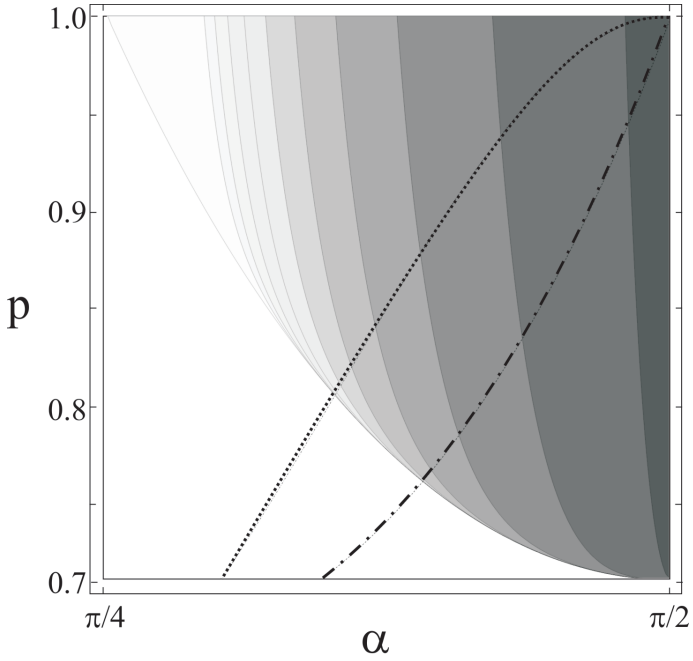
który również jest stanem separowalnym w pewnym zakresie parametrów p i α (patrz Rys. 3.5).

Ze względu na symetrię stanów ρ_R oraz ρ_L, powyższa analiza odnosi się również do drugiej (lewej) strony łańcucha stanów.

3.5.2 Użycie tylko jednej klasy stanów

Rozpatrzmy teraz przypadek, kiedy wszyscy użytkownicy otrzymają w wyniku swojego pomiaru Bella wynik $|\Psi^{\pm}\rangle$, jednak pomiary dokonywane będą tylko na jednej klasie stanów. W tym celu rozpatrzmy prawą stronę łańcucha oraz użytkowników $P_1, ..., P_{n+1}$ $(n \geqslant 2)$ współdzielących stany ρ_R. Po krótkiej analizie możemy się przekonać, że dokonanie $n-1$ wymian splątania powoduje modyfikację stanu $|\Psi_R\rangle$ w relacji:

$$|\Psi_R\rangle \longrightarrow |\Psi'_{R,n}\rangle, \tag{3.22}$$



Rys. 3.5: Zakres stanów, dla których wynik pomiaru $|\Phi^{\pm}\rangle$ prowadzi do powstania stanu separowalnego $\rho_{RR}^{\Phi^{\pm}}$ (odpowiednio $\rho_{1R}^{\Phi^{\pm}}$), w przypadku gdy pomiar dokonywany jest na stanie $\rho_R \otimes \rho_R$ ($\rho_1 \otimes \rho_R$), przedstawia obszar poniżej linii kropkowanej (kropkowano-przerywanej). W przypadku $\rho_{1R}^{\Phi^{\pm}}$ wykorzystano stan ρ_1 z $p_1 \leqslant \frac{1}{\sqrt{2}}$. Źródło: [13].

gdzie nieunormowany stan

$$|\Psi'_{R,n}\rangle = \sin^n \alpha |01\rangle + \cos^n \alpha |10\rangle. \tag{3.23}$$

Z tego też względu, dokonanie $n-1$ procedur wymiany splątania przez użytkowników $P_2, ..., P_n$ (przy uzyskaniu wyniku $|\Psi^{\pm}\rangle$ przez każdego z nich) użytkownicy P_1 i P_{n+1} będą współdzielić stan (po możliwej korekcji fazy):

$$\rho_{R,n} = p'_{R,n} |\Psi'_{R,n}\rangle\langle\Psi'_{R,n}| + (1 - p'_{R,n} \mathrm{Tr} |\Psi'_{R,n}\rangle\langle\Psi'_{R,n}|)|00\rangle\langle 00|, \tag{3.24}$$

z pewnym współczynnikiem $p'_{R,n}$, które wyznaczymy poniżej.

Załóżmy teraz, że wymiany splątania dokonałby jeszcze jeden użytkownik. Wiemy, że wymiana splątania wykonywana na stanach ρ_R nie wyprowadza poza klasę stanów (3.24), a modyfikacji ulega jedynie stan splątany $|\Psi'_{R,n}\rangle \longrightarrow |\Psi'_{R,n+1}\rangle$. Dokładniejsza analiza prowadzi do stwierdzenia, że po owej dodatkowej wymianie splątania, stan współdzielony przez odpowiednich użytkowników będzie postaci:

$$\begin{aligned} & p p'_{R,n} |\Psi'_{R,n+1}\rangle\langle\Psi'_{R,n+1}| \\ & \quad + \{p'_{R,n} \sin^{2n} \alpha (1-p) + [1 - p'_{R,n} (\sin^{2n} \alpha + \cos^{2n} \alpha)] p \cos^2 \alpha\} |00\rangle\langle 00|. \end{aligned} \tag{3.25}$$

Unormowanie powyższego stanu oraz zastąpienie czynnika przy $|\Psi'_{R,n+1}\rangle\langle\Psi'_{R,n+1}|$ przez

$p'_{R,n+1}$ prowadzi do równania rekurencyjnego, którego rozwiązanie daje czynnik $p'_{R,n+1}$:

$$p'_{R,n+1} = \left\{ \sin^{2(n+1)}\alpha + \cos^{2(n+1)}\alpha + \left(\frac{1}{p} - 1\right)\sin^{2n}\alpha + \left[\frac{1}{p'_{R,n}} - (\sin^{2n}\alpha + \cos^{2n}\alpha)\right]\right\}^{-1}. \tag{3.26}$$

Jeśli teraz unormujemy stan splątany (3.23) do postaci

$$|\Psi_{R,n}\rangle = \sin\alpha_n|01\rangle + \cos\alpha_n|10\rangle, \tag{3.27}$$

gdzie

$$\sin\alpha_n = \frac{\sin^n\alpha}{\sqrt{\sin^{2n}\alpha + \cos^{2n}\alpha}}, \tag{3.28}$$

to wykorzystując relację

$$p_{R,n} = p'_{R,n}(\sin^{2n}\alpha + \cos^{2n}\alpha), \tag{3.29}$$

stan (3.24) będziemy mogli zapisać jako

$$\rho_{R,n} = p_{R,n}|\Psi_{R,n}\rangle\langle\Psi_{R,n}| + (1 - p_{R,n})|00\rangle\langle 00|, \tag{3.30}$$

gdzie

$$p_{R,n} = \frac{-p\cos 2\alpha}{1 - p - p\cos 2\alpha + \frac{2(p-1)}{1+\mathrm{tg}^{2n}\alpha}}. \tag{3.31}$$

Możemy się przekonać, że dla $\pi/4 < \alpha < \pi/2$ otrzymujemy $\alpha_n > \alpha$, natomiast po przekształceniu wyrażenia (3.31) do postaci

$$p_{R,n} = p\left[1 + \frac{2(1-p)}{\mathrm{tg}^2\alpha - 1}\left(1 - \frac{1+\mathrm{tg}^2\alpha}{1+\mathrm{tg}^{2n}\alpha}\right)\right]^{-1}, \tag{3.32}$$

i uwzględnieniu faktu, że dla $\pi/4 < \alpha < \pi/2$ mamy $1 < \mathrm{tg}^2\alpha < \mathrm{tg}^{2n}\alpha$, otrzymujemy ostatecznie $p_{R,n} < p$. Tym samym, jeśli ograniczymy się początkowo do stanów lokalnych ρ_R, po dowolnie wielu wymianach splątania pozostajemy w obszarze parametrów, dla których niemożliwe będzie złamanie nierówności CHSH (por. Rys. 3.2).

3.5.3 Niesymetryczny łańcuch stanów

Do rozpatrzenia pozostaje przypadek niesymetrycznego łańcucha, kiedy to wymiana splątania dokonywana jest w łańcuchu stanów z różą liczbą stanów ρ_L i ρ_R po obu jego stronach. Załóżmy, że w dowolnie długim łańcuchu użytkownicy P_{-k+1}, P_{-k+2},..., P_{k-1}, P_{k+1},..., P_{k+n-1} otrzymują w wyniku pomiaru Bella stany $|\Psi^{\pm}\rangle$. Z wcze-

śniejszych analiz wynika, że użytkownicy P_{-k} i P_k będą wtedy współdzielić stan (3.14) z kolei użytkownicy P_k i P_{k+n} stan (3.30). Ponadto wiadomo, iż ten ostatni stan nie pozwala na złamanie nierówności CHSH.

Niech teraz użytkownik P_k jako wynik pomiaru otrzyma $|\Psi^{\pm}\rangle$. Wtedy użytkownicy P_{-k} i P_{k+n} będą współdzielić stan

$$\begin{aligned}\rho_{R,n,k} &= \frac{\mathrm{Tr}_{P_k}[(id_{2\times 2}\otimes|\Psi^+\rangle\langle\Psi^+|\otimes id_{2\times 2})\rho_k\otimes\rho_{R,n}]}{\mathrm{Tr}[(id_{2\times 2}\otimes|\Psi^+\rangle\langle\Psi^+|\otimes id_{2\times 2})\rho_k\otimes\rho_{R,n}]}\\ &= p_{R,n,k}|\Psi_{R,n}\rangle\langle\Psi_{R,n}| + (1-p_{R,n,k})|00\rangle\langle 00|,\end{aligned} \tag{3.33}$$

gdzie

$$p_{R,n,k} = p_{R,n}\left(1+\frac{2p_{R,n}(\frac{1}{p_k}-1)}{1+\mathrm{tg}^{2n}\,\alpha}\right)^{-1}. \tag{3.34}$$

Zauważmy, że powyższy stan jest również postaci (3.30) oraz, że $p_{R,n,k} < p_{R,n}$. Wobec tego stan $\rho_{R,n,k}$ nie pozwala na złamanie nierówności CHSH.

Rozdział 4

Miary kontekstualności

4.1. Wstęp

W bieżącym rozdziale wprowadzimy definicje miar pozwalających na ilościowanie kontekstualności dla rodzin rozkładów prawdopodobieństw, tj. „układów” (ang. „boxes”) w teoriach probabilistycznych [14]. Po omówieniu ich podstawowych własności obliczymy wartości szczególnych rodzajów miar dla wybranych układów kontekstualnych.

4.2. Wprowadzenie miar kontekstualności

4.2.1 Podstawowe pojęcia

Niech V będzie zbiorem k różnych obserwabli $V = \{A_1, ..., A_k\}$ takich, że pewne z nich są współmierzalne. Każdy podzbiór zbioru V, w którym wszystkie obserwable są współmierzalne nazywamy *kontekstem*. Różne konteksty oznaczamy indeksem c. Łączny rozkład prawdopodobieństwa wyników współmierzalnych obserwabli w danym kontekście c oznaczamy przez $g(\lambda_c)$, natomiast zbiór rozkładów $g(\lambda_c)$ dla wszystkich możliwych kontekstów $\{g(\lambda_c)\}$ nazywamy *układem*.

Układ, dla którego rozkłady prawdopodobieństw $g(\lambda_c)$ dla wszystkich kontekstów c otrzymać można z *jednego* łącznego rozkładu prawdopodobieństwa dla wszystkich obserwabli V jako rozkłady brzegowe otrzymywane przez zawężanie ich do obserwabli występujących w danym kontekście, nazywamy *układem niekontekstualnym*. W przeciwnym przypadku układ nazywamy *układem kontekstualnym*.

Wprowadźmy teraz pojęcie hipergrafu $G = (V_G, E_G)$ złożonego ze zbioru k wierzchołków V_G oraz zbioru krawędzi E_G. Każdy wierzchołek utożsamić możemy z konkretną obserwablą w taki sposób, że $V_G = \{A_1, ..., A_k\}$, natomiast każdą krawędź hipergrafu utożsamić możemy ze zbiorem współmierzalnych obserwabli tak, że dla każdego kontekstu c mamy $c = \{A_{i_1}, ..., A_{i_{|c|}}\} \in E_G$, gdzie przez $|c|$ oznaczyliśmy liczbę współmierzalnych obserwabli należących do kontekstu c.

Niezależnie od pojęcia układu scharakteryzowanego wyżej, opisać możemy układy stowarzyszone z danym hipergrafem. Wejściem takiego układu jest wektor $\mathbf{x}$. Ilość różnych wejść układu (wektorów $\mathbf{x}$) jest równa liczbie krawędzi hipergrafu $n = |E_G|$ (czyli liczbie wszystkich kontekstów danego grafu). Wyjściem układu jest wektor $\mathbf{a}$. Ilość różnych wyjść układu d jest równa iloczynowi ilości wyników wszystkich obserwabli A_i wchodzącym w skład danego kontekstu. Zbiór takich układów oznaczamy

przez $B_G^{(k)}$. Układ nazywamy *układem stowarzyszonym z danym hipergrafem G* jeśli jest ono rodziną n rozkładów prawdopodobieństwa takich, że dla wszystkich kontekstów $c \in E_G$ istnieje odpowiadający mu rozkład prawdopodobieństwa w tej rodzinie na przestrzeni probabilistycznej rozpiętej na $\Omega(A_{i_1}) \times ... \times \Omega(A_{i_{|c|}})$. Taką rodzinę rozkładów prawdopodobieństw oznaczamy przez $\{P(\mathbf{a}|x_i)\}$, gdzie $x_i \in E_G$.

Dla danego hipergrafu $G = (V_G, E_G)$ układ $B \in B_G^{(k)}$ nazywamy *układem zgodnym* jeśli dla wszystkich par różnych kontekstów $c, c' \in E_G$ i dla wszystkich podzbiorów obserwabli S takich, że $S = c \cap c' \neq \emptyset$, zachodzi:

$$\forall_s \sum_t P(S = s, T = t|x = c) = \sum_{t'} P(S = s, T' = t'|x = c'), \tag{4.1}$$

gdzie $T = c - S$ oraz $T' = c' - S$. Zbiór wszystkich zgodnych układów stowarzyszonych z danym hipergrafem G oznaczamy przez $C_G^{(n)}$. Zgodność układu stanowi w szczególności, że rozkład brzegowy dla pewnej obserwabli należącej do różnych kontekstów c i c', liczony z dwóch różnych rozkładów prawdopodobieństwa, tj. $g(\lambda_c)$ i $g(\lambda_{c'})$ musi być taki sam. Tak zdefiniowana zgodność układu jest uogólnieniem warunku niesygnalizowania dla układów nielokalnych w ogólnych teoriach probabilistycznych.

Układem niekontekstualnym stowarzyszonym z danym hipergrafem G nazywamy zgodny układ, dla którego istnieje łączny rozkład prawdopodobieństwa dla wszystkich obserwabli w V_G. Zbiór wszystkich układów niekontekstualnych stowarzyszonych z G oznaczamy przez NC_G. Z kolei wszystkie układy nie spełniające danego warunku nazywamy *układami kontekstualnymi*.

W dalszej części rodziny rozkładów prawdopodobieństwa odpowiadające układom kontekstualnym ($B \in C_G^{(n)}$) będziemy oznaczać przez $\{g(\lambda_c)\}$, gdzie c numeruje różne konteksty od 1 do n. Jeśli nie jest zaznaczone inaczej zakładamy, że $n \geq 3$, jako że dla $n \leq 2$ wszystkie układy stowarzyszone z hipergrafem G są niekontekstualne. Jeśli natomiast układ jest niekontekstualny ($B \in NC_G$), odpowiadające mu rodziny rozkładów prawdopodobieństw będziemy oznaczać przez $\{p(\lambda_c)\}$, wreszcie przez $p(\lambda)$ oznaczać będziemy łączny rozkład prawdopodobieństwa dla wszystkich obserwabli w V_G, z którego wszystkie rozkłady $\{p(\lambda_c)\}$ otrzymujemy jako odpowiednie rozkłady brzegowe dla danych kontekstów c.

Lemat 1. *Zgodny układ stowarzyszony z hipergrafem $G = (\{A_1, ..., A_k\}, E_G)$ jest niekontekstualne wtedy i tylko wtedy, gdy może być wyrażone jako kombinacja wypukła układów deterministycznych, tj. takich, że łączny rozkład prawdopodobieństwa dla wszystkich obserwabli $A_1, ..., A_k$ jest równy $\delta_{\mathbf{a_0},\mathbf{a}}$ dla pewnego ustalonego $\mathbf{a_0}$.*

Dowód. Wynika to wprost z definicji układów niekontekstualnych. Łączny rozkład prawdopodobieństwa dla wszystkich obserwabli $A_1, ..., A_k$ jest sumą wypukłą łącznych rozkładów prawdopodobieństw będących rozkładami deterministycznymi.■

Przykład 1. Rozważmy hipergraf z Rys. 4.1, w którym obserwable ze zbioru $V_G = \{A_1, A_2, A_3, A_4\}$ tworzą 4 konteksty współmierzalnych par dychotomicznych obserwa-

bli $A_i, A_{i+1 \bmod 4}$. Niech teraz rozkłady prawdopodobieństw będą takie, że dla $i \in \{1,2,3,4\}$:

$$p(00|A_iA_{i+1}) = p(01|A_iA_{i+1}) = p(10|A_iA_{i+1}) = p(11|A_iA_{i+1}) = \frac{1}{4}. \tag{4.2}$$

Łatwo zauważyć, że taki układ jest układem zgodnym, gdyż np.:

$$\forall_s \sum_k p(ks|A_1A_2) = \sum_l p(sl|A_2A_3) = \frac{1}{2}, \tag{4.3}$$

$(k,l,s \in \{0,1\})$ oraz analogicznie dla każdej pary kontekstów posiadających wspólną obserwablę. Ponadto, zgodny układ stowarzyszony z takim hipergrafem jest układem niekontekstualnym, ponieważ rozkład prawdopodobieństwa $p(kl|A_iA_{i+1})$ dla każdego kontekstu możemy otrzymać z łącznego rozkładu prawdopodobieństwa dla wszystkich obserwabli $p(klmn|A_1A_2A_3A_4)$ takiego, że dla dowolnych $k,l,m,n \in \{0,1\}$ mamy:

$$p(klmn|A_1A_2A_3A_4) = \frac{1}{16}. \tag{4.4}$$

Istotnie, mamy bowiem np.:

$$\forall_{k,l}\ p(kl|A_1A_2) = \sum_{m,n} p(klmn|A_1A_2A_3A_4) = \frac{1}{4}, \tag{4.5}$$

i analogicznie dla pozostałych kontekstów.

Zauważmy również, że zgodnie z Lematem 1 łączny rozkład prawdopodobieństwa dla wszystkich obserwabli $p(klmn|A_1A_2A_3A_4)$ może być wyrażony jako kombinacja wypukła szesnastu układów deterministycznych $\tilde{p}_\nu$ $(\nu \in \{1,\ldots,16\})$ takich, że:

$$\tilde{p}_\nu(klmn|A_1A_2A_3A_4) = \begin{cases} 1 & \text{dla } klmn = \mathbf{a}_\nu \\ 0 & \text{dla } klmn \neq \mathbf{a}_\nu, \end{cases} \tag{4.6}$$

gdzie $\mathbf{a}_\nu$ są różnymi czteroelementowymi ciągami 0 i 1.

Przykład 2. Rozważmy hipergraf z Rys. 4.1, w którym obserwable ze zbioru $V_G = \{A_1,A_2,A_3,A_4\}$ tworzą 4 konteksty współmierzalnych par dychotomicznych obserwabli $A_i, A_{i+1 \bmod 4}$. Niech teraz rozkłady prawdopodobieństw będą takie, że:

$$g(00|A_iA_{i+1}) = g(11|A_iA_{i+1}) = \frac{1}{2}, \tag{4.7}$$

dla $i = 1,2,3$ oraz

$$g(01|A_iA_{i+1}) = g(10|A_iA_{i+1}) = \frac{1}{2}, \tag{4.8}$$

dla $i = 4$. Zgodny układ stowarzyszony z takim hipergrafem jest znanym w literaturze *układem Popescu-Rohrlicha* [2, 3], który oznaczać będziemy jako *PR*.

W dalszej części, wszystkie układy oparte na dychotomicznych obserwablach posiadające jedynie dwa typy rozkładów prawdopodobieństw: równoprawdopodobne ciągi wyników z parzystością 0 bądź równoprawdopodobne ciągi wyników z parzystością 1, nazywać będziemy *układami binarnymi*.

Przykład 3. Powyższy przykład można uogólnić na przypadek hipergrafu z Rys. 4.2, w którym obserwable ze zbioru $V_G = \{A_1, A_2, ..., A_n\}$ tworzą n kontekstów par współmierzalnych dychotomicznych obserwabli $A_i, A_{i+1 \bmod n}$. Niech również i w tym przypadku rozkłady prawdopodobieństw będą określone w taki sposób, że dla wszystkich kontekstów z wyjątkiem ostatniego mamy równoprawdopodobne ciągi wyników z parzystością 0, natomiast dla ostatniego kontekstu mamy równoprawdopodobne ciągi wyników z parzystością 1. Zgodny układ stowarzyszony z takim hipergrafem nazywać będziemy *układem łańcuchowym* (por. [85]), który oznaczać będziemy jako $CH_{(n)}$. Zauważmy, że układ łańcuchowy $CH_{(4)}$ jest równoważny układowi *PR*.

Przykład 4. Rozważmy hipergraf z Rys. 4.3, w którym obserwable ze zbioru $V_G = \{A_1, A_2, ..., A_9\}$ tworzą 6 kontekstów trójek współmierzalnych dychotomicznych obserwabli $E_G = \{\{A_1, A_2, A_3\}, \{A_4, A_5, A_6\}, \{A_7, A_8, A_9\}, \{A_1, A_4, A_7\}, \{A_2, A_5, A_8\}, \{A_3, A_6, A_9\}\}$. Niech teraz rozkłady prawdopodobieństw będą takie, że:

$$g(001|A_3A_6A_9) = g(010|A_3A_6A_9) = g(100|A_3A_6A_9) = g(111|A_3A_6A_9) = \frac{1}{4}, \tag{4.9}$$

dla kontekstu $\{A_3, A_6, A_9\}$ oraz

$$g(011|A_iA_jA_k) = g(101|A_iA_jA_k) = g(110|A_iA_jA_k) = g(000|A_iA_jA_k) = \frac{1}{4}, \tag{4.10}$$

dla pozostałych kontekstów. Zgodny układ stowarzyszony z takim hipergrafem nazywać będziemy *układem Peresa-Mermina* [6–8], który oznaczać będziemy jako *PM*. Układ *PM* oddaje statystykę wyników obserwabli z kwadratu Mermina mierzonych na stanie maksymalnie mieszanym dwóch qubitów [8].

W ogólności dla układów binarnych będziemy mieć do czynienia jedynie z dwoma typami rozkładów:

$$g(\lambda_c) = P^{(|c|)}_{\text{parz}} \equiv \frac{1}{2^{|c|-1}}, \tag{4.11}$$

dla kontekstów, dla których wyniki obserwabli a_i spełniają $\bigoplus_i a_i = 0$, oraz

$$g(\lambda_c) = P^{(|c|)}_{\text{nparz}} \equiv \frac{1}{2^{|c|-1}}, \tag{4.12}$$

dla kontekstów, dla których wyniki obserwabli a_i spełniają $\bigoplus_i a_i = 1$. Przykładowo: $g(\lambda_c) = P^{(2)}_{\text{parz}}$ dla kontekstu $c = \{A_2, A_3\}$ w przypadku układu PR, $g(\lambda_c) = P^{(3)}_{\text{nparz}}$ dla kontekstu $c = \{A_3, A_6, A_9\}$ w przypadku układu PM, etc.

Przykład 5. Rozważmy hipergraf z Rys. 4.4, w którym obserwable ze zbioru $V_G = \{A, B, C, D, E, a, b, c, d, e\}$ tworzą 5 kontekstów czwórek współmierzalnych dychotomicznych obserwabli $E_G = \{\{B, e, a, D\}, \{D, b, c, A\}, \{A, d, e, C\}, \{C, a, b, E\}, \{E, c, b, D\}\}$. Niech teraz rozkłady prawdopodobieństw będą takie, że $g(\lambda_c) = P^{(4)}_{\text{parz}}$ dla kontekstu $c = \{E, c, b, D\}$, oraz $g(\lambda_c) = P^{(4)}_{\text{nparz}}$ dla pozostałych kontekstów. Zgodny układ stowarzyszony z takim hipergrafem nazywać będziemy *układem Mermina* [7,8], który oznaczać będziemy jako M. Układ M oddaje statystykę wyników obserwabli z „gwiazdy" Mermina mierzonych na stanie maksymalnie mieszanym trzech qubitów [8].

Dla danego układu binarnego B zdefiniujmy również *układ przeciwny do B*, który oznaczać będziemy przez B', jako układ, dla którego rozkłady

$$g(\lambda_c) = P^{(|c|)}_{\text{nparz}}, \tag{4.13}$$

dla kontekstów, dla których wyniki obserwabli a_i spełniają $\bigoplus_i a_i = 0$, oraz

$$g(\lambda_c) = P^{(|c|)}_{\text{parz}}, \tag{4.14}$$

dla kontekstów, dla których wyniki obserwabli a_i spełniają $\bigoplus_i a_i = 1$, natomiast rozkłady $P^{(|c|)}_{\text{nparz}}$ oraz $P^{(|c|)}_{\text{parz}}$ dla układu B dane są przez odpowiednio (4.12) oraz (4.11).

Tym samym, hipergraf $G_{B'}$ dla układu B' będzie miał identyczny rozkład krawędzi i wierzchołków co hipergraf G_B dla układu B, przy czym dla B' krawędzie ciągłe zastąpione będą przez przerywane, natomiast krawędzie przerywane przez ciągłe.

Zwróćmy uwagę, że zarówno układ PM jak i M powstają naturalnie na gruncie mechaniki kwantowej poprzez przypisanie poszczególnym węzłom grafów pewnych określonych obserwabli: dobranie obserwabli w taki sposób jak dokonane zostało w [7, 8], działających na przestrzeniach Hilberta odpowiednio $\mathscr{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$ oraz $\mathscr{H} = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$, prowadzi do logicznej sprzeczności bez względu na postać stanu kwantowego, na którym mierzone są obserwable (tzw. kontekstualność stanowo-niezależna). Tym samym, statystykę wyników pomiarów postaci (4.11) oraz (4.12) możemy otrzymać mierząc wybrane obserwable na stanie maksymalnie mie-

szanym.

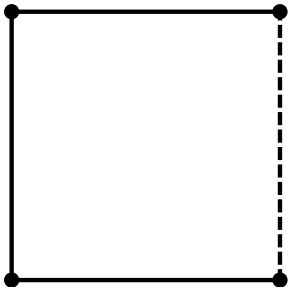
W przypadku układów łańcuchowych $CH_{(n)}$ statystyki dane przez (4.11) oraz (4.12) nie są osiągalne w ramach mechaniki kwantowej, a zatem w ogólności układy łańcuchowe opisywalne są w ramach szerszych teorii probabilistycznych. W dalszej części jednak wprowadzimy pojęcie tzw. *układów izotropowych*, które są klasą uogólniającą omawianych wyżej rodzajów układów. W zależności od stopnia korelacji pewne układy izotropowe można otrzymać w ramach mechaniki kwantowej [92], jednak dalsze rozważania prowadzone będą bez odwoływania się do ograniczeń z niej wynikających.

Do tej pory rozpatrywaliśmy układy $B \in C_G^{(n)}$ stowarzyszone z hipergrafem G, którego wierzchołki V_G utożsamiane były z obserwablami, z kolei krawędzie E_G wyznaczały konteksty współmierzalnych obserwabli. Możemy jednakże ograniczyć tak zdefiniowaną klasę układów i przyjąć w miejsce obserwabli projektory o rzędzie 1. Tym samym, współmierzalność dwóch projektorów oznaczać będzie ich wzajemną ortogonalność. Tak zdefiniowany hipergraf będziemy nazywać *hipergrafem ortogonalności*

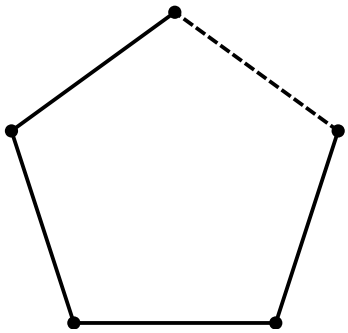
Przykład 6. Rozważmy hipergraf z Rys. 4.5, w którym obserwable ze zbioru $V_G = \{\Pi_1, \Pi_2, \Pi_3, \Pi_4, \Pi_5\}$ tworzą 5 kontekstów par ortogonalnych operatorów rzutowych rzędu jeden $\{\Pi_i, \Pi_{i+1 \bmod 5}\}$. Widzimy, że struktura grafu jest identyczna z hipergrafem $G_{CH}^{(5)}$. W tym przypadku jednak, w odróżnieniu od wcześniejszych przykładów mamy dodatkowe ograniczenie: współmierzalność operatorów rzutowych oznacza ich ortogonalność, co wymusza zerowe prawdopodobieństwo otrzymania wyniku 1 dla dwóch sąsiednich operatorów jednocześnie. Ponadto, wybierzmy stan kwantowy, na którym dokonywane mają być pomiary działające w przestrzeni Hilberta $\mathbb{H} = \mathbb{C}^3$ w taki sposób, by dla każdego kontekstu prawdopodobieństwa otrzymania określonych wyników pomiarów wynosiły [9]:

$$\begin{aligned} g(01|\Pi_i\Pi_{i+1}) = g(10|\Pi_i\Pi_{i+1}) &= \frac{1}{\sqrt{5}}, \\ g(11|\Pi_i\Pi_{i+1}) &= 0, \\ g(00|\Pi_i\Pi_{i+1}) &= 1 - \frac{2}{\sqrt{5}}. \end{aligned} \tag{4.15}$$

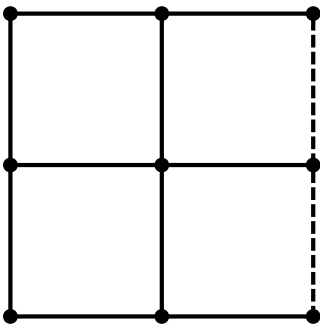
Tak zdefiniowany układ nazywać będziemy *układem KCBS* [9], który oznaczać będziemy jako K. Układ K, skonstruowany ze stanu czystego qutritu wraz z odpowiednio dobranym zbiorem pięciu operatorów rzutowych, jest najprostszym kwantowym układem, dla którego można wykazać cechy kontekstualności; nie istnieje bowiem kontekstualność dla pojedynczego qubitu [89].



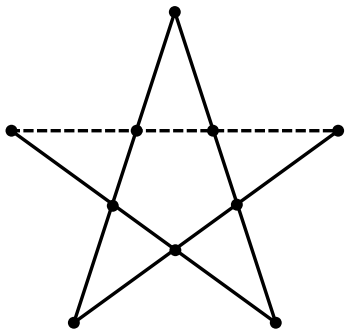
Rys. 4.1: Hipergraf G_{PR} dla układu PR. Krawędzie ciągłe przedstawiają konteksty posiadające równoprawdopodobne ciągi wyników z parzystością 0, natomiast krawędź przerywana przedstawia kontekst posiadający równoprawdopodobne ciągi wyników z parzystością 1.



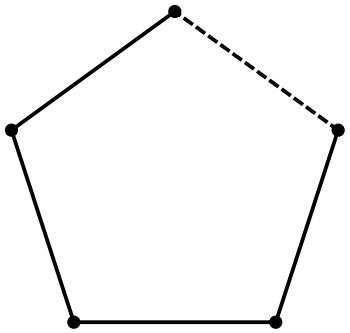
Rys. 4.2: Przykładowy hipergraf $G_{CH_{(n)}}$ $(n = 6)$ dla układu łańcuchowego $CH_{(n)}$.



Rys. 4.3: Hipergraf G_{PM} dla układu PM.



Rys. 4.4: Hipergraf G_M dla układu M.



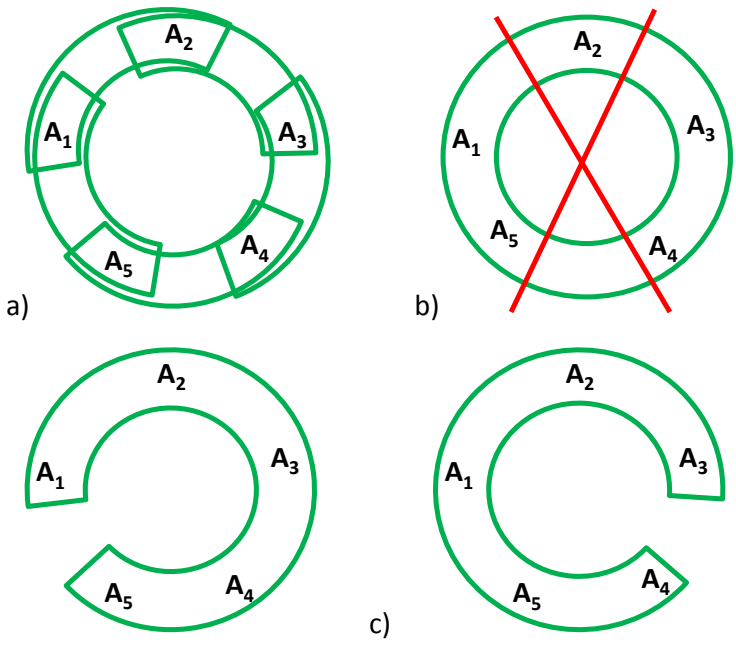
Rys. 4.5: Hipergraf $G_{CH_{(5)}}$ dla układu K. Krawędzie przedstawiają relacje ortogonalności projektorów utożsamianych z węzłami grafu.

4.2.2 Wzajemna informacja kontekstualności

Istnieją stany kwantowe i zbiory obserwabli, dla których statystyka nie może być jednoznacznie opisana przez jeden łączny rozkład prawdopodobieństwa [93]. Innymi słowy, jeśli w zbiorze obserwabli V_G istnieją różne konteksty, to jeden łączny rozkład prawdopodobieństwa oddający poprawne statystyki dla pewnych kontekstów nie będzie mógł jednocześnie oddawać poprawnych statystyk dla innych kontekstów. Tym samym, chcąc poprawnie opisać statystykę danego układu, potrzeba przynajmniej dwóch *różnych* łącznych rozkładów prawdopodobieństwa: jeden rozkład prawdopodobieństwa poprawnie opisujący statystykę wyników pomiarów dla jednego zbioru kontekstów a drugi dla drugiego zbioru kontekstów (zob. Rys. 4.6). Statystyka wyników pomiarów dla każdego kontekstu jest poprawnie opisana przynajmniej przez jeden rozkład prawdopodobieństwa. W szczególności, możemy zażądać, by dla każdego kontekstu istniał inny łączny rozkład prawdopodobieństwa (tylko dla obserwabli wchodzących w skład danego kontekstu) poprawnie go opisujący. Widzimy, że dla takich kontekstualnych układów wybrane (łączne) rozkłady prawdopodobieństwa są skorelowane z kontekstami, których statystykę poprawnie opisują. Możemy więc mówić, że w danym układzie istnieje ukryta informacja mówiąca o wyborze danego kontekstu, która jest nieobecna dla układów niekontekstualnych, a zatem takich, które mogą być opisywane przez jeden łączny rozkład prawdopodobieństwa dla wszystkich obserwabli.

W tym miejscu wprowadzimy korelacyjną definicję miary kontekstualności, która dla danego układu ilościuje korelacje pomiędzy zmienną losową stanowiącą o wyborze danego kontekstu a rozkładem zmiennych dla wszystkich obserwabli w V_G.

Bieżące zagadnienie zilustrujemy za pomącą gry „zgadnij kontekst” pomiędzy trzema użytkownikami: nadawcą (Alicja), odbiorcą (Bolek) oraz przeciwnikiem (Czarek). Początkowo użytkownicy ustalają, że gra oparta będzie na wcześniejszym wyborze pewnego ustalonego układu $B = \{g(\lambda_c)\}$ posiadanego przez Alicję. Celem gry jest zakomunikowanie Bolkowi informacji dotyczącej tego, który kontekst wybrała Alicja (informacją będzie wtedy liczba c numerująca wybrany przez Alicję kontekst), przy



Rys. 4.6: Schematyczne przedstawienie układu kontekstualnego dla obserwabli $A_1,...,A_5$ (konteksty tworzą pary sąsiednich obserwabli): a) Statystyka układu opisana jest przed pięć niezależnych łącznych rozkładów, z których każdy opisuje statystykę wybranego kontekstu; b) Nie istnieje natomiast jeden łączny rozkład prawdopodobieństwa dla wszystkich obserwabli; c) Możliwy opis statystyki układu z użyciem dwóch *różnych* łącznych rozkładów prawdopodobieństwa, każdy z nich nie oddaje poprawnie statystyki pewnego kontekstu: lewy A_1A_5, prawy A_3A_4. Źródło: [14].

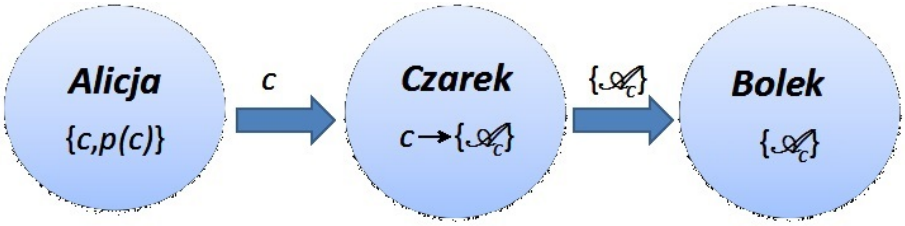
czym komunikat do Bolka przechodzi przez ręce Czarka (zob. Rys. 4.7). W tym celu Alicja musi wybrać najlepszy możliwy rozkład $\{p(c)\}$ (celem maksymalizacji powodzenia zakomunikowania wybranego kontekstu), po czym posyła do Czarka liczbę c wybraną zgodnie z rozkładem $\{p(c)\}$.

Zadaniem Czarka jest minimalizacja wzajemnej informacji między Alicją i Bolkiem, przy założeniu, że konstruuje on zmienną $\mathscr{A}_c$ dla wszystkich zmiennych w zbiorze obserwabli V_G o odpowiednim rozkładzie prawdopodobieństwa, takim że rozkład ten będzie zgodny z rozkładem zawężonym do wybranego kontekstu c, czyli $g(\lambda_c)$. Zauważmy, że im bardziej rozróżnialne będą rozkłady prawdopodobieństwa zmiennej $\mathscr{A}_c$ w zależności od wybranego numeru kontekstu c, tym łatwiej Bolkowi zgadnąć wybraną przez Alicję wartość c.

Ilość korelacji między Alicją i Bolkiem osiąganych w tej grze, przy założeniu, że Alicja dokonuje wyboru c mającego rozkład $\{p(c)\}$, wynosi:

$$I_{\{p(c)\}}(B) := \min_{\mathscr{A}_c} I(\sum_c p(c)|c\rangle\langle c| \otimes \mathscr{A}_c), \tag{4.16}$$

która to wielkość definiuje *wzajemną informację kontekstualności przy rozkładzie* $\{p(c)\}$



Rys. 4.7: Schemat gry „zgadnij kontekst": Alicja ogłasza wybrany przez siebie kontekst c, Czarek przygotowuje $\mathscr{A}_c$ o odpowiednim rozkładzie, Bolek znając rozkłady układu wyjściowego wnioskuje wybrany przez nią kontekst. Źródło: [14].

układu B.

Maksymalna wielkość korelacji między Alicją i Bolkiem osiągana będzie przy wyborze takiego rozkładu $\{p(c)\}$, dla którego powyższa wielkość będzie maksymalna. Tym samym, *wzajemną informację kontekstualności* układu B zdefiniujemy jako:

$$I_{\max}(B) := \sup_{\{p(c)\}} I_{\{p(c)\}}(B), \tag{4.17}$$

która mierzy maksymalną ilość korelacji między Alicją i Bolkiem osiąganą w tej grze.

W powyższych definicjach użyliśmy notacji Diraca celem jawnego wskazania na charakter korelacji, które w przypadku tej gry zaistnieć miałyby między rejestrem informującym o wyborze konkretnego kontekstu c z ogólnymi rozkładami zmiennej $\mathscr{A}_c$, które przy zawężeniu do odpowiednich kontekstów dają rozkład $g(\lambda_c)$.

Zauważmy teraz, że zdefiniowane powyżej wielkości służyć mogą do określenia jak bardzo kontekstualne jest dany układ B. W przypadku, gdy B jest układem niekontekstualnym, z definicji istnieje jeden łączny rozkład prawdopodobieństwa $\mathscr{A}$ dla wszystkich obserwabli w zbiorze V_G, który zawężony do dowolnych wybranych kontekstów c daje rozkłady $g(\lambda_c)$. Wobec tego Czarek zawsze może przygotować łączny rozkład prawdopodobieństwa $\mathscr{A}$ taki sam dla wszystkich kontekstów, który jest całkowicie niezależny od numeru wybranego kontekstu, co powoduje zerowe korelacje między rozkładem $\mathscr{A}$ a numerem kontekstu c. Z tego też względu dla niekontekstualnych układów B mamy: $I_{\max}(B) = 0$.

Jeśli natomiast Alicja i Bolek dysponują kontekstualnym układem B, Czarek musi przygotować przynajmniej dwa różne łączne rozkłady $\mathscr{A}_c$ dla wszystkich obserwabli w zbiorze V_G (które również zawężone do wybranych kontekstów c dają rozkłady $g(\lambda_c)$). Wtedy jednak z definicji wzajemnej informacji mamy: $I_{\max}(B) > 0$.

Przykład. Niech umówionym układem w grze „zgadnij kontekst" będzie układ PR. Alicja z prawdopodobieństwem $p(c)$ wybiera liczbę $c \in \{1,2,3,4\}$. Zadaniem Czarka jest skonstruowanie 4 rozkładów prawdopodobieństwa $p(\lambda)$ ($\lambda = ijkl$, $i,j,k,l \in \{0,1\}$) zmiennej $\mathscr{A}_c$ możliwie najmniej rozróżnialnych od siebie. Rozkłady prawdopodobieństwa zmiennej $\mathscr{A}_c$ mogą być przykładowo skonstruowane w następujący spo-

sób:

- dla $c = 1$ rozkład zmiennej $\mathscr{A}_1$ jest taki, że wyniki $ijkl$ są równoprawdopodobne ($p(ijkl) = \frac{1}{6}$) z wyjątkiem wyników 1000 i 0111, dla których $p(1000) = p(0111) = 0$;
- dla $c = 2$ rozkład zmiennej $\mathscr{A}_2$ jest taki, że wyniki $ijkl$ są równoprawdopodobne ($p(ijkl) = \frac{1}{6}$) z wyjątkiem wyników 1100 i 0011, dla których $p(1100) = p(0011) = 0$;
- dla $c = 3$ rozkład zmiennej $\mathscr{A}_3$ jest taki, że wyniki $ijkl$ są równoprawdopodobne ($p(ijkl) = \frac{1}{6}$) z wyjątkiem wyników 1110 i 0001, dla których $p(1000) = p(0111) = 0$;
- dla $c = 4$ rozkład zmiennej $\mathscr{A}_4$ jest taki, że wyniki $ijkl$ są równoprawdopodobne ($p(ijkl) = \frac{1}{6}$) z wyjątkiem wyników 1111 i 0000, dla których $p(0000) = p(1111) = 0$.

Zauważmy, że mamy 4 różne rozkłady prawdopodobieństwa, które jednakże przy zawężeniu do odpowiednich kontekstów (tj. przykładowo rozkładu zmiennej $\mathscr{A}_3$ do trzeciego kontekstu) oddają poprawnie statystykę poszczególnych kontekstów $g(\lambda_c)$ dla układu PR. Jak się okazuje, taki wybór dokonany przez Czarka będzie optymalny, tj. minimalizujący wyrażenie (4.16).

4.2.3 Względna entropia kontekstualności

W tym miejscu wprowadzimy miarę kontekstualności, opartą na pojęciu względnej entropii (dywergencji Kullbacka-Leiblera) [94].

Przy wprowadzaniu *wzajemnej informacji kontekstualności* dla danego układu B korzystaliśmy z faktu, iż nieprawdą jest, że dla kontekstualnego układu B istnieje jeden łączny rozkład prawdopodobieństwa $p(\lambda)$ oddający prawidłowe statystyki dla wszystkich poszczególnych kontekstów $p(\lambda_c)$. Dla danego kontekstu c_0 wybranego przez Alicję, Czarek konstruował łączny rozkład prawdopodobieństwa $\mathscr{A}_{c_0}$ dający prawidłową statystykę dla wybranego kontekstu: $p(\lambda_{c_0})$. W ogólności jednak istnieje taki kontekst c_1, że zawężenie łącznego rozkładu prawdopodobieństwa $\mathscr{A}_{c_0}$ do tego kontekstu nie oddaje poprawnie statystyki $p(\lambda_{c_1})$ dla danego układu B.

Obecnie spojrzymy na problem nieco odmiennie. Mianowicie, zamiast konstruować łączne rozkłady $\mathscr{A}_c$ w zależności od wybranego kontekstu c, spróbujmy skonstruować jeden łączny rozkład prawdopodobieństwa $p(\lambda)$ dla wszystkich obserwabli w zbiorze V_G. Jest oczywiste, że w przypadku układów kontekstualnych łączny rozkład $p(\lambda)$ nie może poprawnie oddawać statystyk dla wszystkich kontekstów jednocześnie. Jeśli z kolei dla danego układu B udałoby się dobrać taki rozkład, wtedy układ ten byłoby niekontekstualny.

Mając pewien układ $B = \{g(\lambda_c)\}$ oraz pewien łączny rozkład prawdopodobieństwa $p(\lambda)$ nad $\Omega(A_1) \times ... \times \Omega(A_k)$ możemy zapytać jak bardzo układ B różniłby się

od układu powstałego w wyniku zawężeń łącznego rozkładu $p(\lambda)$ do poszczególnych kontekstów. Statystyczną odległość między poszczególnymi rozkładami $g(\lambda_c)$ oraz $p(\lambda_c)$ dla wybranego kontekstu c wyrazić można za pomocą względnej entropii (zob. w tym kontekście [18]):

$$D(g(\lambda_c)||p(\lambda_c)) = \sum_i g(\lambda_c)_i \log \frac{g(\lambda_c)_i}{p(\lambda_c)_i}, \tag{4.18}$$

gdzie sumowanie przebiega po wszystkich możliwych wynikach pomiarów w obrębie kontekstu c.

Tym samym dla danego układu $B = \{g(\lambda_c)\} \in C_G^{(n)}$ możemy zdefiniować miarę, która określa jak odległy w sensie statystycznym jest dany układ względem „najbliższego" łącznego rozkładu $p(\lambda)$:

$$X_{\mathrm{u}}(B) := \min_{p(\lambda)} \sum_{c \in E_G} \frac{1}{n} D(g(\lambda_c)||p(\lambda_c)), \tag{4.19}$$

gdzie sumowanie przebiega po wszystkich kontekstach c, a n jest liczbą wszystkich kontekstów (por. [18]). Powyższą wielkość nazwiemy *jednorodną względną entropią kontekstualności*. W przypadku, gdy chcemy określić miarę kontekstualności dla danego układu B, musimy znaleźć minimum po wszystkich możliwych łącznych rozkładach $p(\lambda)$ dających w zawężeniu do danych kontekstów poszczególne rozkłady $p(\lambda_c)$.

Ze względu na charakter pewnych układów B może się zdarzyć, że niektóre konteksty byłyby w pewnym sensie „wyróżnione", tj. względna entropia dla danych kontekstów byłaby większa niż dla pozostałych. Tym samym możemy zdefiniować miarę, która w przypadku układów kontekstualnych mogłaby wyróżniać konteksty, dla których odległość w sensie statystycznym między rozkładami danego układu a wynikającymi z rozkładu łącznego $p(\lambda)$ będzie większa niż dla pozostałych:

$$X_{\max}(B) := \sup_{p(c)} \min_{p(\lambda)} \sum_{c \in E_G} p(c) D(g(\lambda_c)||p(\lambda_c)), \tag{4.20}$$

gdzie supremum brane jest po rozkładach prawdopodobieństw $p(c)$ wyboru poszczególnych kontekstów $\{1, ..., n\}$. Powyższą wielkość nazywamy *względną entropią kontekstualności*.

W ogólności, dla dowolnego układu B zachodzi:

$$X_{\max}(B) \geq X_{\mathrm{u}}(B), \tag{4.21}$$

ponieważ w przypadku *jednorodnej względnej entropii kontekstualności* mamy po prostu $p(c) = \frac{1}{n}$ dla wszystkich c.

4.3. Własności miar kontekstualności

4.3.1 Równość $I_{\max}$ oraz $X_{\max}$

Poniżej wykażemy równoważność dwóch zdefiniowanych wcześniej miar: wzajemnej informacji kontekstualności (4.17) oraz względnej entropii kontekstualności (4.20). Zanim przejdziemy do formalnego twierdzenia, wprowadźmy wykorzystywane w dalszej części wielkości, które odnoszą się do odpowiednich miar kontekstualności dla układu $B = \{g(\lambda_c)\}$ przy ustalonym rozkładzie prawdopodobieństwa wyboru kontekstów $p(c)$: względną entropię kontekstualności dla ustalonego rozkładu $\{p(c)\}$:

$$X_{\{p(c)\}}(B) := \min_{p(\lambda)} \sum_c p(c) D(g(\lambda_c)||p(\lambda_c)), \tag{4.22}$$

oraz wzajemną informację kontekstualności dla łącznych rozkładów $\{g(\lambda|c)\}$:

$$I'_{\{p(c)\}}(B) := \min_{\{g(\lambda|c):g(\lambda_c|c)=g(\lambda_c)\},p(\lambda)} \sum_c p(c) D(g(\lambda|c)||p(\lambda)), \tag{4.23}$$

gdzie $g(\lambda|c)$ są dowolnymi łącznymi rozkładami prawdopodobieństwa dla wszystkich obserwabli, dla których rozkład brzegowy ograniczony do kontekstu c jest zgodny z rozkładem $g(\lambda_c)$.

Ponadto zdefiniujmy maksymalną wzajemną informację kontekstualności dla łącznych rozkładów $\{g(\lambda|c)\}$:

$$I'_{\max}(B) := \sup_{\{p(c)\}} I'_{\{p(c)\}}(B), \tag{4.24}$$

z kolei supremum po $\{p(c)\}$ miary $X_{\{p(c)\}}(B)$ daje oczywiście względną entropię kontekstualności (4.20).

Twierdzenie 1. *Dla dowolnego układu $B = \{g(\lambda_c)\} \in C_G^{(n)}$ mamy:*

$$I_{\max}(B) = X_{\max}(B). \tag{4.25}$$

Dowód.

Kluczową kwestią jest tutaj wykazanie równości trzech miar (4.22), (4.23) oraz (4.16):

$$X_{\{p(c)\}}(B) \overset{(1)}{=} I'_{\{p(c)\}}(B) \overset{(2)}{=} I_{\{p(c)\}}(B), \tag{4.26}$$

dla pewnego ustalonego rozkładu $p(c)$.

Równość (1).

Zauważmy, że miarę $X_{\{p(c)\}}(B)$ otrzymać możemy z $I'_{\{p(c)\}}(B)$ zawężając w tej drugiej rozkłady $g(\lambda|c)$ oraz $p(\lambda)$ do rozkładów brzegowych dla danych kontekstów odpowiednio $g(\lambda_c)$ i $p(\lambda_c)$ tożsamych z rozkładami prawdopodobieństwa występują-

cymi w $X_{\{p(c)\}}(B)$. Proces zawężania łącznych rozkładów prawdopodobieństwa do odpowiednich rozkładów brzegowych nie prowadzi natomiast do zwiększenia względnej entropii, tym samym mamy

$$I'_{\{p(c)\}}(B) \geq X_{\{p(c)\}}(B). \tag{4.27}$$

Niech teraz $p^*(\lambda)$ będzie rozkładem prawdopodobieństwa minimalizującym wielkość $X_{\{p(c)\}}(B)$, dla którego rozkłady brzegowe dla poszczególnych kontekstów wynoszą odpowiednio $p^*(\lambda_c)$. Dla każdego kontekstu możemy znaleźć rozkład prawdopodobieństw warunkowych dla obserwabli nie wchodzących w skład danego kontekstu $p^*(\lambda'_c|\lambda_c)$ taki, że $p^*(\lambda'_c|\lambda_c)p^*(\lambda_c) = p^*(\lambda)$, gdzie $\lambda = \lambda'_c\lambda_c$. Dzięki temu możemy zdefiniować rozszerzenia rozkładów $g(\lambda_c)$ na wszystkie obserwable: $g(\lambda|c) = p^*(\lambda'_c|\lambda_c)g(\lambda_c)$, które z założenia są zgodne z rozkładami brzegowymi ograniczonymi do wybranych kontekstów. Łatwo sprawdzić, że taki wybór optymalnego rozkładu prawdopodobieństwa $p^*(\lambda)$ wysyca nierówność (4.27), dając oczekiwaną równość. Mamy bowiem następujący ciąg równości:

$$\begin{aligned} I'_{\{p(c)\}}(B) &= \sum_c p(c)D(g(\lambda|c)||p^*(\lambda)) \\ &= \sum_c p(c)D(p^*(\lambda'_c|\lambda_c)g(\lambda_c)||p^*(\lambda'_c|\lambda_c)p^*(\lambda_c)) \\ &= \sum_c p(c)D(g(\lambda_c)||p^*(\lambda_c)) \\ &= X_{\{p(c)\}}(B). \end{aligned} \tag{4.28}$$

Równość (2).

Aby wykazać równość $I_{\{p(c)\}}(B)$ oraz $I'_{\{p(c)\}}(B)$ skorzystamy z Lematu 3 podanego w dalszej części rozdziału otrzymując relację:

$$\begin{aligned} I(\sum_c p(c)|c\rangle\langle c| \otimes \mathscr{A}_c) &\equiv \sum_c p(c)D(g(\lambda|c)||\sum_c p(c)g(\lambda|c)) \\ &= \min_{p(\lambda)} \sum_c p(c)D(g(\lambda|c)||p(\lambda)), \end{aligned} \tag{4.29}$$

gdzie rozkład prawdopodobieństwa zmiennej $\mathscr{A}_c$ dany jest przez $g(\lambda|c)$. Minimalizując następnie wyrażenia po rozkładach $g(\lambda|c)$ mających rozkłady brzegowe $g(\lambda_c)$ dla danego układu B otrzymujemy żądaną równość.

Otrzymawszy równość

$$X_{\{p(c)\}}(B) = I_{\{p(c)\}}(B) \tag{4.30}$$

dla dowolnego rozkładu prawdopodobieństwa $p(c)$ bierzemy supremum tych wielkości po tym rozkładzie, co dowodzi równości miar $I_{\max}(B)$ i $X_{\max}(B)$ dla dowolnych zgodnych układów B.■

W dowodzie powyższego twierdzenia wykorzystaliśmy podany w dalszej części Lemat 3, do którego wykazania niezbędny jest poniższy rezultat.

Lemat 2. *Niech będzie dany stan kwantowy ρ_{AB}, dla którego odpowiednie podukłady będą dane przez $\rho_A = \mathrm{Tr}_B \rho_{AB}$ oraz $\rho_B = \mathrm{Tr}_A \rho_{AB}$. Zachodzi relacja*

$$\inf_{\sigma_A,\sigma_B} S(\rho_{AB}||\sigma_A \otimes \sigma_B) = S(\rho_{AB}||\rho_A \otimes \rho_B) \tag{4.31}$$

gdzie S jest kwantową względną entropią [95], natomiast infimum wzięte jest po dowolnych stanach kwantowych σ_A oraz σ_B.

Dowód.

Zauważmy, że $\log(\sigma_A \otimes \sigma_B) = (\log \sigma_A) \otimes id_B + id_A \otimes (\log \sigma_B)$, gdzie id_A oraz id_B są operatorami identycznościowymi na podukładach odpowiednio A i B. Zachodzi:

$$\begin{aligned}
&S(\rho_{AB}||\sigma_A \otimes \sigma_B) \\
&= \mathrm{Tr}\rho_{AB} \log \rho_{AB} - \mathrm{Tr}\rho_{AB} \log(\sigma_A \otimes \sigma_B) \\
&= -S(\rho_{AB}) - Tr\rho_A \log \sigma_A - Tr\rho_B \log \sigma_B \\
&= -S(\rho_{AB}) + S(\rho_A) + S(\rho_B) + [-S(\rho_A) - Tr\rho_A \log \sigma_A] + [-S(\rho_B) - Tr\rho_B \log \sigma_B] \\
&= S(\rho_{AB}||\rho_A \otimes \rho_B) + S(\rho_A||\sigma_A) + S(\rho_B||\sigma_B),
\end{aligned} \tag{4.32}$$

gdzie w drugiej równości skorzystaliśmy z faktu $\mathrm{Tr}\rho_{AB} M \otimes id_B = \mathrm{Tr}\rho_A M$ oraz analogicznie dla drugiego podukładu, natomiast wielkość $S(\rho_{AB}||\rho_A \otimes \rho_B)$ jest równa kwantowej wzajemnej informacji stanu ρ_{AB}, $I(\rho_{AB}) = S(\rho_A) + S(\rho_B) - S(\rho_{AB})$. Minimalizując wyrażenia w relacji (4.32) po dowolnych stanach σ_A oraz σ_B dostajemy tezę lematu.■

Lemat 3. *Dla dowolnego zespołu kwantowego $\{p(c), \rho_c\}$, zachodzi:*

$$I(\sum_c p(c)|c\rangle\langle c| \otimes \rho_c) = \inf_\sigma \sum_c p(c) S(\rho_c||\sigma), \tag{4.33}$$

gdzie infimum jest wzięte po dowolnych stanach kwantowych σ.

Dowód.

Oznaczmy stan ρ jako:

$$\rho := \sum_c p(c)|c\rangle\langle c| \otimes \rho_c, \tag{4.34}$$

którego podukłady to odpowiednio $\sum_c p(c)|c\rangle\langle c|$ oraz $\sum_c p(c)\rho_c =: \sigma$. Tym samym kwantowa wzajemna informacja stanu ρ wyraża się przez

$$I(\rho) \equiv S\left(\sum_c p(c)|c\rangle\langle c| \otimes \rho_c || (\sum_c p(c)|c\rangle\langle c|) \otimes (\sum_c p(c)\rho_c)\right). \tag{4.35}$$

Na mocy Lematu 2 mamy:

$$S\left(\rho||(\sum_c p(c)|c\rangle\langle c|)\otimes(\sum_c p(c)\rho_c)\right)=\inf_{\sigma_A,\sigma_B} S(\rho||\sigma_A\otimes\sigma_B). \tag{4.36}$$

Z kolei wiedząc, że podukład $\sum_c p(c)|c\rangle\langle c|$ jest optymalnym stanem σ_A minimalizującym powyższą relację, otrzymujemy:

$$S\left(\rho||(\sum_c p(c)|c\rangle\langle c|)\otimes(\sum_c p(c)\rho_c)\right)=\inf_{\sigma} S\left(\rho||(\sum_c p(c)|c\rangle\langle c|)\otimes\sigma\right). \tag{4.37}$$

Teraz

$$\begin{aligned}
&S\left(\rho||(\sum_c p(c)|c\rangle\langle c|)\otimes\sigma\right)\\
&=S(\sum_c p(c)|c\rangle\langle c|)+S(\sigma)-S(\rho)\\
&=H(\{p(c)\})-\mathrm{Tr}\sigma\log\sigma-H(\{p(c)\})-\sum_c p(c)S(\rho_c)\\
&=-\mathrm{Tr}(\sum_c p(c)\rho_c)\log\sigma-\sum_c p(c)S(\rho_c)\\
&=-\sum_c p(c)\mathrm{Tr}\rho_c\log\sigma+\sum_c p(c)\mathrm{Tr}\rho_c\log\rho_c\\
&=\sum_c p(c)\left(\mathrm{Tr}\rho_c\log\rho_c-\mathrm{Tr}\rho_c\log\sigma\right)\\
&=\sum_c p(c)S(\rho_c||\sigma),
\end{aligned} \tag{4.38}$$

co dowodzi tezy lematu.■

W tym miejscu należy dodatkowo uzupełnić pozorną rozbieżność między znajdywaniem *minimum* w definicji miary $X_{\max}$ a *infimum* w powyższym lemacie. Zauważmy, że zbiór stanów σ jest wypukły i zwarty. Teraz, ze względu na fakt, że względna entropia jest półciągła z dołu [96], zawsze istnieje pewien stan σ^*, dla którego jest osiągane minimum.

Uwaga. W powyższych lematach wykorzystaliśmy pojęcia i formalizm mechaniki kwantowej. Należy przy tym zaznaczyć, że otrzymane rezultaty odnoszą się również do klasycznych rozkładów prawdopodobieństw, które można opisać w ramach formalizmu kwantowego. W szczególności, w przypadku klasycznych rozkładów prawdopodobieństwa mamy równoważność względnej entropii oraz jej kwantowego odpowiednika. Użycie stanów kwantowych ρ_c odpowiada użyciu zmiennych $\mathscr{A}_c$ o rozkładach $g(\lambda|c)$, minimalizacja po stanach σ odpowiada minimalizacji po rozkładach $p(\lambda)$, etc.

4.3.2 Relacja między $X_{\max}$ oraz X_{u}

W podrozdziale 4.4.3 pokażemy, że w przypadku układów izotropowych miary X_{u} oraz $X_{\max}$ są równe. W bieżącym podrozdziale wykażemy jednak, że miary X_{u} oraz $X_{\max}$ w ogólnym przypadku są różne. Jako przykład podamy wartość miar dla *sum prostych* układów. W pierwszej kolejności zdefiniujmy pojęcia, które pojawią się w dalszej części.

Niech dowolne dwa układy $B_1 = \{g(\lambda_{c^1})\}_{c^1 \in E_{G_1}}$ oraz $B_2 = \{g(\lambda_{c^2})\}_{c^2 \in E_{G_2}}$ będą układami stowarzyszonymi z hipergrafami odpowiednio $G_1 = (V_{G_1}, E_{G_1})$ oraz $G_2 = (V_{G_2}, E_{G_2})$. Wprowadźmy pojęcie sumy prostej dwóch hipergrafów jako:

$$G_1 \oplus G_2 := (V_{G_1 \oplus G_2}, E_{G_1 \oplus G_2}), \tag{4.39}$$

gdzie $V_{G_1 \oplus G_2} = V_{G_1} \cup V_{G_2}$ oraz $E_{G_1 \oplus G_2} = E_{G_1} \cup E_{G_2}$. *Sumą prostą dwóch układów* będziemy nazywali

$$B_1 \oplus B_2 := \{g(\lambda_c)\}_{c \in E_{G_1 \oplus G_2}}. \tag{4.40}$$

Niech V oznacza pewien wybrany zbiór obserwabli $V = \{A_1, ..., A_n\}$. Łączny rozkład prawdopodobieństwa dla otrzymanych wyników wszystkich obserwabli ze zbioru V oznaczać będziemy przez $p(\lambda)[V]$. Jeśli teraz ze zbioru V wyodrębnimy pewien jego podzbiór V_i, to łączny rozkład prawdopodobieństwa dla otrzymanych wyników obserwabli ze zbioru V_i, będący rozkładem brzegowym pochodzącym z $p(\lambda)[V]$, oznaczać będziemy przez $p(\lambda)[V]|_{V_i}$. Oznaczmy również przez $D(g(\lambda_c)||p(\lambda_c))\big|_{p(\lambda)}$ względną entropię rozkładów prawdopodobieństw dla wybranego kontekstu c, takich że rozkład $g(\lambda_c)$ brany jest z układu $\{g(\lambda_c)\}$, natomiast $p(\lambda_c)$ pochodzi z pewnego ustalonego łącznego rozkładu $p(\lambda)$.

Lemat 4. *Dla dowolnych układów $B_1 \in C_{G_1}^{(n_1)}$ oraz $B_2 \in C_{G_2}^{(n_2)}$ $(n_1, n_2 \geq 1)$ stowarzyszonych z hipergrafami odpowiednio $G_1 = (V_{G_1}, E_{G_1})$ oraz $G_2 = (V_{G_2}, E_{G_2})$, zachodzi:*

$$X_{\mathrm{u}}(B_1 \oplus B_2) = \min_{p(\lambda)[V_{G_1}]p(\lambda)[V_{G_2}]} \sum_{c \in E_{G_1 \oplus G_2}} \frac{1}{n_1 + n_2} D(g(\lambda_c)||p(\lambda_c)), \tag{4.41}$$

oraz

$$X_{\max}(B_1 \oplus B_2) = \sup_{\{p(c)\}} \min_{p(\lambda)[V_{G_1}]p(\lambda)[V_{G_2}]} \sum_{c \in E_{G_1 \oplus G_2}} p(c) D(g(\lambda_c)||p(\lambda_c)), \tag{4.42}$$

gdzie minimum brane jest po łącznym rozkładzie będącym rozkładem produktowym rozkładów łącznych $p(\lambda)[V_{G_1}]$ oraz $p(\lambda)[V_{G_2}]$.

Dowód.

Zauważmy, że dla dowolnego rozkładu prawdopodobieństwa dla kontekstów $\{p(c)\}$

oraz dla dowolnych rozkładów łącznych $p(\lambda)[V_{G_1\oplus G_2}]$, zachodzi

$$\begin{aligned}
&\sum_{c\in E_{G_1}\cup E_{G_2}} p(c)D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_1\oplus G_2}]} \\
&= \sum_{c\in E_{G_1}} p(c)D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_1\oplus G_2}]|_{V_{G_1}}} + \sum_{c\in E_{G_2}} p(c)D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_1\oplus G_2}]|_{V_{G_2}}} \\
&= \sum_{c\in E_{G_1}\cup E_{G_2}} p(c)D(g(\lambda_c)||p(\lambda_c))|_{p(\lambda)[V_{G_1\oplus G_2}]|_{V_{G_1}}\, p(\lambda)[V_{G_1\oplus G_2}]|_{V_{G_2}}},
\end{aligned} \tag{4.43}$$

gdzie skorzystaliśmy z faktu, iż z definicji sumy prostej układów $B_1 \oplus B_2$ konteksty ze zbioru E_{G_1} (odpowiednio E_{G_2}) zależą wyłącznie od zmiennych ze zbioru V_{G_1} (V_{G_2}). Tym samym możemy napisać

$$X_{\{p(c)\}}(B_1 \oplus B_2) = \min_{p(\lambda)[V_{G_1}]p(\lambda)[V_{G_2}]} \sum_{c\in E_{G_1}\cup E_{G_2}} p(c)D(g(\lambda_c)||p(\lambda_c)), \tag{4.44}$$

co dla rozkładu $p(c) = \frac{1}{n_1+n_2}$ dla wszystkich kontekstów c prowadzi bezpośrednio do (4.41), natomiast wzięcie supremum po rozkładach $\{p(c)\}$ daje (4.42).■

Powyższy lemat pozwala nam wyrazić miary X_{u} oraz $X_{\max}$ sumy prostej dwóch układów jako funkcje miar X_{u} oraz $X_{\max}$ liczonych dla pojedynczych układów.

Twierdzenie 2. *Dla dowolnych układów $B_1 \in C_{G_1}^{(n_1)}$ oraz $B_2 \in C_{G_2}^{(n_2)}$ stowarzyszonych z hipergrafami G_1 oraz G_2 zachodzi:*

$$X_{\mathrm{u}}(B_1 \oplus B_2) = \frac{n_1}{n_1+n_2} X_{\mathrm{u}}(B_1) + \frac{n_2}{n_1+n_2} X_{\mathrm{u}}(B_2), \tag{4.45}$$

oraz

$$X_{\max}(B_1 \oplus B_2) = \max\{X_{\max}(B_1), X_{\max}(B_2)\}. \tag{4.46}$$

Dowód.

Zauważmy, że dla dowolnego $\{p(c)\}$ takiego, że dla

$$w := \sum_{c\in E_{G_1}} p(c), \tag{4.47}$$

mamy $0 < w < 1$, oraz dla dowolnych rozkładów $p(\lambda)[V_{G_1}]$ oraz $p(\lambda)[V_{G_2}]$, możemy

zapisać:

$$\sum_{c\in E_{G_1}\cup E_{G_2}} p(c)D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_1}]p(\lambda)[V_{G_2}]}$$
$$= w\sum_{c\in E_{G_1}} \frac{p(c)}{w}D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_1}]}$$
$$+(1-w)\sum_{c\in E_{G_2}} \frac{p(c)}{1-w}D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_2}]}. \quad (4.48)$$

Stąd też:

$$\min_{p(\lambda)[V_{G_1}]p(\lambda)[V_{G_2}]} \sum_{c\in E_{G_1}\cup E_{G_2}} p(c)D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_1}]p(\lambda)[V_{G_2}]}$$
$$= w\min_{p(\lambda)[V_{G_1}]} \sum_{c\in E_{G_1}} \frac{p(c)}{w}D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_1}]}$$
$$+(1-w)\min_{p(\lambda)[V_{G_2}]} \sum_{c\in E_{G_2}} \frac{p(c)}{1-w}D(g(\lambda_c)||p(\lambda_c))\Big|_{p(\lambda)[V_{G_2}]}$$
$$\equiv w X_{\{\frac{p(c)}{w}\}_{c\in E_{G_1}}}(B_1)+(1-w)X_{\{\frac{p(c)}{1-w}\}_{c\in E_{G_2}}}(B_2). \quad (4.49)$$

Przyjmując teraz rozkład jednorodny $p(c)=\frac{1}{n_1+n_2}$ dostajemy

$$w=\sum_{c\in E_{G_1}} \frac{1}{n_1+n_2}=\frac{n_1}{n_1+n_2}, \quad (4.50)$$

po czym na mocy Lematu 4:

$$X_{\mathrm{u}}(B_1\oplus B_2)$$
$$=\frac{n_1}{n_1+n_2}[\min_{p(\lambda)[V_{G_1}]} \sum_{c\in E_{G_1}} \frac{1}{n_1}D(g(\lambda_c)||p(\lambda_c))]+\frac{n_2}{n_1+n_2}[\min_{p(\lambda)[V_{G_2}]} \sum_{c\in E_{G_2}} \frac{1}{n_2}D(g(\lambda_c)||p(\lambda_c))]$$
$$\equiv \frac{n_1}{n_1+n_2}X_{\mathrm{u}}(B_1)+\frac{n_2}{n_1+n_2}X_{\mathrm{u}}(B_2), \quad (4.51)$$

co dowodzi (4.45).

Przejdźmy teraz do wykazania zależności (4.46). Bez straty ogólności możemy założyć, że

$$X_{\max}(B_1)\geq X_{\max}(B_2). \quad (4.52)$$

Z definicji miary $X_{\max}$ mamy, że dla dowolnego $\delta_k>0$ istnieje taki rozkład $\{p_k(c)\}$, że

zachodzi

$$X_{\max}(B_1 \oplus B_2) \leq X_{\{p_k(c)\}}(B_1 \oplus B_2) + \delta_k. \tag{4.53}$$

Poniżej wykażemy, że dla dowolnego k zachodzi

$$X_{\max}(B_1 \oplus B_2) \leq \max\{X_{\max}(B_1), X_{\max}(B_2)\} + \delta_k, \tag{4.54}$$

co dla malejącego δ_k przy $k \to \infty$ dowodzi

$$X_{\max}(B_1 \oplus B_2) \leq \max\{X_{\max}(B_1), X_{\max}(B_2)\}. \tag{4.55}$$

Jak pokażemy, równość (4.46) dostajemy przyjąwszy taki ciąg rozkładów $\{p_k(c)\}$, że $w_k := \sum_{c \in E_{G_1}} p_k(c) = 1$ i dla którego osiągane jest supremum w wyrażeniu $X_{\max}(B_1)$ w granicy dużych k.

Aby udowodnić (4.55) musimy rozpatrzyć dwa przypadki: niech w pierwszej kolejności $\{p_k(c)\}$ będzie takim ciągiem rozkładów, że $w_k = 1$. Wtedy mamy:

$$\begin{aligned} X_{\max}(B_1 \oplus B_2) &\leq X_{\{p_k(c)\}}(B_1 \oplus B_2) + \delta_k \\ &\leq X_{\{p_k(c)\}}(B_1) + \delta_k \\ &\leq X_{\max}(B_1) + \delta_k \\ &\leq \max\{X_{\max}(B_1), X_{\max}(B_2)\} + \delta_k, \end{aligned} \tag{4.56}$$

co też chcieliśmy wykazać.

Rozpatrzmy teraz $\{p_k(c)\}$, dla których $0 < w_k < 1$. W tym przypadku możemy przyjąć, że $\{\frac{p_k(c)}{w_k}\}_{c \in E_{G_1}}$ oraz $\{\frac{p_k(c)}{1-w_k}\}_{c \in E_{G_2}}$ są dobrze zdefiniowanymi rozkładami, dla których możemy napisać:

$$\begin{aligned} X_{\max}(B_1 \oplus B_2) &\leq X_{\{p_k(c)\}}(B_1 \oplus B_2) + \delta_k \\ &= w_k X_{\{\frac{p_k(c)}{w_k}\}_{c \in E_{G_1}}}(B_1) + (1 - w_k) X_{\{\frac{p_k(c)}{1-w_k}\}_{c \in E_{G_2}}}(B_2) + \delta_k \\ &\leq w_k X_{\max}(B_1) + (1 - w_k) X_{\max}(B_2) + \delta_k \\ &\leq X_{\max}(B_1) + \delta_k \\ &\leq \max\{X_{\max}(B_1), X_{\max}(B_2)\} + \delta_k, \end{aligned} \tag{4.57}$$

gdzie równość w powyższym wynika z (4.49), z kolei w następującej po niej nierówności wykorzystaliśmy własność miary $X_{\max}$.■

Powyższe Twierdzenie można uogólnić na przypadek dowolnej skończonej sumy prostej układów, dostając w rezultacie:

$$X_{\mathrm{u}}(\bigoplus_i B_i) = \sum_i n_i X_{\mathrm{u}}(B_i) / \sum_i n_i, \tag{4.58}$$

oraz

$$X_{\max}(\bigoplus_i B_i) = \max_i \{X_{\max}(B_i)\}. \tag{4.59}$$

O ile rezultat dotyczący miary X_{u} wydaje się w pełni intuicyjny, to miara $X_{\max}$ dla sumy prostej stwierdza, iż lepiej „faworyzować" wszystkie konteksty z układu, dla którego $X_{\max}$ jest największe, niż najbardziej kontekstualne konteksty układów o mniejszej wartości miary $X_{\max}$.

Z Twierdzenia 2 wynika również:

Wniosek. Dla dowolnych układów $B_1 \in C_{G_1}^{(n_1)}$ oraz $B_2 \in C_{G_2}^{(n_2)}$ stowarzyszonych z hipergrafami G_1 oraz G_2, dla których $X_{\mathrm{u}}(B_1) \neq X_{\mathrm{u}}(B_2)$, zachodzi:

$$X_{\mathrm{u}}(B_1 \oplus B_2) < X_{\max}(B_1 \oplus B_2). \tag{4.60}$$

Dowód.

Załóżmy, że $X_{\mathrm{u}}(B_1) > X_{\mathrm{u}}(B_2)$. Na mocy Twierdzenia 2 otrzymujemy:

$$\begin{aligned} X_{\mathrm{u}}(B_1 \oplus B_2) &= \frac{n_1}{n_1+n_2} X_{\mathrm{u}}(B_1) + \frac{n_2}{n_1+n_2} X_{\mathrm{u}}(B_2) \\ &< X_{\mathrm{u}}(B_1) \\ &\leq X_{\max}(B_1) \\ &\leq \max\{X_{\max}(B_1), X_{\max}(B_2)\} \\ &= X_{\max}(B_1 \oplus B_2), \end{aligned} \tag{4.61}$$

co dowodzi (4.60).■

4.4. Wartości $X_{\max}$ dla wybranych układów

W bieżącym podrozdziale zajmiemy się wyznaczeniem wartości miary X_{u} dla wybranych układów binarnych, takich jak układy PR, PM, M, jak również układ K. Wymienione układy charakteryzują się posiadaniem wielu symetrii, z których możemy skorzystać obliczając dla nich odpowiednią miarę kontekstualności. W dalszej części omówiony zostanie formalizm, dzięki któremu możliwe będzie dokładne scharakteryzowane przekształceń symetrii. Ponadto zostanie opisane pewne odwzorowanie układów w inne układy oparte na tych przekształceniach symetrii. To odwzorowanie pozwoli na wyodrębnienie kilkuparametrowych rodzin rozkładów prawdopodobieństwa, do których wystarczy się ograniczyć minimalizując względną entropię w definicji miary X_{u}.

4.4.1 Symetryzacja i układy izotropowe

Rozważmy hipergraf G z liczbą kontekstów wynoszącą n oraz układ $B \in C_G^{(n)}$. *Automorfizmem zachowującym niekontekstualność układu B* nazywamy każdą operację l zachowującą niekontekstualność i spełniającą $l(B) = B$.

Niech $\mathscr{L}$ będzie skończonym zbiorem automorfizmów zachowujących niekontekstualność. Oznaczmy przez $\mathscr{G}_{\mathscr{L}}$ grupę generowaną przez $\mathscr{L}$. Odwzorowanie zdefiniowane na układach B jako:

$$B \xrightarrow{\tau_B^{\mathscr{L}}} \sum_{l \in \mathscr{G}_{\mathscr{L}}} \frac{1}{|\mathscr{G}_{\mathscr{L}}|} l(B), \tag{4.62}$$

będziemy nazywać $B - \mathscr{L}$*-symetryzacją* i oznaczymy przez $\tau_B^{\mathscr{L}}$.

Zbiorem *układów* $B - \mathscr{L}$*-izotropowych* nazwiemy obraz wszystkich układów poddanych $B - \mathscr{L}$-symetryzacji:

$$\mathscr{I}_B^{\mathscr{L}} := \{D \in C_G^{(n)} : \exists_{F \in C_G^{(n)}} D = \tau_B^{\mathscr{L}}(F)\}. \tag{4.63}$$

W zależności od zbioru automorfizmów $\mathscr{L}$ generujących grupę $\mathscr{G}_{\mathscr{L}}$ możemy mieć do czynienia z różnymi symetryzacjami. Dla wygody, jeśli rozważania prowadzone będą dla pewnego ustalonego zbioru $\mathscr{L}$, będziemy używać skróconego nazewnictwa i oznaczeń: B-symetryzacja (τ_B) oraz zbiór B-izotropowych układów ($\mathscr{I}_B$).

Poniżej zobaczymy, że zbiór B-izotropowych układów $\mathscr{I}_B$ jest równy zbiorowi układów niezmienniczych względem elementów zbioru $\mathscr{L}$.

Twierdzenie 3. *Niech $C_G^{(n)}$ będzie zbiorem zgodnych układów stowarzyszonych z hipergrafem G. Niech dalej $\mathscr{F}$ będzie skończoną grupą liniowych odwzorowań $f : C_G^{(n)} \to C_G^{(n)}$, a $\mathscr{H} = \{h_1, ..., h_n\} \subseteq \mathscr{F}$ podzbiorem jej elementów takich, że dla każdego z nich jego odwrotność h_i^{-1} należy do $\mathscr{F}$. Oznaczmy przez $\mathscr{T}$ rodzinę układów B niezmienniczych ze względu na transformacje h_i:*

$$\mathscr{T} := \{B \in C_G^{(n)} : \forall_i h_i(B) = B\}, \tag{4.64}$$

oraz podgrupę $\mathscr{F}_{\mathscr{H}} \subseteq \mathscr{F}$ generowaną przez $\mathscr{H}$.

Zachodzi:

$$\mathrm{Im}_{\mathscr{F}_{\mathscr{H}}}(C_G^{(n)}) := \{D \in C_G^{(n)} : \exists_{B \in C_G^{(n)}} \sum_{f \in \mathscr{F}_{\mathscr{H}}} \frac{1}{|\mathscr{F}_{\mathscr{H}}|} f(B) = D\} = \mathscr{T}. \tag{4.65}$$

Dowód.

Niech $B \in \mathrm{Im}_{\mathscr{F}_{\mathscr{H}}}(C_G^{(n)})$. Wtedy dla każdego i mamy:

$$h_i(B) = h_i\left(\sum_{f\in\mathscr{F}_{\mathscr{H}}} \frac{1}{|\mathscr{F}_{\mathscr{H}}|} f(B)\right) \tag{4.66}$$

$$= \sum_{f\in\mathscr{F}_{\mathscr{H}}} \frac{1}{|\mathscr{F}_{\mathscr{H}}|} h_i \circ f(B) \tag{4.67}$$

$$= \sum_{\tilde{f}=h_i\circ f\in\mathscr{F}_{\mathscr{H}}} \frac{1}{|\mathscr{F}_{\mathscr{H}}|} \tilde{f}(B) = B, \tag{4.68}$$

gdzie w pierwszym równaniu skorzystaliśmy z liniowości odwzorowań h natomiast w ostatnim skorzystaliśmy z faktu, że sumowanie po $\tilde{f}$ przebiega po całej grupie $\mathscr{F}_{\mathscr{H}}$ (niezależnie w jakiej kolejności wybierzemy jej elementy), ponieważ każdy element h_i ma odwrotność w tej grupie. Z powyższego otrzymujemy, że $B \in \mathscr{T}$, a zatem również $\mathrm{Im}_{\mathscr{F}_{\mathscr{H}}}(C_G^{(n)}) \subseteq \mathscr{T}$.

Z drugiej strony, dla każdego układu $B \in \mathscr{T}$ mamy:

$$B = \sum_{f\in\mathscr{F}_{\mathscr{H}}} \frac{1}{|\mathscr{F}_{\mathscr{H}}|} f(B), \tag{4.69}$$

ponieważ dla dowolnego $f \in \mathscr{F}_{\mathscr{H}}$ każdy element jest złożeniem $f = h_{i_1} \circ \ldots \circ h_{i_n}$ ($h_{i_k} \in \mathscr{H}$), a zatem $f(B) = B$, i stąd też mamy bezpośrednio (4.69). Tym samym pokazaliśmy, że $\mathscr{T} \subseteq \mathrm{Im}_{\mathscr{F}_{\mathscr{H}}}(C_G^{(n)})$ co, łącznie z przeciwną inkluzją dowodzi tezy twierdzenia.■

Rozważmy teraz szczególny przypadek układów binarnych i pewien wybrany zbiór automorfizmów zachowujących niekontekstualność $\mathscr{L}_0$, który będzie składał się z dwóch rodzajów liniowych odwzorowań:

1. π_i - permutacja pewnych obserwabli, co odpowiada zamianie kontekstów,
2. b_i - negacja pewnych wyników obserwabli, tzw. bit-flipów (dla układów binarnych $0 \to 1$ oraz $1 \to 0$).

W przypadku tak wybranych układów i zbioru $\mathscr{L}_0$ udowodnimy twierdzenie, które w znacznym stopniu ułatwia wyznaczenie wartości miary $X_{\mathrm{u}}(B)$.

Twierdzenie 4. *Dla dowolnych układów $B \in C_G^{(n)}$ oraz zbioru $B - \mathscr{L}_0$-izotropowych układów $\mathscr{I}_B^{\mathscr{L}_0}$ mamy:*

$$X_{\mathrm{u}}(B) = \min_{p(\lambda)\in\mathscr{I}_B^{\mathscr{L}_0}} \sum_c \frac{1}{n} D(g(\lambda_c)||p(\lambda_c)), \tag{4.70}$$

gdzie minimum wzięte jest jedynie po takich łącznych rozkładach prawdopodobieństw $p(\lambda)$, z których otrzymać można niekontekstualne układy ze zbioru $B-\mathscr{L}_0$-izotropowych układów $\mathscr{I}_B^{\mathscr{L}_0}$.

Dowód.

Niech $p(\lambda)$ będzie rozkładem prawdopodobieństwa, który jest optymalny dla miary $X_{\mathrm{u}}(B)$, z kolei niekontekstualny układ otrzymany z tego rozkładu oznaczmy przez B_{nc}. Zauważmy, że ustaliwszy zbiór automorfizmów jako ustalony uprzednio $\mathscr{L}_0$, dla dowolnego elementu f w grupie $\mathscr{G}_{\mathscr{L}_0}$, mamy:

$$X_{\mathrm{u}}(B) = \sum_c \frac{1}{n} D(g_f(\lambda_c)||p_f(\lambda_c)), \tag{4.71}$$

gdzie $g_f(\lambda_c)$ oraz $p_f(\lambda_c)$ są rozkładami prawdopodobieństw dla kontekstów c układów odpowiednio $f(B)$ oraz $f(B_{nc})$. Aby to wykazać przypomnijmy, że z definicji zbioru $\mathscr{L}_0$, dowolny jego element f jest złożeniem permutacji obserwabli π_i oraz negacji pewnych wyników obserwabli b_i. Wystarczy zatem dowieść, że powyższa równość zachodzi dla każdej z tej operacji z osobna - wtedy równość ta będzie prawdziwa dla dowolnego złożenia $\pi_i \circ b_j$. Niech w pierwszej kolejności f będzie pewną permutacją obserwabli π_i. Z kolei jeśli $f(B) = B$ to f jest automorfizmem hipergrafu G, z którym stowarzyszony jest układ B, a zatem jest ona tożsama z permutacją obserwabli, która dokonuje permutacji kontekstów. Jeśli tak, to działanie f zmienia jedynie kolejność elementów $D(g(\lambda_c)||p(\lambda_c))$ w sumowaniu po wszystkich kontekstach w definicji miary X_{u}. Z drugiej zaś strony, niech f będzie teraz negacją wybranych wyników obserwabli b_i. Wtedy, dla negacji wyniku wybranej (wybranych) obserwabli zastosowanej równocześnie w $g(\lambda_c)$ i $p(\lambda_c)$, względna entropia nie ulegnie zmianie, jako że jest ona niezmiennicza ze względu na obustronnie zastosowaną odwracalną operację. Możemy zatem zapisać:

$$\begin{aligned} X_{\mathrm{u}}(B) &= \sum_c \frac{1}{n} \sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} D(g_f(\lambda_c)||p_f(\lambda_c)) \\ &\geq \sum_c \frac{1}{n} D\left(\sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} g_f(\lambda_c) || \sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} p_f(\lambda_c) \right), \end{aligned} \tag{4.72}$$

gdzie w drugiej linii skorzystaliśmy z wypukłości względnej entropii [94].

Z powyższego widzimy, że symetryzacja nie zwiększa wartości względnej entropii. Ponadto, jeśli f jest automorfizmem układu B, to dla dowolnego kontekstu c mamy:

$$\sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} g_f(\lambda_c) = g(\lambda_c). \tag{4.73}$$

Teraz, jeśli $B - \mathscr{L}_0$-symetryzacja jest odwzorowaniem zachowującym niekontekstualność, to wyrażenie

$$\sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} p_f(\lambda_c) \tag{4.74}$$

jest z definicji rozkładem prawdopodobieństwa dla odpowiedniego kontekstu układu $\tau_B^{\mathscr{L}_0}(B_{nc}) \in \mathscr{I}_B^{\mathscr{L}_0}$, który jest niekontekstualny, jako że $B-\mathscr{L}_0$-symetryzacja przekształca układy niekontekstualne w niekontekstualne.

Z kolei jeśli układ B_{nc} był układem optymalnym dla miary $X_{\mathrm{u}}(B)$, natomiast podstawienie układu $\tau_B^{\mathscr{L}_0}(B_{nc})$ w miejsce B_{nc} może co najwyżej zachować lub zmniejszyć wartość $X_{\mathrm{u}}(B)$ (ze względu na nierówność (4.72)), to również układ $\tau_B^{\mathscr{L}_0}(B_{nc})$ musi być układem optymalnym dla miary $X_{\mathrm{u}}(B)$. Dzięki temu, w miejsce nierówności (4.72) otrzymujemy równość. ■

4.4.2 Zbiory automorfizmów dla wybranych układów

W dalszej części wyszczególnimy zbiory $\mathscr{L}_0$ operacji zachowujących niekontekstualność dla układów binarnych PM, M oraz $CH_{(n)}$. Mając określony zbiór automorfizmów $\mathscr{L}_0$ wciąż nie jest jednak łatwe znalezienie rodziny układów, która jest obrazem działania grupy $\mathscr{G}_{\mathscr{L}_0}$, jednak dzięki Twierdzeniu 3 obraz ten dostajemy wymuszając niezmienniczość układów ze względu na kilka określonych symetrii. Dzięki takiemu wyszczególnieniu układów izotropowych będziemy mogli wykorzystać Twierdzenie 4 do wyznaczenia miary X_{u} dla wymienionych układów.

W pierwszej kolejności zastosujmy symetryzację do układu PM. Niech $\mathscr{L}_0 = \{h_1, ..., h_8\}$, gdzie elementy h_i są w ogólności złożeniami operacji π_i oraz b_i, takich że:

(i) $h_1, ..., h_6$ - są wszystkimi 3! permutacjami kontekstów $\{A_1, A_2, A_3\}$, $\{A_4, A_5, A_6\}$ oraz $\{A_7, A_8, A_9\}$ (dla hipergrafu G_{PM} z Rys. 4.3 są to permutacje krawędzi poziomych);

(ii) h_7 - permutacja kontekstów $\{A_1, A_4, A_7\}$ oraz $\{A_2, A_5, A_8\}$ (dla hipergrafu G_{PM} permutacja krawędzi pionowych ciągłych);

(iii) h_8 - złożenie permutacji kontekstów zdefiniowanej jako jednoczesna zamiana obserwabli $A_4 \longleftrightarrow A_2$, $A_7 \longleftrightarrow A_3$, $A_7 \longleftrightarrow A_6$ (pozostałe obserwable przechodzą na siebie), oraz negacji wyniku obserwabli A_9 (dla hipergrafu G_{PM} symetria ta jest odbiciem względem diagonali z negacją wyniku odpowiedniej obserwabli, dzięki czemu układ krawędzi ciągłych i krawędzi przerywanej pozostaje niezmieniony).

Taki wybór zbioru $\mathscr{L}_0$ prowadzi do jednoparametrowej rodziny izotropowych układów PM (o czym mówi poniższy lemat), którą oznaczać będziemy przez $\mathscr{I}_{PM}^{\mathscr{L}_0}$.

Lemat 5. *Zachodzi:*

$$\mathscr{I}_{PM}^{\mathscr{L}_0} = \{\alpha PM + (1-\alpha)PM' | \alpha \in [0,1]\}, \tag{4.75}$$

gdzie PM' jest układem przeciwnym do układu PM.

Dowód.

Zgodnie z Twierdzeniem 3 wystarczy wykazać niezmienniczość układu PM ze względu na wybrany zbiór $\mathscr{L}_0$, co pociąga za sobą, że układ PM należy do rodziny $\mathscr{I}_{PM}^{\mathscr{L}_0}$. Celem wyszczególniania poszczególnych kontekstów odniesiemy się do hipergrafu G_{PM} z Rys. 4.3.

Zauważmy, że: po pierwsze, niezmienniczość ze względu na operacje $h_1,...,h_6$ wszystkie konteksty poziome muszą być opisane przez ten sam rozkład prawdopodobieństwa. Po drugie, niezmienniczość ze względu na operację h_8 wymaga, by środkowy kontekst pionowy oraz środkowy kontekst poziomy miały ten sam rozkład prawdopodobieństwa. Po trzecie, niezmienniczość ze względu na operację h_7 pociąga za sobą, iż wszystkie konteksty oznaczone przez krawędzie ciągłe muszą mieć ten sam rozkład prawdopodobieństwa.

Teraz, dla krawędzi ciągłych rozkład prawdopodobieństwa jest wyznaczony jednoznacznie przez 8 prawdopodobieństw $q(ijk)$, gdzie i,j,k są binarnymi wynikami pomiaru obserwabli na poszczególnych kontekstach. Niezmienniczość ze względu na operacje $h_1,...,h_6$ sprawia, że dla krawędzi ciągłych rozkład prawdopodobieństwa jest niezmienniczy ze względu na permutacje obserwabli, tzn.:

$$q(001) = q(010) = q(100), \tag{4.76}$$
$$q(011) = q(101) = q(110). \tag{4.77}$$

Rozkład prawdopodobieństwa jest zatem opisany jedynie przez 4 niezależne parametry $q(000)$, $q(001)$, $q(011)$ oraz $q(111)$ (analogicznie, dla krawędzi przerywanej: $r(000)$, $r(001)$, $r(011)$ oraz $r(111)$). Dalej, niezmienniczość ze względu na operację h_8 wymaga:

$$q(000) = r(001), \tag{4.78}$$
$$q(011) = r(010). \tag{4.79}$$

Ponieważ $r(001) = r(010)$, to również prawdopodobieństwa $q(000)$ i $q(011)$ muszą być sobie równe. Podobnie:

$$q(111) = r(110), \tag{4.80}$$
$$q(010) = r(011), \tag{4.81}$$

co, ze względu na $r(110) = r(011)$ pociąga równość prawdopodobieństw $q(111)$ i $q(010)$. Widzimy więc, że zachodzi równość dla prawdopodobieństw o określonej parzystości wyników obserwabli. Tym samym można przyjąć, że $q(000) = \alpha/4$ dla pewnego parametru $\alpha \in [0,1]$, natomiast $q(111) = (1-\alpha)/4$. Analogicznie, dla krawędzi przerywanej mamy: $r(111) = \alpha/4$ oraz $r(000) = (1-\alpha)/4$.■

Dla układu M schemat sprowadzania go do jednoparametrowej rodziny pudeł

izotropowych jest analogiczny jak dla układu PM. Niech w tym przypadku $\mathscr{L}_0 = \{\tilde{h}_1, ..., \tilde{h}_{10}\}$, gdzie elementy $\tilde{h}_i$ są w ogólności złożeniami operacji π_i oraz b_i, wyszczególnionych poniżej:

(i) $\tilde{h}_1$ - symetryczne odbicie gwiazdy (tj. przemieszczenia obserwabli na odpowiednich węzłach grafu G_M) względem osi symetrii Aa przechodzącej przez węzły A oraz a;

(ii) $\tilde{h}_2$ - odbicie względem osi symetrii Bb oraz negacja wyniku obserwabli B;

(iii) $\tilde{h}_3$ - odbicie względem osi symetrii Cc oraz negacja wyniku obserwabli c;

(iv) $\tilde{h}_4$ - odbicie względem osi symetrii Dd oraz negacja wyniku obserwabli d;

(v) $\tilde{h}_5$ - odbicie względem osi symetrii Ee oraz negacja wyniku obserwabli E;

(vi) $\tilde{h}_{6-10}$ - negacja wyników trzech obserwabli tworzących trójkąty odpowiednio: Acd, Bde, Cea, Dab oraz Ebc.

Lemat 6. *Zachodzi:*

$$\mathscr{I}_M^{\mathscr{L}_0} = \{\alpha M + (1-\alpha)M' | \alpha \in [0,1]\}, \tag{4.82}$$

gdzie M' jest układem przeciwnym do układu M.

Dowód.

Dowód jest analogiczny jak w przypadku układu PM.■

Dla układów łańcuchowych $CH_{(n)}$ rozumowanie przedstawia się podobnie. Niech w tym przypadku $\mathscr{L}_0 = \{\hat{h}_1, ..., \hat{h}_j, ..., \hat{h}_{n-1}\}$, gdzie elementy $\hat{h}_j$ są złożeniami cyklicznych permutacji kontekstów takich, że wszystkie pary obserwabli $\{A_i, A_{i+1}\}$ przechodzą na $\{A_{i+j}, A_{i+j+1}\}$, wraz z negacją wyników dla obserwabli $A_1, ..., A_j$.

Taki wybór zbioru $\mathscr{L}_0$ prowadzi do jednoparametrowej rodziny izotropowych układów łańcuchowych $CH_{(n)}$ (o czym mówi nam poniższy lemat), którą oznaczać będziemy przez $\mathscr{I}_{CH_{(n)}}^{\mathscr{L}_0}$.

Lemat 7. *Zachodzi:*

$$\mathscr{I}_{CH_{(n)}}^{\mathscr{L}_0} = \{\alpha CH_{(n)} + (1-\alpha)CH'_{(n)} | \alpha \in [0,1]\}, \tag{4.83}$$

gdzie $CH'_{(n)}$ jest układem przeciwnym do układu $CH_{(n)}$.

Dowód.

Dowód opiera się na rozumowaniu analogicznym jak w przypadku układów PM oraz M.■

W przypadku układu K niech $\mathscr{L}_0 = \{\check{h}_1, ..., \check{h}_{10}\}$ będzie grupą diedralną $\mathbb{D}_5$, gdzie elementy $\check{h}_i$ są w ogólności operacjami π_i takimi, że:

(i) $\check{h}_1, ..., \check{h}_5$ - są cyklicznymi permutacjami operatorów rzutowych $\Pi_i \to \Pi_{i+k}$, gdzie $k \in \{1, ..., 5\}$;

(ii) $\check{h}_6, ..., \check{h}_{10}$ - są cyklicznymi permutacjami operatorów rzutowych $\Pi_i \to \Pi_{i+k}$, gdzie $k \in \{1, ..., 5\}$ z dodatkową zamianą operatorów rzutowych $\Pi_1 \longleftrightarrow \Pi_5$ oraz $\Pi_2 \longleftrightarrow \Pi_4$.

4.4.3 Wyznaczenie miary X_{u} dla wybranych układów

Dalsze lematy i twierdzenia, o ile nie zostanie zaznaczone inaczej, dotyczyć będą układów PM, M oraz $CH_{(n)}$; rozważania odnoszące się do układu K przedstawione zostaną odrębnie. Dla wygody zastosujmy teraz oznaczenie $B_\alpha \in \mathscr{I} \equiv \mathscr{I}_{PM}^{\mathscr{L}_0} \cup \mathscr{I}_{M}^{\mathscr{L}_0} \cup \mathscr{I}_{CH_{(n)}}^{\mathscr{L}_0}$ dla dowolnego układu należącego do klasy jednoparametrowych układów izotropowych rozważanych powyżej. Zgodnie z Twierdzeniem 4 wiemy, że celem wyznaczenia miary kontekstualności X_{u} wybranego układu B_{α_0} musimy zminimalizować wielkość $\sum_{c \in E_G} \frac{1}{n} D(g(\lambda_c)||p(\lambda_c))$ po wszystkich rozkładach prawdopodobieństwa $p(\lambda)$, z których skonstruować można niekontekstualne układy B_α. Konieczna jest zatem znajomość przedziału wartości parametru α, dla którego układy B_α z danej klasy są układami niekontekstualnymi. Poniżej wprowadzimy pomocniczą *miarę zgodności układów*, dzięki której będziemy mogli wyprowadzić nierówności kontekstualne ze względu na parametr α.

Wprowadźmy notację wektorową charakteryzującą dany układ. Niech $|B\rangle$ oznacza wektor prawdopodobieństw danego układu B. Przykładowo, dla układu $CH_{(4)} = PR$ mającego 4 różne konteksty mamy:

$$\begin{aligned} |PR\rangle = & \left(g(\lambda_{c_1}) \middle| g(\lambda_{c_2}) \middle| g(\lambda_{c_3}) \middle| g(\lambda_{c_4}) \right) \\ = & \left(\tfrac{1}{2}\,0\,0\,\tfrac{1}{2} \middle| \tfrac{1}{2}\,0\,0\,\tfrac{1}{2} \middle| \tfrac{1}{2}\,0\,0\,\tfrac{1}{2} \middle| 0\,\tfrac{1}{2}\,\tfrac{1}{2}\,0 \right) \\ = & \tfrac{1}{2} \left(1\,0\,0\,1 \middle| 1\,0\,0\,1 \middle| 1\,0\,0\,1 \middle| 0\,1\,1\,0 \right). \end{aligned} \tag{4.84}$$

Oznaczmy ponadto przez $g(\lambda_c)_i$ prawdopodobieństwo i-tego wyniku pomiaru dla układu B w obrębie kontekstu c w rozkładzie prawdopodobieństwa $g(\lambda_c)$. Zgodnie z powyższym przykładem, dla układu PR mamy: $g(\lambda_{c_3})_2 = \frac{1}{2}$ jako prawdopodobieństwo otrzymania wyniku $(1, 1)$.

Określmy teraz pewien układ $B = \{g(\lambda_c)\}$ stowarzyszony z hipergrafem G oraz inny układ $\tilde{B} = \{\tilde{g}(\lambda_c)\}$ stowarzyszony z tym samym hipergrafem. Dla układu B określamy miarę *zgodności układu $\tilde{B}$ względem układu B* poprzez:

$$\beta_B(\tilde{B}) := \sum_c \sum_{i \in \mathrm{supp}(g(\lambda_c))} \tilde{g}(\lambda_c)_i, \tag{4.85}$$

gdzie sumowanie po i ograniczone jest do *nośnika rozkładu* $g(\lambda_c)$ (oznaczonego przez $\mathrm{supp}(g(\lambda_c))$), tj. zbioru tych wyników pomiarów dla kontekstu c, dla których prawdopodobieństwo jest niezerowe.

Wartość miary β_B określonej dla pewnego układu $\tilde{B}$ można rozumieć w kategoriach „średniej wartości" ilości kontekstów, dla których przy ogólnych pomiarach otrzymać można te same wyniki pomiarów co dla układu B, tzn. zgodność wyników pomiarów na kontekstach z pominięciem ich statystyki.

Przykład. Rozpatrzmy układ $\hat{B}$, dany przez następujący wektor:

$$|\hat{B}\rangle = \left(\tfrac{1}{4}\,0\,0\,\tfrac{3}{4} \middle| \tfrac{3}{4}\,0\,0\,\tfrac{1}{4} \middle| \tfrac{3}{4}\,0\,0\,\tfrac{1}{4} \middle| \tfrac{1}{4}\,0\,0\,\tfrac{3}{4} \right). \tag{4.86}$$

Łatwo sprawdzić, że układ $\hat{B}$ jest układem zgodnym. Z kolei zgodnie z definicją miary β liczonej względem układu PR, dostajemy:

$$\beta_{PR}(\hat{B}) = 3. \tag{4.87}$$

Lemat 8. *Dla dowolnego układu $\tilde{B} \in C_G^{(n)}$ oraz układu binarnego $B \in C_G^{(n)}$ wybranego jako układ referencyjny w mierze β_B, określonych dla kontekstów składających się z m dychotomicznych obserwabli, zachodzi:*

$$\beta_B(\tilde{B}) = 2^{m-1}\langle \tilde{B}|B\rangle, \tag{4.88}$$

gdzie $\langle .|.\rangle$ oznacza iloczyn skalarny dwóch wektorów. Ponadto, dla określonej symetryzacji $\tau_B^{\mathscr{L}_0}$ zachodzi:

$$\beta_B(\tilde{B}) = \beta_B(\tau_B^{\mathscr{L}_0}(\tilde{B})). \tag{4.89}$$

Dowód.

Zauważmy, że dla układów binarnych jedynymi występującymi w nich rozkładami prawdopodobieństwa są (4.11) oraz (4.12), a zatem wektor $2^{m-1}|B\rangle$ będzie wektorem jedynek w miejscach wyników, dla których prawdopodobieństwa ich otrzymania są niezerowe. Tym samym iloczyn skalarny sumuje wszystkie prawdopodobieństwa z nośnika układu B.

Własność (4.89) dowodzimy wykazując następujący ciąg równości:

$$\begin{aligned}\beta_B(\tau_B^{\mathscr{L}_0}(\tilde{B})) &= 2^{m-1}\langle \sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} f(\tilde{B})|B\rangle \\ &= \sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} 2^{m-1}\langle f(\tilde{B})|B\rangle \\ &= \sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} 2^{m-1}\langle f(\tilde{B})|f(B)\rangle \\ &= \sum_{f\in\mathscr{G}_{\mathscr{L}_0}} \frac{1}{|\mathscr{G}_{\mathscr{L}_0}|} 2^{m-1}\langle \tilde{B}|B\rangle \\ &= \beta_B(\tilde{B}),\end{aligned} \tag{4.90}$$

gdzie druga równość wynika z liniowości iloczynu skalarnego, trzecia równość jest konsekwencją faktu, że dla wybranej symetryzacji operacja f jest automorfizmem układu B, z kolei w czwartej równości wykorzystaliśmy fakt, że dowolne operacje f, będące w ogólności złożeniami permutacji kontekstów i negacją wyników wybranych obserwabli, działają jednakowo w obrębie obu wektorów prawdopodobieństw, co nie wpływa na wartość iloczynu skalarnego.■

Zdefiniowawszy miarę zgodności układów, w dalszym toku udowodnimy twierdzenia pozwalające wyznaczyć zakres parametru α, dla którego jednoparametrowe układy izotropowe B_α są układami niekontekstualnymi.

Twierdzenie 5. *Dla dowolnego układu binarnego $B \in C_G^{(n)}$ z jednym kontekstem mającym rozkład P_{nparz}, takiego, że każdy wierzchołek z V_G należy do parzystej liczby krawędzi z E_G, i dla niekontekstualnego układu $\tilde{B} \in C_G^{(n)}$ zachodzi:*

$$\beta_B(\tilde{B}) \leq n-1. \tag{4.91}$$

Ponadto, powyższa nierówność jest wysycana.

Dowód.

Celem wykazania kontekstualności układu M zastosujemy argument użyty w pracy [8].

Zgodnie z Lematem 1, układ niekontekstualny przedstawić można jako probabilistyczną mieszaninę układów deterministycznych, tj. takich układów, dla których wyniki poszczególnych pomiarów są *a priori* ściśle określone. Z kolei z lematu 8 wiemy, że miara $\beta_B(\tilde{B})$ jest liniowa, a zatem w dowodzie możemy ograniczyć się do układów deterministycznych. Ponieważ dla układów deterministycznych $\tilde{B}$ poszczególne wyniki ogólnych pomiarów zachodzą z prawdopodobieństwem równym 1 lub 0, wartość miary $\beta_B(\tilde{B})$ przyjmować może jedynie całkowite wartości. Załóżmy zatem, że $\beta_B(\tilde{B}) = n$, tj. wszystkie warunki zachowania parzystości dla układu kontekstual-

nego są spełnione. Wtedy dla $n-1$ kontekstów suma wszystkich wyników pomiarów jest parzysta ($\bigoplus_i a_i = 0$), natomiast dla wyszczególnionego kontekstu mającego rozkład P_{nparz} w układzie B, suma wyników pomiarów jest nieparzysta ($\bigoplus_i a_i = 1$). W konsekwencji suma po wszystkich kontekstach i pomiarach daje wartość nieparzystą ($\bigoplus_{i,c} a_i = 1$). Z drugiej strony, dla układu deterministycznego suma ta przyjmuje wartość parzystą: $\bigoplus_{i,c} a_i = \bigoplus_i n_i a_i = 0$. Wynika to z faktu, że każda obserwabla występuje w parzystej ilości n_i kontekstów. Otrzymujemy zatem sprzeczność, co dowodzi $\beta_B(\tilde{B}) < n$ (z definicji mamy bowiem $\beta_B(\tilde{B}) \leq n$).

Miara $\beta_B(\tilde{B})$ może jednak osiągnąć wartość $n-1$: wystarczy przypisać wyniki pomiarów dla wszystkich obserwabli równe 0, dzięki czemu wszystkie konteksty będą miały rozkład prawdopodobieństwa P_{parz}, a zatem zachowana zostaje parzystość na $n-1$ kontekstach, co nasyca nierówność.■

Twierdzenie 6. *Dla dowolnego układu binarnego $B \in C_G^{(n)}$ (z n parzystym) z jednym kontekstem mającym rozkład P_{nparz}, takiego, że każdy wierzchołek z V_G należy do parzystej liczby krawędzi z E_G, i dla niekontekstualnego układu $\tilde{B} \in C_G^{(n)}$, zachodzi:*

$$\beta_B(\tilde{B}) \geq 1. \tag{4.92}$$

Ponadto, jeśli liczba wierzchołków w każdym kontekście jest nieparzysta, to powyższa nierówność jest wysycana.

Dowód.

Dowód jest analogiczny do dowodu Twierdzenia 5. Podobnie jak poprzednio, ograniczymy rozważania tylko do układów deterministycznych. Załóżmy zatem, że dla układu $\tilde{B}$ istnieje takie przypisanie *a priori* wyników pomiarów obserwabli, dla których zachodzi $\beta_B(\tilde{B}) = 0$. Wtedy układ $\tilde{B}$ nie spełnia warunków parzystości dla żadnego kontekstu układu B, natomiast spełnia wszystkie warunki parzystości dla układu przeciwnego (również kontekstualnego) B'. Dla tego układu suma wyników obserwabli dla $n-1$ kontekstów jest nieparzysta ($\bigoplus_i a_i = 1$), natomiast dla jednego kontekstu suma wyników obserwabli jest parzysta ($\bigoplus_i a_i = 0$). Tym samym sumowanie po wszystkich kontekstach i pomiarach daje wynik nieparzysty ($\bigoplus_{i,c} a_i = 1$). Znów jednak, w przypadku układu deterministycznego suma ta przyjmuje wartość parzystą: $\bigoplus_{i,c} a_i = \bigoplus_i n_i a_i = 0$. Wynika to z faktu, że każda obserwabla występuje w parzystej ilości n_i kontekstów.

Miara $\beta_B(\tilde{B})$ może jednak w tym szczególnym przypadku osiągnąć wartość 1: wystarczy przypisać wyniki pomiarów dla wszystkich obserwabli równe 1, dzięki czemu ze względu na nieparzystą ilość obserwabli w każdym kontekście będą one miały rozkład prawdopodobieństwa P_{nparz}. Tym samym zgodność parzystości zachowana zostaje tylko na jednym kontekście z układu B, co wysyca nierówność.■

Pokażemy teraz, że dodatkowe założenie poczynione w Twierdzeniu 6 na temat parzystości liczby kontekstów n była konieczna:

Twierdzenie 7. *Dla układów $B \in \{M, CH_{(n)}\}$ z nieparzystą liczbą kontekstów n, istnieje niekontekstualny układ $\tilde{B}$, dla którego:*

$$\beta_B(\tilde{B}) = 0. \tag{4.93}$$

Dowód.

Istnieje deterministyczne przypisanie wartości wynikom obserwabli, dla którego $\beta_B(\tilde{B}) = 0$. Skonstruujmy układ $\tilde{B}$ w następujący sposób: wpierw przypiszmy wynikom wszystkich obserwabli wartości 1, następnie wybierzmy wszystkie obserwable nienależące do kontekstu mającego dla układu B rozkład P_{nparz}, ale tylko takie, które nie przynależą do tych samych kontekstów (przykładowo: dla układu M są to np. obserwable A i a) i dla tych obserwabli zmieńmy wartości na 0. Takie przypisanie wyników prowadzi o powstania układu przeciwnego do B, a zatem mamy z definicji:

$$\beta_B(\tilde{B}) \equiv \beta_B(B') = 0.\blacksquare \tag{4.94}$$

Korzystając z powyższych twierdzeń możemy dla jednoparametrowych układów izotropowych znaleźć ograniczenie na wartość parametru α, dla których układy B_α są układami niekontekstualnymi.

Lemat 9. *Dla niekontekstualnych izotropowych układów $B_\alpha \in \mathscr{I} \equiv \mathscr{I}_{PM}^{\mathscr{L}_0} \cup \mathscr{I}_{M}^{\mathscr{L}_0} \cup \mathscr{I}_{CH_{(n)}}^{\mathscr{L}_0}$ mających n kontekstów, zachodzi:*

$$\alpha \leq \frac{n-1}{n}. \tag{4.95}$$

Dodatkowo dla parzystych n:

$$\alpha \geq \frac{1}{n}, \tag{4.96}$$

a dla nieparzystych n:

$$\alpha \geq 0. \tag{4.97}$$

Ponadto, powyższe nierówności są ciasne.

Dowód.

Z definicji miary β_B wiemy, że dla układu przeciwnego B' zachodzi: $\beta_B(B') = 0$. Z liniowości miary β_B mamy natomiast, że dla układów izotropowych postaci $B_\alpha = \alpha B + (1-\alpha)B'$ zachodzi:

$$\beta_B(B_\alpha) = \alpha\beta_B(B) + (1-\alpha)\beta_B(B') = n\alpha, \tag{4.98}$$

i w konsekwencji dla niekontekstualnych układów izotropowych $B_\alpha \in \mathscr{I}$ na mocy

Twierdzenia 5 dostajemy:

$$\alpha \leq \frac{n-1}{n}. \tag{4.99}$$

Drugą nierówność dowodzimy w sposób analogiczny mając na uwadze, że układy PM oraz $CH_{(n)}$ dla parzystych n spełniają założenia Twierdzenia 6.

Podobnie postępujemy w przypadku trzeciej nierówności, gdzie układy M oraz $CH_{(n)}$ dla nieparzystych n spełniają założenia Twierdzenia 7.

Zobaczymy teraz, że graniczne wartości α są osiągalne przez układy niekontekstualne. W pierwszej kolejności, z Twierdzenia 5 wiemy, że istnieje niekontekstualny układ $\tilde{B}$, dla którego $\beta_B(\tilde{B}) = n-1$. Na mocy Lematu 5, po zastosowaniu odpowiedniej symetryzacji $\tau_B^{\mathscr{L}_0}$ układ $\tilde{B}$ należeć będzie do klasy układów izotropowych $\mathscr{I}_B^{\mathscr{L}_0}$, a tym samym będzie miało ono formę jednoparametrowego układu postaci:

$$B_\alpha = \alpha B + (1-\alpha)B'. \tag{4.100}$$

Na mocy Lematu 8 mamy natomiast:

$$\beta_B(\tilde{B}) = \beta_B(\tau_B^{\mathscr{L}_0}(\tilde{B})) = n-1. \tag{4.101}$$

Z równania (4.98) $\beta_B(\tau_B^{\mathscr{L}_0}(\tilde{B})) = n\alpha$, dzięki czemu układ $\tau_B^{\mathscr{L}_0}(\tilde{B})$ osiąga wartość $\alpha = \frac{n-1}{n}$. Ze względu na fakt, że odwzorowanie $\tau_B^{\mathscr{L}_0}$ jest losowym zastosowaniem operacji zachowujących niekontekstualność, również układ powstały z $\tilde{B}$ po symetryzacji jest układem niekontekstualnym.

Analogicznie dowodzi się osiągania granicznych wartości w nierównościach (4.96) oraz (4.97), korzystając z Twierdzeń 6 oraz 7.■

Poniżej przedstawimy lemat, który dalej uprości wyznaczenie miary X_{u} dla układów izotropowych.

Lemat 10. *Niech układ $B = \{g(\lambda_c)\} \in C_G^{(n)}$ oraz dla dowolnego rozkładu prawdopodobieństwa $p(c)$*

$$X_{\{p(c)\}}(B) = \min_{p(\lambda)\in\mathscr{S}} \sum_c p(c) D(g(\lambda_c)||p(\lambda_c)), \tag{4.102}$$

dla pewnej klasy układów $\mathscr{S} \subseteq NC_G$.

Jeśli dla dowolnego rozkładu $p(\lambda) \in \mathscr{S}$ dla dowolnych kontekstów c istnieje odwracalna operacja $\mathfrak{P}_c$ taka, że

$$\mathfrak{P}_c(p(\lambda_c)) = p(\lambda_{c_0}), \tag{4.103}$$

i równocześnie

$$\mathfrak{P}_c(g(\lambda_c)) = g(\lambda_{c_0}), \tag{4.104}$$

wtedy zachodzi:

$$X_{\{p(c)\}}(B) = \min_{p(\lambda)\in\mathscr{S}} D(g(\lambda_{c_0})||p(\lambda_{c_0})). \tag{4.105}$$

Dowód.

Wystarczy zauważyć, że względna entropia jest niezmiennicza ze względu na jednoczesne zastosowanie odwracalnej operacji dla porównywanych rozkładów prawdopodobieństwa [94]

$$D\left(\mathfrak{P}_c(g(\lambda_c))||\mathfrak{P}_c(p(\lambda_c))\right) = D(g(\lambda_{c_0})||p(\lambda_{c_0})), \tag{4.106}$$

w szczególności, gdy operacją tą jest negacja wyników bądź permutacja obserwabli. ■

Dla przykładu, rozpatrzmy względną entropię dla rozkładów prawdopodobieństwa na kontekście składającym się z dwóch dychotomicznych obserwabli takich, że dla dowolnego wyniku z $\{00, 01, 10, 11\}$ mamy określone pewne prawdopodobieństwo. Niech operacja $\mathfrak{P}_c$ będzie zdefiniowana jako negacja wyników obu obserwabli. Zastosowanie tej operacji do rozkładów $p(\lambda_c)$ oraz $g(\lambda_c)$ zmienia wyłącznie indeksowanie poszczególnych prawdopodobieństw nie zmieniając ich wartości, tym samym zastosowanie $\mathfrak{P}_c$ do $D\left(\mathfrak{P}_c(g(\lambda_c))||\mathfrak{P}_c(p(\lambda_c))\right)$ zmienia jedynie porządek sumowania w definicji względnej entropii.

Korzystając z powyższych Lematów i Twierdzeń przedstawimy poniżej zwięzłe formuły wyrażające wartość miary X_{u} dla wszystkich izotropowych układów omawianych w tym rozdziale.

Twierdzenie 8. *Dla układów izotropowych* $B_\alpha \in \mathscr{I} \equiv \mathscr{I}_{PM}^{\mathscr{L}_0} \cup \mathscr{I}_{M}^{\mathscr{L}_0} \cup \mathscr{I}_{CH_{(n)}}^{\mathscr{L}_0}$ *z liczbą kontekstów* $n \geq 3$ *oraz* $\alpha \geq \frac{n-1}{n}$ *zachodzi:*

$$X_{\mathrm{u}}(B_\alpha) = \log_2[n(n-1)^{-\alpha}] - h(\alpha). \tag{4.107}$$

Dodatkowo dla parzystych wartości n *oraz* $\alpha \leq \frac{1}{n}$ *zachodzi:*

$$X_{\mathrm{u}}(B_\alpha) = \log_2[n(n-1)^{(\alpha-1)}] - h(\alpha), \tag{4.108}$$

gdzie $h(\alpha) = -\alpha\log_2\alpha - (1-\alpha)\log_2(1-\alpha)$.

Rys. 4.8 przedstawia wykres miary $X_{\mathrm{u}}(B_\alpha)$ w zależności od parametru α dla kilku wybranych układów.

Dowód.

W każdym przypadku, tj. dla układów izotropowych PM_α, M_α oraz $CH_\alpha^{(n)}$ spełnione są założenia Twierdzenia 4, na mocy którego:

$$X_{\mathrm{u}}(B_\alpha) = \min_{p(\lambda)\in\mathscr{I}_{B_\alpha}^{\mathscr{L}_0}} \sum_c \frac{1}{n} D(g(\lambda_c)||p(\lambda_c)), \tag{4.109}$$

gdzie minimum wzięte jest po takich łącznych rozkładach prawdopodobieństw $p(\lambda)$, z których skonstruować można niekontekstualne układy ze zbioru $B-\mathscr{L}_0$-izotropowych układów $\mathscr{I}_{B_\alpha}^{\mathscr{L}_0}$. Dla wygody, zbiór tych wszystkich niekontekstualnych układów izotropowych oznaczmy przez $\mathscr{I}_{B_{\tilde{\alpha}}}^{\mathscr{L}_0}$, natomiast układ z tego zbioru przez $B_{\tilde{\alpha}}$.

Zauważmy teraz, że w ogólności układy izotropowe B_α postaci $\alpha B+(1-\alpha)B'$ mają jedynie dwa rodzaje rozkładów:

$$P^\alpha_{\text{parz}} \equiv \alpha P_{\text{parz}} + (1-\alpha)P_{\text{nparz}}, \tag{4.110}$$

oraz

$$P^\alpha_{\text{nparz}} \equiv \alpha P_{\text{nparz}} + (1-\alpha)P_{\text{parz}}. \tag{4.111}$$

Możemy zatem napisać:

$$X_{\text{u}}(B_\alpha) = \min_{\tilde{\alpha}} \frac{1}{n}\left((n-1)D(P^\alpha_{\text{parz}}||P^{\tilde{\alpha}}_{\text{parz}}) + D(P^\alpha_{\text{nparz}}||P^{\tilde{\alpha}}_{\text{nparz}})\right), \tag{4.112}$$

gdzie $\tilde{\alpha}$ ma ograniczenia wynikające z Lematu 9.

Zauważmy, że ponieważ dla układów izotropowych X_{u} ma w definicji minimum po układach izotropowych, mamy spełnione założenia Lematu 10. Istotnie, dla $\mathscr{S} = \mathscr{I}_{B_{\tilde{\alpha}}}^{\mathscr{L}_0}$, natomiast operację $\mathfrak{P}_c$ zdefiniujmy jako identyczność na wszystkich kontekstach $n-1$, dla których układ B_α ma rozkład P^α_{parz}, natomiast dla kontekstu mającego rozkład P^α_{nparz} jako negację wyniku jednej wybranej obserwabli. Wówczas otrzymujemy:

$$X_{\text{u}}(B_\alpha) = \min_{\tilde{\alpha}} D(P^\alpha_{\text{parz}}||P^{\tilde{\alpha}}_{\text{parz}}). \tag{4.113}$$

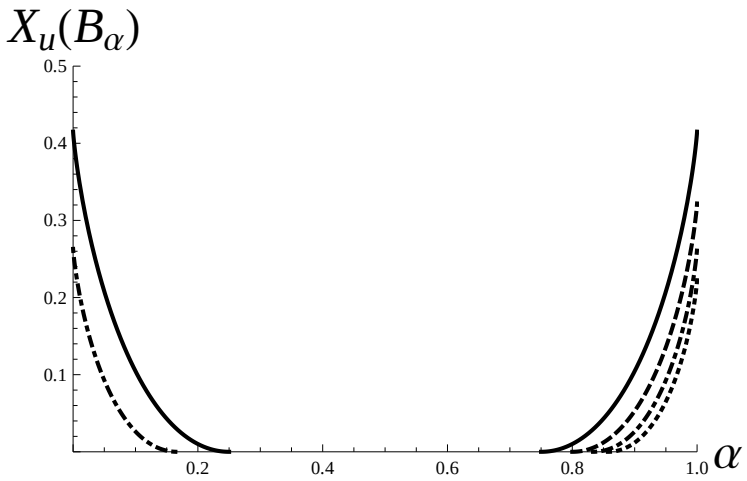
Teraz, dla wybranego kontekstu c_0 określonego układu binarnego mającego m obserwabli w każdym kontekście możemy napisać:

$$\begin{aligned} &D(P^\alpha_{\text{parz}}||P^{\tilde{\alpha}}_{\text{parz}}) = \\ &= \sum_{\lambda_{c_0}} g(\lambda_{c_0}) \log_2 \frac{g(\lambda_{c_0})}{p(\lambda_{c_0})} \\ &= 2^{m-1}\frac{\alpha}{2^{m-1}} \log_2 \left(\frac{\alpha}{2^{m-1}} \Big/ \frac{2^{m-1}}{\tilde{\alpha}}\right) + 2^{m-1}\frac{1-\alpha}{2^{m-1}} \log_2 \left(\frac{1-\alpha}{2^{m-1}} \Big/ \frac{2^{m-1}}{1-\tilde{\alpha}}\right) \\ &= \log_2 \frac{1}{\tilde{\alpha}} \left(\frac{1}{\tilde{\alpha}} - 1\right)^{\alpha-1} - h(\alpha). \end{aligned} \tag{4.114}$$

Łatwo sprawdzić, że:

$$\frac{\partial}{\partial \tilde{\alpha}} D(P^\alpha_{\text{parz}}||P^{\tilde{\alpha}}_{\text{parz}}) = \frac{\tilde{\alpha}-\alpha}{\tilde{\alpha}(1-\tilde{\alpha})}, \tag{4.115}$$

a zatem dla $\alpha \geq \tilde{\alpha}$ wyrażenie (4.113) jest funkcją malejącą z $\tilde{\alpha}$. Zgodnie z Lematem



Rys. 4.8: Wykres zależności miary X_{u} postaci (4.107) od parametru α dla różnych izotropowych układów łańcuchowych: $CH_\alpha^{(4)}$ (linia ciągła), $CH_\alpha^{(5)}$ (linia przerywana), $CH_\alpha^{(6)}$ (linia przerywano-kropkowana) oraz $CH_\alpha^{(7)}$ (linia kropkowana).

9 maksymalna wartość dla niekontekstualnych układów izotropowych jest osiągalna dla $\tilde{\alpha} = \frac{(n-1)}{n}$, a zatem dla tej wartości funkcja osiąga minimum, stąd (4.107).

Analogicznie, w przypadku parzystej wartości n, dla $\alpha \leq \tilde{\alpha}$ funkcja jest rosnąca, a zatem osiąga ona minimalną wartość dla $\tilde{\alpha} = \frac{1}{n}$, stąd (4.108).■

Do tej pory zajmowaliśmy się wyłącznie wyznaczeniem miary X_{u} dla układów izotropowych. Poniżej zobaczymy, że w przypadku układów izotropowych miary X_{u} oraz $X_{\max}$ są równe.

Twierdzenie 9. *Dla układów izotropowych* $B_\alpha \in \mathscr{I} \equiv \mathscr{I}_{PM}^{\mathscr{L}_0} \cup \mathscr{I}_{M}^{\mathscr{L}_0} \cup \mathscr{I}_{CH_{(n)}}^{\mathscr{L}_0}$ *zachodzi:*

$$X_{\max}(B_\alpha) = X_{\mathrm{u}}(B_\alpha). \tag{4.116}$$

Dowód.

Niech $p^*(\lambda)$ będzie łącznym rozkładem prawdopodobieństwa optymalnym dla miary $X_{\mathrm{u}}(B_\alpha)$. Ponieważ B_α jest układem izotropowym, to zgodnie z Twierdzeniem 4 układ skonstruowany z rozkładu $p^*(\lambda)$ również może być wzięty jako izotropowy. Wyznaczmy wartość miary pomocniczej zdefiniowanej jako:

$$X^*(B_\alpha) := \sum_c p(c) D(g(\lambda_c)||p(\lambda_c)), \tag{4.117}$$

gdzie rozkłady prawdopodobieństwa $p(\lambda_c)$ są rozkładami brzegowymi pochodzącymi

z rozkładu $p^*(\lambda)$. Teraz, zgodnie z dowodem Lematu 10 mamy, że dla $X^*(B_\alpha)$:

$$\begin{aligned} X^*(B_\alpha) &= \sum_c p(c) D(g(\lambda_c)||p(\lambda_c)) \\ &= \sum_c p(c) D(g(\lambda_{c_0})||p(\lambda_{c_0})) \\ &= D(g(\lambda_{c_0})||p(\lambda_{c_0})). \end{aligned} \tag{4.118}$$

Z drugiej strony, zgodnie z Lematem 10 dla rozkładu $p(c) = \frac{1}{n}$ otrzymujemy

$$X_{\mathrm{u}}(B_\alpha) = D(g(\lambda_{c_0})||p(\lambda_{c_0})), \tag{4.119}$$

co w rezultacie prowadzi do równości

$$X^*(B_\alpha) = X_{\mathrm{u}}(B_\alpha). \tag{4.120}$$

Z definicji miary $X_{\{p(c)\}}$ mamy oczywiście:

$$X_{\{p(c)\}}(B_\alpha) \leq X^*(B_\alpha), \tag{4.121}$$

a zatem

$$X_{\{p(c)\}}(B_\alpha) \leq X_{\mathrm{u}}(B_\alpha), \tag{4.122}$$

z kolei wzięcie supremum obustronnie daje:

$$X_{\max}(B_\alpha) \leq X_{\mathrm{u}}(B_\alpha). \tag{4.123}$$

Zważywszy jednak na przeciwną nierówność (4.21):

$$X_{\max}(B_\alpha) \geq X_{\mathrm{u}}(B_\alpha), \tag{4.124}$$

otrzymujemy równość obu miar dla układów izotropowych.■

Zajmijmy się teraz układem K, dla którego zbiór $\mathscr{L}_0$ został zdefiniowany w sekcji 4.4.2. Na tym, stosunkowo przejrzystym przykładzie pokażemy wprost rezultat zastosowania symetryzacji $\tau_B^{\mathscr{L}_0}$ działającej na odpowiednie rozkłady prawdopodobieństwa.

Celem wyznaczenia miary X_{u} ponownie wykorzystamy Twierdzenie 4. Klasyczny łączny rozkład prawdopodobieństwa $p(\lambda)$ dla 5 operatorów rzutowych Π_i w ogólności posiada $2^5 = 32$ punkty ekstremalne, tj. rozkłady prawdopodobieństw P_{ijklm} takie, że prawdopodobieństwo

$$p(ijklm|\Pi_1\Pi_2\Pi_3\Pi_4\Pi_5) = 1, \tag{4.125}$$

dla $i,j,k,l,m \in \{0,1\}$. Po dokonaniu symetryzacji $\tau_B^{\mathscr{L}_0}$ punkty ekstremalne zostają odwzorowane na 8 rozkładów niezmienniczych ze względu na działanie grupy $\mathbb{D}_5$:

$$\tilde{p}^1(\lambda) = P_{00000}, \tag{4.126}$$
$$\tilde{p}^2(\lambda) = \frac{1}{5}\left(P_{00001} + P_{00010} + P_{00100} + P_{01000} + P_{10000}\right), \tag{4.127}$$
$$\tilde{p}^3(\lambda) = \frac{1}{5}\left(P_{00011} + P_{00110} + P_{01100} + P_{11000} + P_{10001}\right), \tag{4.128}$$
$$\tilde{p}^4(\lambda) = \frac{1}{5}\left(P_{00101} + P_{01010} + P_{10100} + P_{01001} + P_{10010}\right), \tag{4.129}$$
$$\tilde{p}^5(\lambda) = \frac{1}{5}\left(P_{00111} + P_{01110} + P_{11100} + P_{11001} + P_{10011}\right), \tag{4.130}$$
$$\tilde{p}^6(\lambda) = \frac{1}{5}\left(P_{01101} + P_{11010} + P_{10101} + P_{01011} + P_{10110}\right), \tag{4.131}$$
$$\tilde{p}^7(\lambda) = \frac{1}{5}\left(P_{01111} + P_{11110} + P_{11101} + P_{11011} + P_{10111}\right), \tag{4.132}$$
$$\tilde{p}^8(\lambda) = P_{11111}. \tag{4.133}$$

Po tak dokonanej symetryzacji rozkładu prawdopodobieństwa $p(\lambda)$ widzimy, że zawężając go do dowolnego kontekstu, rozkłady są niezależne od wyboru kontekstu, a zatem wystarczy szczegółowo rozpatrzyć np. pierwszy kontekst, tj. parę operatorów rzutowych $c_1 = \{\Pi_1, \Pi_2\}$. Weźmy zatem dla przykładu rozkład $\tilde{p}^6(\lambda)$. Zawężając go do kontekstu c_1 dostajemy następujące prawdopodobieństwa otrzymania określonych wyników pomiarów:

$$\begin{aligned} p(01|\Pi_1\Pi_2) &= \frac{2}{5}, \\ p(10|\Pi_1\Pi_2) &= \frac{2}{5}, \\ p(11|\Pi_1\Pi_2) &= \frac{1}{5}. \end{aligned} \tag{4.134}$$

W analogiczny sposób możemy znaleźć rozkłady prawdopodobieństwa w pozostałych przypadkach, otrzymując dla pierwszego kontekstu:

$$\tilde{p}^1(\lambda_{c_1}) = \{p(00) = 1\}, \tag{4.135}$$
$$\tilde{p}^2(\lambda_{c_1}) = \left\{p(00) = \frac{3}{5}, p(01) = \frac{1}{5}, p(10) = \frac{1}{5}\right\}, \tag{4.136}$$
$$\tilde{p}^3(\lambda_{c_1}) = \left\{p(00) = \frac{2}{5}, p(01) = \frac{1}{5}, p(10) = \frac{1}{5}, p(11) = \frac{1}{5}\right\}, \tag{4.137}$$
$$\tilde{p}^4(\lambda_{c_1}) = \left\{p(00) = \frac{1}{5}, p(01) = \frac{2}{5}, p(10) = \frac{2}{5}\right\}, \tag{4.138}$$
$$\tilde{p}^5(\lambda_{c_1}) = \left\{p(00) = \frac{1}{5}, p(01) = \frac{1}{5}, p(10) = \frac{1}{5}, p(11) = \frac{2}{5}\right\}, \tag{4.139}$$

$$\tilde{p}^6(\lambda_{c_1}) = \left\{ p(01) = \frac{2}{5}, p(10) = \frac{2}{5}, p(11) = \frac{1}{5} \right\}, \tag{4.140}$$

$$\tilde{p}^7(\lambda_{c_1}) = \left\{ p(01) = \frac{1}{5}, p(10) = \frac{1}{5}, p(11) = \frac{3}{5} \right\}, \tag{4.141}$$

$$\tilde{p}^8(\lambda_{c_1}) = \left\{ p(11) = 1 \right\}. \tag{4.142}$$

Zbiór powyższych zsymetryzowanych rozkładów prawdopodobieństwa przedstawić można obrazowo na wykresie trójkątnym jako zbiór punktów na płaszczyźnie wyznaczonej przez względny udział poszczególnych prawdopodobieństw w odpowiednich rozkładach (Rys. 4.9). Widzimy, że dowolny zsymetryzowany niekontekstualny rozkład prawdopodobieństwa dla wybranego kontekstu $\tilde{p}(\lambda_{c_1})$ otrzymać możemy jako mieszaninę czterech rozkładów: $\tilde{p}^1(\lambda_{c_1})$, $\tilde{p}^4(\lambda_{c_1})$, $\tilde{p}^6(\lambda_{c_1})$ oraz $\tilde{p}^8(\lambda_{c_1})$. Tym samym możemy napisać:

$$X_{\mathrm{u}}(K) = \min_{\tilde{p}(\lambda_{c_1})} D(g(\lambda_{c_1}) || \tilde{p}(\lambda_{c_1})), \tag{4.143}$$

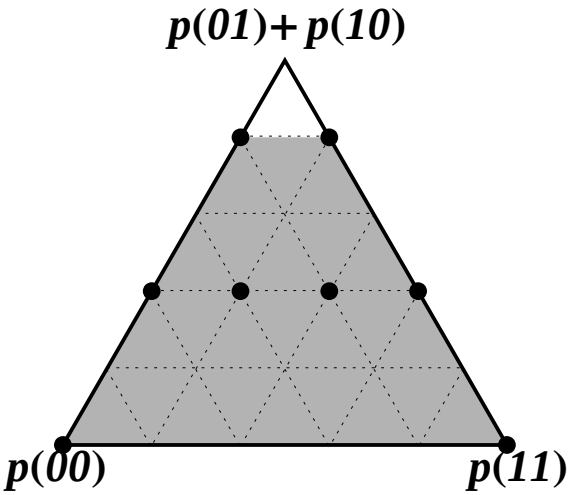
gdzie wykorzystaliśmy równoważność wszystkich kontekstów ($p(c) = \frac{1}{5}$). Powyższe wyrażenie należy zatem zminimalizować po czterech niezależnych parametrach: $p(00)$, $p(01)$, $p(10)$ oraz $p(11)$. Zauważmy jednak, że w każdym przypadku zachodzi $p(01) = p(10)$. Warunek normalizacji z kolei redukuje liczbę niezależnych parametrów w minimalizacji powyższego wyrażenia do dwóch, np. $p(00)$ i $p(11)$. Zauważmy również, że ze względu na fakt $g(00) = 0$ warto przyjąć $p(00) = 0$. W przeciwnym bowiem przypadku zmniejszeniu ulegną pozostałe prawdopodobieństwa $p(00)$, $p(01)$ i $p(10)$, co przyczyni się do wzrostu względnej entropii. Aby zatem otrzymać wartość miary X_{u} wystarczy zminimalizować funkcję

$$\begin{aligned} & D(g(\lambda_{c_1}) || \tilde{p}(\lambda_{c_1})) = \\ & = \sum_{\lambda_{c_1}} g(\lambda_{c_1}) \log_2 \frac{g(\lambda_{c_1})}{p(\lambda_{c_1})} \\ & = g(00) \log_2 \frac{g(00)}{p(00)} + 2g(01) \log_2 \frac{g(01)}{p(01)} \\ & = \log_2 \frac{1}{p(00)} \left(\frac{1}{p(00)} - 1 \right)^{g(00)-1} - h(g(00)), \end{aligned} \tag{4.144}$$

ze względu na parametr $p(00)$, który dla układów niekontekstualnych zawiera się w granicach $\frac{1}{5} \leq p(00) \leq 1$. Łatwo sprawdzić, że:

$$\frac{\partial}{\partial p(00)} D(g(\lambda_{c_1}) || \tilde{p}(\lambda_{c_1})) = \frac{p(00) - g(00)}{p(00)(1 - p(00))}, \tag{4.145}$$

więc zgodnie z $p(00) > g(00)$ wyrażenie (4.143) jest funkcją rosnącą z $p(00)$. Minimalna wartość $p(00)$ dla niekontekstualnych układów izotropowych jest osiągalna dla $p(00) = \frac{1}{5}$, a zatem dla tej wartości funkcja osiąga minimum.



Rys. 4.9: Orbity powstałe przez zsymetryzowanie punktów ekstremalnych rozkładu $p(\lambda)$ będących niezmiennikami działania grupy $\mathbb{D}_5$, w parametryzacji prawdopodobieństw rozkładu dla pojedynczego kontekstu. Zbiór zsymetryzowanych niekontekstualnych rozkładów prawdopodobieństwa zgodnych z układem K jest równoważny uwypukleniu zbioru rozkładów $\tilde{p}^1(\lambda_{c_1})$, $\tilde{p}^4(\lambda_{c_1})$, $\tilde{p}^6(\lambda_{c_1})$ oraz $\tilde{p}^8(\lambda_{c_1})$ (zacieniony obszar).

Ze wspomnianej wcześniej równoważności wszystkich kontekstów, dla układu K, podobnie jak w przypadku układów izotropowych, zachodzi $X_{\max} = X_{\mathrm{u}}$.

4.4.4 Zestawienie wyników

Zauważmy, że zgodnie z tezą Twierdzenia 8 w przypadku układów izotropowych dla parzystej liczby kontekstów n wartość miary $X_{\mathrm{u}}(B_\alpha)$ jest równa wartości $X_{\mathrm{u}}(B_{\alpha'})$, gdzie $\alpha' = 1 - \alpha$. Wynika to z faktu, że układ przeciwny $B'_\alpha \equiv B_{\alpha'} = B_{1-\alpha}$ otrzymać możemy z układu B_α poprzez negację wyników odpowiednich obserwabli, która nie zmienia wartości względnej entropii w definicji $X_{\mathrm{u}}(B_\alpha)$.

Korzystając z otrzymanych wyników możemy w łatwy sposób znaleźć wartości miary $X_{\max}(B_\alpha)$ dla rozważanych ekstremalnych układów izotropowych ($\alpha = 1$) oraz układu K:

$$X_{\max}(PR) = \log_2 \frac{4}{3} \approx 0,4150, \tag{4.146}$$

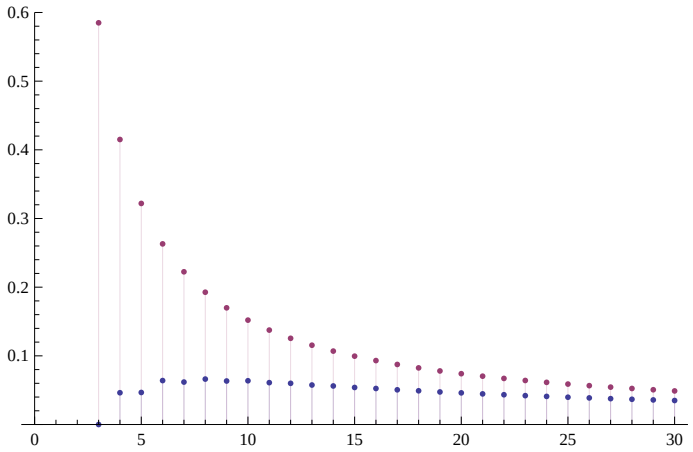
$$X_{\max}(CH_{(n)}) = \log_2 \frac{n}{n-1}, \tag{4.147}$$

$$X_{\max}(PM) = \log_2 \frac{6}{5} \approx 0,2630, \tag{4.148}$$

$$X_{\max}(M) = \log_2 \frac{5}{4} \approx 0,3219, \tag{4.149}$$

$$X_{\max}(K) \approx 0,0467. \tag{4.150}$$

Warto zwrócić uwagę na różnicę niemal rzędu wielkości między miarą $X_{\max}$ dla ukła-



Rys. 4.10: Wartość miary $X_{\max}$ dla układów łańcuchowych w zależności od liczby kontekstów n. Górne punkty odpowiadają wartościom miary dla maksymalnie kontekstualnych układów ($\alpha = 1$), dolne punkty odpowiadają wartościom miary dla maksymalnie kontekstualnych układów kwantowych. Źródło: [14].

dów kwantowych PM i M a układem K.

Dla porównania, w przypadku maksymalnego łamania nierówności CHSH odpowiadający temu (maksymalnie kontekstualny) kwantowy układ to $B_{CHSH} \equiv CH_{\alpha}^{(4)}$ z parametrem α równym $\alpha = \cos^2 \frac{\pi}{8}$. Dla tak zdefiniowanego układu otrzymujemy:

$$X_{\max}(B_{CHSH}) \approx 0.0463. \tag{4.151}$$

W ogólności, w przypadku układów łańcuchowych, maksymalnie kwantowe układy [92] otrzymujemy dla parametru α równego:

$$\alpha = \frac{2\cos\frac{\pi}{n}}{1+\cos\frac{\pi}{n}} \tag{4.152}$$

dla nieparzystych n oraz:

$$\alpha = \frac{1}{2}(1+\cos\frac{\pi}{n}) \tag{4.153}$$

dla parzystych n. W szczególności, dla $n = 5$ mamy również $K \equiv CH_{\alpha}^{(5)}$ z parametrem α równym $\alpha = \frac{2}{\sqrt{5}}$ oraz

$$X_{\max}(CH_{\alpha}^{(5)}) \approx 0,0467. \tag{4.154}$$

Zwróćmy teraz uwagę, że w przypadku układów łańcuchowych miara $X_{\max}(CH_{(n)})$ dąży do zera wraz ze wzrastającą wartością liczby kontekstów n (Rys. 4.10). Z kolei miara X_{u} (w przypadku układów izotropowych równa $X_{\max}$) poprzez średniowanie dla wszystkich kontekstów określa poniekąd wartość kontekstualności przypadają-

cej na jeden kontekst. Tym samym, jeśli wzięlibyśmy pod uwagę wartość nX_{u} moglibyśmy mówić o całkowitej kontekstualności „zawartej" w danym układzie. W tym przypadku jednak, dla rosnącej liczby kontekstów wartość nX_{u} zarówno w przypadku układów maksymalnie kontekstualnych jak i maksymalnie kontekstualnych układów kwantowych dążyłaby do wartości $\log_2 e \approx 1,4427$, przy czym w przypadku układów kwantowych byłaby ona ograniczeniem górnym, z kolei dla układów maksymalnie kontekstualnych - dolnym. Warto jednak tutaj zauważyć, że dla dużych n wartości parametru α dla układów kwantowych (4.152) oraz (4.153) dążą do 1, tym samym układy kwantowe „upodobniają się" do układów maksymalnie kontekstualnych.

Rozdział 5

Zasady wykluczania informacji

5.1. Wstęp

W bieżącym rozdziale wprowadzimy relacje wykluczania informacji ograniczające sumę dwóch wzajemnych informacji dla odpowiednich par obserwabli mierzonych na dowolnych stanach dwuukładowych [19]. Pierwsza relacja dotyczy przypadku dwóch wzajemnych informacji, gdzie na pierwszym podukładzie mierzona jest tylko jedna obserwabla, a na drugim podukładzie mierzona jest jedna z dwóch obserwabli. Druga relacja dotyczy przypadku dwóch wzajemnych informacji, gdzie na każdym podukładzie może być mierzona jedna z dwóch obserwabli.

5.2. Relacja wykluczania informacji w przypadku 1:2 obserwabli

Załóżmy, że dany jest dwuukładowy stan kwantowy ρ_{AB}, na którym dokonywane mogą być pomiary przez dwoje użytkowników, Alicję oraz Bolka. Niech Alicja dokonuje pomiaru w bazie

$$\mathscr{A} = \{|a_k\rangle\}, \tag{5.1}$$

z odpowiadającymi jej jednowymiarowymi operatorami rzutowymi $\{P_k = |a_k\rangle\langle a_k|\}$, z kolei Bolek niech dokonuje pomiaru w jednej z dwóch baz

$$\mathscr{B}^{(1)} = \{|b_i^{(1)}\rangle\}, \tag{5.2}$$

oraz

$$\mathscr{B}^{(2)} = \{|b_j^{(2)}\rangle\}, \tag{5.3}$$

z odpowiadającymi im jednowymiarowymi operatorami rzutowymi $\{Q_j^{(s)} = |b_j^{(s)}\rangle\langle b_j^{(s)}|\}$ $(s = \{1, 2\})$. Stan Alicji po pomiarze przyjmuje postać

$$\rho_A' = \sum_k p_k P_k, \tag{5.4}$$

z kolei stan Bolka po pomiarze przyjmuje postać

$$\rho'_{B^{(s)}} = \sum_j p_j^{(s)} Q_j^{(s)}, \tag{5.5}$$

gdzie p_k ($p_j^{(s)}$) to prawdopodobieństwo otrzymania wyniku k (j), natomiast łączny stan Alicji i Bolka po pomiarze ma postać

$$\rho'_{AB^{(s)}} = \sum_{k,j} p_{kj}^{(s)} P_k \otimes Q_j^{(s)}. \tag{5.6}$$

Oznaczmy przez $H(\mathscr{A})$ ($H(\mathscr{B}^{(s)})$) entropię rozkładu prawdopodobieństwa $\{p_k\}$ ($\{p_j^{(s)}\}$) oraz przez $H(\mathscr{A}, \mathscr{B}^{(s)})$ entropię łącznego rozkładu prawdopodobieństwa $\{p_{kj}^{(s)}\}$. Zdefiniujmy wzajemną informację $I(\mathscr{A} : \mathscr{B}^{(s)})$ przez:

$$I(\mathscr{A} : \mathscr{B}^{(s)}) = H(\mathscr{A}) + H(\mathscr{B}^{(s)}) - H(\mathscr{A}, \mathscr{B}^{(s)}). \tag{5.7}$$

Pokażemy, że zachodzi następująca relacja, którą dalej będziemy nazywać *relacją wykluczania informacji*:

$$I(\mathscr{A} : \mathscr{B}^{(1)}) + I(\mathscr{A} : \mathscr{B}^{(2)}) \leq \log_2 d + \log_2 c, \tag{5.8}$$

gdzie d jest wymiarem przestrzeni Hilberta każdego podukładu, a współczynnik c zdefiniowany jest jako suma d pierwszych co do wielkości współczynników c_{ij}, gdzie:

$$c_{ij} = |\langle b_i^{(1)} | b_j^{(2)} \rangle|^2. \tag{5.9}$$

Powyższa relacja wykluczania informacji ogranicza wzajemne korelacje między wynikami par pomiarów: jeśli korelacje między wynikami obserwabli A oraz $B^{(1)}$ są silne, to równocześnie korelacje między wynikami obserwabli A oraz $B^{(2)}$ stają się coraz silniej ograniczone, i vice versa.

Przykład. Porównajmy relację wykluczania informacji (5.8) z relacją wykluczania informacji Halla (2.76) [4]. W tym celu załóżmy, że Alicja oraz Bolek współdzielą stan klasycznie skorelowany postaci

$$\rho_{AB} = \sum_{i=0}^{d-1} p_i |i\rangle\langle i|_A \otimes |i\rangle\langle i|_B. \tag{5.10}$$

Niech Alicja dokonuje pomiaru w bazie

$$\mathscr{A} = \{|i\rangle : i = 0, \ldots, d-1\}, \tag{5.11}$$

z kolei Bolek niech dokonuje pomiaru w jednej baz

$$\mathscr{B}^{(1)} = \{|i\rangle : i = 0, ..., d-1\}, \tag{5.12}$$

oraz

$$\mathscr{B}^{(2)} = \{|0\rangle, |\tilde{j}\rangle : \tilde{j} = 1, ..., d-1\}, \tag{5.13}$$

gdzie

$$|\tilde{j}\rangle = \frac{1}{\sqrt{d-1}} \sum_{k=1}^{d-1} \exp(\frac{2\pi i \tilde{j} k}{d-1})|k\rangle. \tag{5.14}$$

Zauważmy, że obserwable dobrane zostały w taki sposób, że posiadają jeden wspólny wektor, a pozostałe wektory są wzajemnie komplemenarne, tj. przekrycie między dwoma dowolnymi wektorami wynosi $\frac{1}{d-1}$. Rys. 5.1 przedstawia graficzną reprezentację wektorów tworzących bazy $\mathscr{B}^{(1)}$ oraz $\mathscr{B}^{(2)}$ w przypadku $d = 3$. Macierz c_{ij} dla zdefiniowanych powyżej baz jest postaci:

$$c_{ij} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{d-1} & \cdots & \frac{1}{d-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{1}{d-1} & \cdots & \frac{1}{d-1} \end{pmatrix}, \tag{5.15}$$

z czego otrzymujemy $c = 2$. Widzimy więc, że w rozważanym przypadku odpowiednia relacja wykluczania informacji jest postaci:

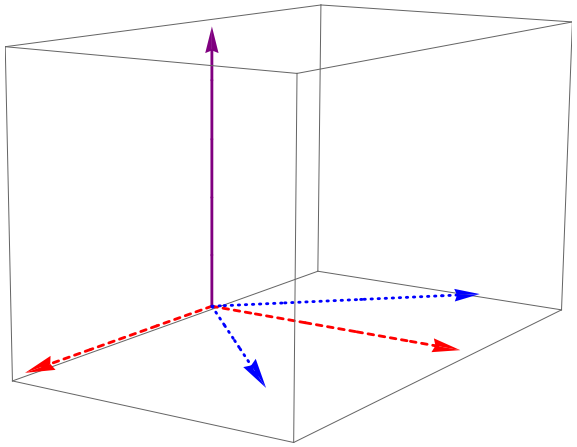
$$I(\mathscr{A} : \mathscr{B}^{(1)}) + I(\mathscr{A} : \mathscr{B}^{(2)}) \leq \log_2 d + 1. \tag{5.16}$$

Z kolei relacja nieoznaczoności Halla

$$I(\mathscr{A} : \mathscr{B}^{(1)}) + I(\mathscr{A} : \mathscr{B}^{(2)}) \leq 2\log_2 d + \log_2 \max_{i,j} c_{ij}, \tag{5.17}$$

daje trywialne ograniczenie o wartości $2\log_2 d$, ponieważ bazy $\mathscr{B}^{(1)}$ oraz $\mathscr{B}^{(2)}$ posiadają przynajmniej jeden wspólny wektor. Tym samym, dla tak dobranych baz, o ile w przypadku podukładów o wymiarze $d = 2$ obie porównywalne relacje są równoważne, to dla podukładów o wymiarze $d = 2^n$ różnica ograniczeń wynosi $n-1$ bitów informacji.

W tym miejscu warto wspomnieć, że wadliwość relacji wykluczania informacji Halla bierze się z faktu, iż relacja Halla wynika bezpośrednio z relacji nieoznaczoności Maassena-Uffinka [5]. Skądinąd wiadomo [20], że komplementarność informacji



Rys. 5.1: Wektory bazowe obserwabli (5.12) (przerywane) oraz (5.13) (kropkowane) w przypadku $d = 3$. Wspólny wektor bazowy $|0\rangle$ przedstawiony został linią ciągłą.

winna być charakteryzowana innymi zasadami niż tymi, które są optymalne w przypadku zasad nieoznaczoności. Niemniej jednak warto zastanowić się, czy sama relacja Maassena-Uffinka mogłaby w prosty sposób zostać ulepszona. W tym celu zapiszmy relację Maassena-Uffinka (2.63) w następujący sposób:

$$H(\mathscr{A}) + H(\mathscr{B}) \geq \min H_\infty(\{c_{ij}\}), \tag{5.18}$$

gdzie $H_\infty(\{c_{ij}\})$ jest min-entropią liczoną dla dowolnego rozkładu z macierzy bistochastycznej c_{ij} (dowolnego wiersza lub kolumny). Zauważmy, że $H_\infty(\{c_{ij}\})$ jest szczególnym przypadkiem entropii Rényi'ego definiowanej dla określonego rozkładu prawdopodobieństwa p_i jako [97]

$$H_\alpha(\{p_i\}) = \frac{1}{1-\alpha} \log_2 \sum_i p_i^\alpha. \tag{5.19}$$

Tym samym, wiedząc, że dla $\beta < \alpha$ zachodzi

$$H_\beta(\{p_i\}) \geq H_\alpha(\{p_i\}), \tag{5.20}$$

możemy zapytać, czy możliwe jest ulepszenie zasady nieoznaczoności do postaci

$$H(\mathscr{A}) + H(\mathscr{B}) \geq \min \left[\min_i H_\alpha(\{c_{ij}\}), \min_j H_\alpha(\{c_{ij}\}) \right], \tag{5.21}$$

dla pewnego $\alpha < \infty$. Wyniki numeryczne sugerują jednak, że powyższa relacja nie jest w ogólności słuszna dla $\alpha < \infty$: wraz ze zwiększaniem wymiaru układu znaleźć można kontrprzykłady dla relacji (5.21) dla rosnących wartości α.

Poniżej przedstawimy dowód relacji wykluczania informacji (5.8) dla pewnej ogra-

niczonej klasy stanów. Wcześniej jednak przedstawimy pomocniczy Lemat.

Lemat 11. *Załóżmy, że Alicja i Bolek współdzielą pewien stan ρ_{AB}. Niech Alicja dokonuje pomiaru w bazie $\mathscr{A} = \{|a_k\rangle\}$ z odpowiadającymi jej jednowymiarowymi operatorami rzutowymi $\{P_k = |a_k\rangle\langle a_k|\}$, z kolei Bolek niech dokonuje pomiaru w jednej z dwóch baz $\mathscr{B}^{(s)} = \{|b_j^{(s)}\rangle\}$ $(s = \{1,2\})$ z odpowiadającymi im jednowymiarowymi operatorami rzutowymi $\{Q_j^{(s)} = |b_j^{(s)}\rangle\langle b_j^{(s)}|\}$. Jeśli po pomiarze Alicji współdzielony stan jest diagonalny w bazie $|a_k\rangle \otimes |b_i^{(1)}\rangle$, tzn. jest postaci*

$$\begin{aligned}\rho_{AB^{(1)}} &= \sum_{ki} p_{ki} |a_k\rangle\langle a_k| \otimes |b_i^{(1)}\rangle\langle b_i^{(1)}| \\ &\equiv \sum_{ki} p_{ki} P_k \otimes Q_i^{(1)},\end{aligned} \tag{5.22}$$

wtedy spełniona jest relacja

$$I(\mathscr{A} : \mathscr{B}^{(1)}) + I(\mathscr{A} : \mathscr{B}^{(2)}) \leq \log_2 d + \log_2 \textstyle\sum_{ij} c_{ij}^2, \tag{5.23}$$

gdzie d jest wymiarem przestrzeni Hilberta każdego podukładu, a współczynniki c_{ij} zdefiniowane są równaniem:

$$c_{ij} = |\langle b_i^{(1)} | b_j^{(2)} \rangle|^2. \tag{5.24}$$

Dowód.

Jeśli po pomiarze Alicji w bazie $\mathscr{A} = \{|a_k\rangle\}$ na stanie $\rho_{AB^{(1)}}$ Bolek dokona pomiaru w bazie $\mathscr{B}^{(1)} = \{|b_j^{(1)}\rangle\}$, to stan nie ulegnie zmianie. Z kolei jeśli Bolek dokona pomiaru w bazie $\mathscr{B}^{(2)} = \{|b_j^{(2)}\rangle\}$, wtedy stan przybierze postać

$$\begin{aligned}\rho_{AB^{(2)}} &= \sum_{ijk} p_{ki} P_k \otimes Q_j^{(2)} Q_i^{(1)} Q_j^{(2)} \\ &\equiv \sum_{ijk} p_{ki} c_{ij} P_k \otimes Q_j^{(2)}.\end{aligned} \tag{5.25}$$

Dla uproszczenia możemy założyć, że stan po pomiarze Alicji jest postaci

$$\rho_{AB^{(1)}} = \sum_i p_i P_i \otimes Q_i^{(1)}, \tag{5.26}$$

który nie ulegnie zmianie po pomiarze Bolka dokonanym w bazie $\mathscr{B}^{(1)}$. Jeśli natomiast Bolek wykona pomiar w bazie $\mathscr{B}^{(2)}$, wtedy stan przybierze postać

$$\rho_{AB^{(2)}} = \sum_{ij} p_i c_{ij} P_i \otimes Q_j^{(2)}. \tag{5.27}$$

Obliczając sumę dwóch wzajemnych informacji dla poszczególnych par pomiarów,

dostajemy:

$$\begin{aligned} I(\mathscr{A}:\mathscr{B}^{(1)})+I(\mathscr{A}:\mathscr{B}^{(2)}) &= H(\mathscr{A})+I(\mathscr{A}:\mathscr{B}^{(2)}) \\ &= H(\mathscr{A})+H(\mathscr{B}^{(2)})-H(\mathscr{B}^{(2)}|\mathscr{A}) \\ &= H(\mathscr{A})+H(\mathscr{B}^{(2)})+\sum_i p_i \sum_j c_{ij}\log_2 c_{ij} \\ &\leq H(\mathscr{A})+H(\mathscr{B}^{(2)})+\sum_i p_i \log_2(\sum_j c_{ij}^2) \\ &= H(\mathscr{B}^{(2)})+\sum_i p_i \log_2(\frac{\sum_j c_{ij}^2}{p_i}) \\ &\leq \log_2 d+\log_2(\sum_{ij} c_{ij}^2), \end{aligned} \tag{5.28}$$

gdzie w czwartej i szóstej linijce wykorzystaliśmy własność wklęsłości logarytmu.

Udowodniwszy prawdziwość relacji (5.23) dla stanów postaci (5.26) zauważamy, że implikuje ona prawdziwość relacji dla stanów ogólnych postaci (5.22). Stan (5.22) można otrzymać z (5.26) po zastosowaniu lokalnego kanału po stronie Alicji, który: po pierwsze, nie zwiększa wartości wzajemnych informacji $I(\mathscr{A}:\mathscr{B}^{(1)})$ oraz $I(\mathscr{A}:\mathscr{B}^{(2)})$ (co wynika z nierówności przetwarzania danych [94]); po drugie, komutuje z obserwablami po stronie Bolka, $\mathscr{B}^{(1)}$ oraz $\mathscr{B}^{(2)}$; po trzecie, nie zwiększa wartości entropii $H(\mathscr{B}^{(2)})$.■

Zauważmy, że dla macierzy bistochastycznej c_{ij} ustalenie dowolnego z dwóch indeksów, np. $i=i_0$, definiuje bezpośrednio rozkład prawdopodobieństwa $\{c_{i_0 j}\}$. Tym samym, mamy

$$\sum_i p_i \sum_j c_{ij}\log_2 c_{ij} \leq \log_2 \max_{i,j} c_{ij}, \tag{5.29}$$

co przy skorzystaniu w czwartej linijce z (5.28) prowadzi bezpośrednio do relacji wykluczania informacji postaci

$$I(\mathscr{A}:\mathscr{B}^{(1)})+I(\mathscr{A}:\mathscr{B}^{(2)}) \leq 2\log_2 d+\log_2 \max_{i,j} c_{ij}, \tag{5.30}$$

która jest szczególnym przypadkiem relacji wykluczania informacji Halla (prawdziwej dla stanu (5.22)).

Powyższy lemat umożliwia nam wykazanie prawdziwości relacji wykluczania informacji (5.8) ograniczonej do szczególnych przypadków.

Twierdzenie 10. *Przy założeniach Lematu 11 zachodzi:*

$$I(\mathscr{A} : \mathscr{B}^{(1)}) + I(\mathscr{A} : \mathscr{B}^{(2)}) \leq \log_2 d + \log_2 c. \tag{5.31}$$

Dowód.

Zauważmy, że dla ustalonego i, $\{c_{ij}\}$ jest rozkładem prawdopodobieństwa. Tym samym zachodzi:

$$\sum_j c_{ij}^2 \leq \sum_j c_{ij} \max_j c_{ij} = \max_j c_{ij}. \tag{5.32}$$

Sumując powyższą zależność dochodzimy do

$$\sum_{ij} c_{ij}^2 \leq \sum_i \max_j c_{ij} \leq \sum c_{ij}, \tag{5.33}$$

gdzie ostatnie wyrażenie jest sumą d pierwszych co do wielkości współczynników c_{ij}.■

Na poniższym przykładzie pokażemy jednak, że relacja (5.23) nie jest prawdziwa w ogólności (co nie podważa pierwotnego przypuszczenia o prawdziwości relacji (5.8)).

Przykład. Załóżmy, że Alicja i Bolek współdzielą stan, którego każdy podukład jest wymiaru $d = 3$. Niech Alicja dokonuje pomiaru w bazie:

$$\begin{aligned} |a_1\rangle &= (1,0,0), \\ |a_2\rangle &= (0,1,0), \\ |a_3\rangle &= (0,0,1), \end{aligned} \tag{5.34}$$

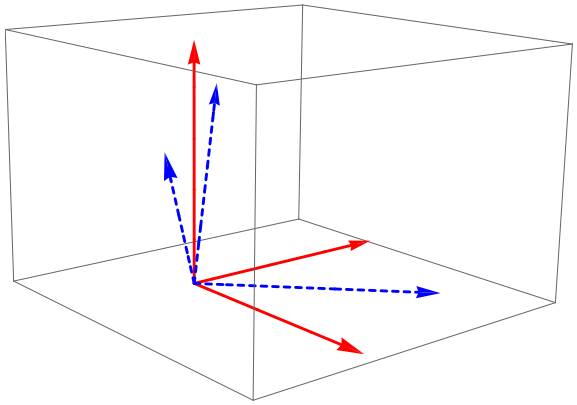
natomiast Bolek niech dokonuje pomiaru w jednej z dwóch baz:

$$\begin{aligned} |b_1^{(1)}\rangle &= (1,0,0), \\ |b_2^{(1)}\rangle &= (0,1,0), \\ |b_2^{(1)}\rangle &= (0,0,1), \end{aligned} \tag{5.35}$$

albo:

$$\begin{aligned} |b_1^{(2)}\rangle &= (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), \\ |b_2^{(2)}\rangle &= (\frac{1}{2}, -\frac{1}{2}, \frac{1}{\sqrt{2}}), \\ |b_2^{(2)}\rangle &= (-\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}), \end{aligned} \tag{5.36}$$

(patrz Rys. 5.2).



Rys. 5.2: Wektory bazowe obserwabli (5.34) oraz (5.35) (ciągłe) oraz (5.36) (przerywane). Wektor $|b_1^{(1)}\rangle$ leży w płaszczyźnie rozpiętej przez wektory $|b_2^{(2)}\rangle$ oraz $|b_3^{(2)}\rangle$. Podobnie, wektor $|b_1^{(2)}\rangle$ leży w płaszczyźnie rozpiętej przez wektory $|b_2^{(1)}\rangle$ oraz $|b_3^{(1)}\rangle$.

Rozpatrzmy przypadek, kiedy powyżej zdefiniowane pomiary dokonywane są na stanie

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}}(|a_1\rangle \otimes |b_3^{(1)}\rangle + |a_2\rangle \otimes |b_1^{(2)}\rangle). \tag{5.37}$$

Suma dwóch wzajemnych informacji dla wybranych obserwabli wynosi:

$$I(\mathscr{A} : \mathscr{B}^{(1)}) + I(\mathscr{A} : \mathscr{B}^{(2)}) = 2. \tag{5.38}$$

Z drugiej strony, macierz c_{ij} w tym przypadku jest postaci:

$$c_{ij} = |\langle b_i^{(1)}|b_j^{(2)}\rangle|^2 = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \tag{5.39}$$

z czego dla prawej strony nierówności (5.23) mamy:

$$\log_2 d + \log_2 \sum_{ij} c_{ij}^2 = \log_2 3 + \log_2 \frac{5}{4} < 2, \tag{5.40}$$

a zatem relacja (5.23) nie jest słuszna w ogólności.

5.2.1 Ogólny dowód relacji (5.8) przedstawiony w pracy (20)

Poniżej przedstawimy pełny dowód relacji wykluczania informacji (5.8), który został przedstawiony w pracy [20]. Główną tezą tej pracy jest ulepszona wersja entropowej relacji nieoznaczoności z uwzględnieniem pamięci kwantowej [66]. Mianowicie, pokazano, że dla stanu dwuukładowego ρ_{AB}, gdzie A, B oznaczają tutaj układy od-

powiednio Alicji i Bolka, zachodzi następująca relacja nieoznaczoności dla wyników pomiarów w bazach $\mathscr{B}^{(s)}$ ($s \in \{1,2\}$):

$$S(\mathscr{B}^{(1)}|A) + S(\mathscr{B}^{(2)}|A) \geq h + S(B|A), \tag{5.41}$$

gdzie

$$h = \max\{h^{(1)}, h^{(2)}\}, \tag{5.42}$$

przy czym $h^{(s)}$ zdefiniowane są w następujący sposób:

$$h^{(1)} = -\sum_i p_i^{(1)} \log_2 \max_j c_{ij}, \tag{5.43}$$

$$h^{(2)} = -\sum_j p_j^{(2)} \log_2 \max_i c_{ij}, \tag{5.44}$$

natomiast $p_k^{(s)}$ to prawdopodobieństwo otrzymania k-tego wyniku pomiaru, jeśli pomiar jest wykonywany w bazie $\mathscr{B}^{(s)}$. $S(B|A)$ jest warunkową entropią von Neumanna dla podukładów A i B:

$$S(B|A) = S(\rho_{AB}) - S(\rho_A), \tag{5.45}$$

i analogicznie

$$S(\mathscr{B}^{(s)}|A) = S(\rho_{AB^{(s)}}) - S(\rho_A), \tag{5.46}$$

gdzie stan układu po pomiarze Bolka w bazie $B^{(s)}$ jest postaci

$$\begin{aligned} \rho_{AB^{(s)}} &= \sum_i id_{d\times d} \otimes Q_i^{(s)} \rho_{AB} id_{d\times d} \otimes Q_i^{(s)} \\ &\equiv \sum_i p_i^{(s)} \rho_{A^{(s)}}^i \otimes Q_i^{(s)}. \end{aligned} \tag{5.47}$$

W powyższym równaniu $\rho_{A^{(s)}}^i$ są stanami warunkowymi po pomiarze Bolka zredukowanymi do podukładu Alicji.

Jeśli teraz dodatkowo Alicja dokona pomiaru w bazie $\mathscr{A}$ na swoim układzie, to stan układu po pomiarze będzie postaci

$$\begin{aligned} \rho_{AB^{(s)}} &= \sum_{k,i} P_k \otimes Q_i^{(s)} \rho_{AB} P_k \otimes Q_i^{(s)} \\ &\equiv \sum_{k,i} p_{ki}^{(s)} P_k \otimes Q_i^{(s)}, \end{aligned} \tag{5.48}$$

natomiast dla otrzmanych stanów klasycznie skorelowanych dostaniemy

$$S(\mathscr{B}^{(s)}|A) = H(\mathscr{B}^{(s)}|\mathscr{A}), \tag{5.49}$$

oraz

$$S(B|\mathscr{A}) = H(B|\mathscr{A}). \tag{5.50}$$

Zapiszmy teraz entropię warunkową jako

$$H(\mathscr{B}^{(s)}|\mathscr{A}) = H(\mathscr{B}^{(s)}) - I(\mathscr{A} : \mathscr{B}^{(s)}), \tag{5.51}$$

i podstawmy to wyrażenie do nierówności (5.41). Po przekształceniach dostajemy:

$$\begin{aligned} I(\mathscr{A} : \mathscr{B}^{(1)}) + I(\mathscr{A} : \mathscr{B}^{(2)}) &\leq H(\mathscr{B}^{(1)}) + H(\mathscr{B}^{(2)}) - h - H(\mathscr{B}|\mathscr{A}) \\ &\leq H(\mathscr{B}^{(1)}) + \log_2 d - h^{(1)} \\ &= \log_2 d + \sum_i p_i^{(1)} \log_2 \frac{\max_j c_{ij}}{p_i^{(1)}} \\ &\leq \log_2 d + \log_2 \sum_i \max_j c_{ij}. \end{aligned} \tag{5.52}$$

W drugiej linijce wykorzystaliśmy fakt, że entropia $H(B^{(2)}) \leq \log_2 d$, relację $h^{(1)} \leq h$ oraz fakt, że entropia warunkowa $H(B|A) \geq 0$. W trzeciej linijce wykorzystaliśmy definicję entropii $H(B^{(1)})$. Wreszcie w czwartej linijce zastosowaliśmy własność wklęsłości logarytmu.

Zauważmy teraz, że relacja (5.52) implikuje relację wykluczania informacji (5.8). Istotnie, suma d pierwszych co wielkości współczynników w macierzy c_{ij} jest zawsze niemniejsza niż suma największych współczynników poszczególnych wierszy (bądź kolumn) macierzy c_{ij}. Tym samym mamy:

$$\log_2 \sum_i \max_j c_{ij} \leq c. \tag{5.53}$$

przez co ograniczenie (5.52) jest silniejsze niż (5.8).

5.3. Relacja wzajemnej nieoznaczoności dla dwóch par obserwabli

Załóżmy, że dany jest dwuukładowy stan maksymalnie splątany, na którym dokonywane mogą być pomiary przez dwoje użytkowników, Alicję oraz Bolka. Niech Alicja dokonuje pomiaru w jednej z dwóch baz $\mathscr{A}^{(1)}$ oraz $\mathscr{A}^{(2)}$:

$$\mathscr{A}^{(s)} = \{|a_k^{(s)}\rangle\}, \tag{5.54}$$

gdzie $s \in \{1,2\}$. Podobnie, niech Bolek dokonuje pomiaru jednej z dwóch obserwabli $\mathscr{B}^{(1)}$ oraz $\mathscr{B}^{(2)}$ w bazach

$$\mathscr{B}^{(s)} = \{|b_j^{(s)}\rangle\}, \tag{5.55}$$

gdzie $s \in \{1,2\}$.

Pokażemy, że zachodzi następująca relacja ograniczająca sumę wzajemnych informacji, z których pierwsza określa korelacje dla pierwszej pary pomiarów ($\mathscr{A}^{(1)}$ oraz $\mathscr{B}^{(1)}$), a druga określa korelacje dla drugiej pary pomiarów ($\mathscr{A}^{(2)}$ oraz $\mathscr{B}^{(2)}$):

$$I(\mathscr{A}^{(1)} : \mathscr{B}^{(1)}) + I(\mathscr{A}^{(2)} : \mathscr{B}^{(2)}) \leq 2\log_2 d + \log_2 c', \tag{5.56}$$

gdzie d jest wymiarem przestrzeni Hilberta każdego podukładu, z kolei współczynnik c' dany jest wyrażeniem:

$$c' = \max_V \max_{i,j} |\langle b_i^{(1)}|VU^TV^\dagger|b_j^{(2)}\rangle|^2, \tag{5.57}$$

przy czym U jest operatorem unitarnym transformującym dwie bazy $\{|a_k^{(1)}\rangle\}$ oraz $\{|a_k^{(2)}\rangle\}$ w następujący sposób:

$$U^\dagger|a_k^{(2)}\rangle = |a_k^{(1)}\rangle, \tag{5.58}$$

a V jest dowolnym operatorem unitarnym.

Zwróćmy uwagę, że współczynnik c' zdefiniowany jest w taki sposób, by w odpowiednich przypadkach odpowiadał on współczynnikowi c w relacji nieoznaczoności Maassena-Uffinka. Istotnie, jeśli przyjmiemy, że U jest operatorem identycznościowym, wtedy pomiary Alicji będą ograniczone tylko do jednej bazy $\{|a_k^{(1)}\rangle\}$, i współczynnik $c' = \max_{i,j} |\langle b_i^{(1)}|b_j^{(2)}\rangle|^2$. W konsekwencji, relacja (5.56) przechodzi bezpośrednio w relację wykluczania informacji Halla.

Poniżej przedstawimy dowód relacji (5.56). W tym celu udowodnimy odpowiednie twierdzenia, odnoszące się w pierwszej kolejności do relacji wykluczania informacji zależnej od stanu, natomiast w drugiej kolejności udowodnimy twierdzenie odnoszące się do relacji wykluczania informacji niezależnej od stanu. Zacznijmy od sformułowania dwóch pomocniczych lematów.

Lemat 12. *Niech Alicja dokonuje pomiaru w bazie $\mathscr{A}^{(2)} = \{|a_k^{(2)}\rangle\}$ z odpowiadającymi mu jednowymiarowymi operatorami rzutowymi $\{P_k^{(2)} = |a_k^{(2)}\rangle\langle a_k^{(2)}|\}$, z kolei Bolek niech dokonuje pomiaru w bazie $\mathscr{B}^{(2)} = \{|b_j^{(2)}\rangle\}$ z odpowiadającymi mu jednowymiarowymi operatorami rzutowymi $\{Q_j^{(2)} = |b_j^{(2)}\rangle\langle b_j^{(2)}|\}$.*

Wzajemna informacja między podukładami Alicji i Bolka liczona na stanie

$$\sum_{k,j}\langle\Phi^+|_{AB}P_k^{(2)}\otimes Q_j^{(2)}|\Phi^+\rangle_{AB}P_k^{(2)}\otimes Q_j^{(2)} \tag{5.59}$$

jest równa wzajemnej informacji między podukładami Alicji i Bolka liczonej na stanie

$$\sum_{k,j}\langle\Phi^+|_{AB}U^\dagger P_k^{(2)}U\otimes U^T Q_j^{(2)}U^*|\Phi^+\rangle_{AB}U^\dagger P_k^{(2)}U\otimes U^T Q_j^{(2)}U^*, \tag{5.60}$$

gdzie U jest dowolnym operatorem unitarnym oraz

$$|\Phi^+\rangle_{AB}=\frac{1}{\sqrt{d}}\sum_{i=0}^{d-1}|i\rangle_A\otimes|i\rangle_B. \tag{5.61}$$

Dowód.

Zauważmy, że wzajemna informacja nie ulega zmianie, gdy któryś z użytkowników dokona lokalnej operacji unitarnej. Jeśli zatem początkowo Alicja i Bolek dzielili stan postaci (5.59), to po wykonaniu przez nich operacji $U\otimes U^*$ będą oni współdzielić stan

$$\begin{aligned}&\sum_{k,j}\langle\Phi^+|_{AB}P_k^{(2)}\otimes Q_j^{(2)}|\Phi^+\rangle_{AB}U^\dagger P_k^{(2)}U\otimes U^T Q_j^{(2)}U^*\\&=\sum_{k,j}\langle\Phi^+|_{AB}U^\dagger P_k^{(2)}U\otimes U^T Q_j^{(2)}U^*|\Phi^+\rangle_{AB}U^\dagger P_k^{(2)}U\otimes U^T Q_j^{(2)}U^*,\end{aligned} \tag{5.62}$$

gdzie w powyższej równości wykorzystana została niezmienniczość stanu $|\Phi^+\rangle_{AB}$ ze względu na operację $U\otimes U^*$:

$$U\otimes U^*|\Phi^+\rangle_{AB}=|\Phi^+\rangle_{AB}.\blacksquare \tag{5.63}$$

Lemat 13. *Niech Alicja i Bolek współdzielą stan maksymalnie splątany $|\Phi^+\rangle_{AB}$. Niech dalej Alicja dokonuje pomiaru w jednej z dwóch baz $\mathscr{A}^{(s)}=\{|a_k^{(s)}\rangle\}$ ($s\in\{1,2\}$) z odpowiadającymi mu jednowymiarowymi operatorami rzutowymi $\{P_k^{(s)}=|a_k^{(s)}\rangle\langle a_k^{(s)}|\}$, z kolei Bolek niech dokonuje pomiaru w jednej z baz $\mathscr{B}^{(s)}=\{|b_j^{(s)}\rangle\}$ z odpowiadającymi mu jednowymiarowymi operatorami rzutowymi $\{Q_j^{(s)}=|b_j^{(s)}\rangle\langle b_j^{(s)}|\}$. Zachodzi:*

$$I(\mathscr{A}^{(1)}:\mathscr{B}^{(1)})+I(\mathscr{A}^{(2)}:\mathscr{B}^{(2)})\leq 2\log_2 d+\log_2\tilde{c}, \tag{5.64}$$

gdzie d jest wymiarem przestrzeni Hilberta każdego podukładu, a współczynnik $\tilde{c}$ dany jest wyrażeniem:

$$\tilde{c}=\max_{i,j}|\langle b_i^{(1)}|U^T|b_j^{(2)}\rangle|^2. \tag{5.65}$$

W powyższym równaniu U jest operacją zmiany bazy $\{|a_k^{(1)}\rangle\}$ w $\{|a_k^{(2)}\rangle\}$.

Dowód.

Zgodnie z Lematem 12 wzajemna informacja $I(\mathscr{A}^{(2)} : \mathscr{B}^{(2)})$ nie ulegnie zmianie, jeśli przed pomiarem Alicja i Bolek wykonają na współdzielonym stanie $|\Phi^+\rangle_{AB}$ operację $U \otimes U^*$. Pomiar w bazach $\mathscr{A}^{(2)}$ oraz $\mathscr{B}^{(2)}$ na tak zmienionym stanie będzie odpowiadać pomiarom w bazach odpowiednio $\{U^\dagger P_k^{(2)} U\} \equiv \{P_k^{(1)}\}$ dokonanym przez Alicję oraz $\{U^T Q_j^{(2)} U^*\} =: B_U^{(2)}$ dokonanym przez Bolka. Mamy wtedy:

$$\begin{aligned} &I(\mathscr{A}^{(1)} : \mathscr{B}^{(1)}) + I(\mathscr{A}^{(2)} : \mathscr{B}^{(2)}) \\ &= I(\mathscr{A}^{(1)} : \mathscr{B}^{(1)}) + I(\mathscr{A}^{(1)} : \mathscr{B}_U^{(2)}) \\ &= H(\mathscr{B}^{(1)}) - H(\mathscr{B}^{(1)}|\mathscr{A}^{(1)}) + H(\mathscr{B}_U^{(2)}) - H(\mathscr{B}_U^{(2)}|\mathscr{A}^{(1)}) \\ &\leq 2\log_2 d - [H(\mathscr{B}^{(1)}|\mathscr{A}^{(1)}) + H(\mathscr{B}_U^{(2)}|\mathscr{A}^{(1)})]. \end{aligned} \tag{5.66}$$

Zgodnie z definicją entropii warunkowej mamy:

$$H(\mathscr{B}^{(1)}|\mathscr{A}^{(1)}) + H(\mathscr{B}_U^{(2)}|\mathscr{A}^{(1)}) = \sum_k p_k \left(H(\mathscr{B}^{(1)})_{\rho_B^k} + H(\mathscr{B}_U^{(2)})_{\rho_B^k} \right), \tag{5.67}$$

gdzie entropie liczone są na stanach warunkowych po pomiarze Alicji zredukowanych do podukładu Bolka $\rho_B^k = \mathrm{Tr}_A \rho_{AB}^k$, gdzie:

$$\begin{aligned} \rho_{AB}^k &= \frac{P_k^{(1)} \otimes id_{d\times d} |\Phi^+\rangle\langle\Phi^+|_{AB} P_k^{(1)} \otimes id_{d\times d}}{\mathrm{Tr}(P_k^{(1)} \otimes id_{d\times d} |\Phi^+\rangle\langle\Phi^+|_{AB})} \\ &= d P_k^{(1)} \otimes id_{d\times d} |\Phi^+\rangle\langle\Phi^+|_{AB} P_k^{(1)} \otimes id_{d\times d}, \end{aligned} \tag{5.68}$$

natomiast p_k jest prawdopodobieństwem uzyskania k-tego wyniku przez Alicję:

$$p_k = \mathrm{Tr}(P_k^{(1)} \otimes id_{d\times d} |\Phi^+\rangle\langle\Phi^+|_{AB}) = \frac{1}{d}. \tag{5.69}$$

Zgodnie z zasadą nieoznaczoności Maassena-Uffinka [5] w przypadku pomiarów w bazach $\{|b_i^{(1)}\rangle\}$ oraz $\{U^T |b_j^{(2)}\rangle\}$, dla każdego wyniku k mamy:

$$H(\mathscr{B}^{(1)})_{\rho_B^k} + H(\mathscr{B}_U^{(2)})_{\rho_B^k} \geq -\log_2 \tilde{c}. \tag{5.70}$$

Podstawiając teraz (5.69) oraz (5.70) do (5.67) dostajemy:

$$H(\mathscr{B}^{(1)}|\mathscr{A}^{(1)}) + H(\mathscr{B}_U^{(2)}|\mathscr{A}^{(1)}) \geq -\sum_k \frac{1}{d} \log_2 \tilde{c}. \tag{5.71}$$

Następnie uzyskaną nierówność podstawimy do (5.66), otrzymując:

$$I(\mathscr{A}^{(1)} : \mathscr{B}^{(1)}) + I(\mathscr{A}^{(2)} : \mathscr{B}^{(2)}) \leq 2\log_2 d + \log_2 \tilde{c}. \blacksquare \tag{5.72}$$

Twierdzenie 11. *(dla relacji stanowo-zależnej) Niech Alicja i Bolek współdzielą stan maksymalnie splątany* $|\Phi\rangle_{AB}$ *taki, że:*

$$|\Phi\rangle_{AB} = id_{d\times d} \otimes V|\Phi^+\rangle_{AB}, \tag{5.73}$$

gdzie V jest operatorem unitarnym działającym na podukładzie Bolka. Niech dalej Alicja dokonuje pomiaru w bazie $\mathscr{A}^{(s)} = \{|a_k^{(s)}\rangle\}$ *(*$s \in \{1,2\}$*) z odpowiadającymi mu jednowymiarowymi operatorami rzutowymi* $\{P_k^{(s)} = |a_k^{(s)}\rangle\langle a_k^{(s)}|\}$*, z kolei Bolek niech dokonuje pomiaru w bazie* $\mathscr{B}^{(s)} = \{|b_j^{(s)}\rangle\}$ *z odpowiadającymi mu jednowymiarowymi operatorami rzutowymi* $\{Q_j^{(s)} = |b_j^{(s)}\rangle\langle b_j^{(s)}|\}$*. Wtedy zachodzi:*

$$I(\mathscr{A}^{(1)} : \mathscr{B}^{(1)}) + I(\mathscr{A}^{(2)} : \mathscr{B}^{(2)}) \leq 2\log_2 d + \log_2 \tilde{c}', \tag{5.74}$$

gdzie d jest wymiarem przestrzeni Hilberta każdego podukładu, a współczynnik $\tilde{c}'$ *zdefiniowany jest jako:*

$$\tilde{c}' = \max_{i,j} |\langle b_i^{(1)}|VU^TV^\dagger|b_j^{(2)}\rangle|^2, \tag{5.75}$$

gdzie U jest operacją zmiany bazy $\{|a_k^{(1)}\rangle\}$ *w* $\{|a_k^{(2)}\rangle\}$*.*

Dowód.

Zauważmy, że pomiary Alicji i Bolka w bazach danych przez operatory rzutowe odpowiednio $P_k^{(s)}$ i $Q_j^{(s)}$ dokonywane na stanie maksymalnie splątanym $|\Phi\rangle_{AB} = I \otimes V|\Phi^+\rangle_{AB}$ są równoważne pomiarom w bazach danych przez operatory rzutowe odpowiednio $P_k^{(s)}$ i $V^\dagger Q_j^{(s)} V$ dokonywanym na stanie maksymalnie splątanym $|\Phi^+\rangle_{AB}$. Relację (5.74) dostajemy bezpośrednio z Lematu 13 po odpowiedniej zmianie baz Bolka.■

Twierdzenie 12. *(dla relacji stanowo-niezależnej) Niech Alicja i Bolek współdzielą dowolny stan maksymalnie splątany* $|\Phi\rangle_{AB}$*. Niech dalej Alicja dokonuje pomiaru w bazie* $\mathscr{A}^{(s)} = \{|a_k^{(s)}\rangle\}$ *(*$s \in \{1,2\}$*) z odpowiadającymi jej jednowymiarowymi operatorami rzutowymi* $\{P_k^{(s)} = |a_k^{(s)}\rangle\langle a_k^{(s)}|\}$*, z kolei Bolek niech dokonuje pomiaru w bazie* $\mathscr{B}^{(s)} = \{|b_j^{(s)}\rangle\}$ *z odpowiadającymi jej jednowymiarowymi operatorami rzutowymi* $\{Q_j^{(s)} = |b_j^{(s)}\rangle\langle b_j^{(s)}|\}$*. Wtedy zachodzi:*

$$I(\mathscr{A}^{(1)} : \mathscr{B}^{(1)}) + I(\mathscr{A}^{(2)} : \mathscr{B}^{(2)}) \leq 2\log_2 d + \log_2 c' \tag{5.76}$$

gdzie d jest wymiarem przestrzeni Hilberta każdego podukładu, a współczynnik c' *zdefiniowany jest jako:*

$$c' = \max_V \max_{i,j} |\langle b_i^{(1)}|VU^TV^\dagger|b_j^{(2)}\rangle|^2. \tag{5.77}$$

W powyższym wyrażeniu V jest operatorem unitarnym, a U jest operacją zmiany bazy $\{|a_k^{(1)}\rangle\}$ *w* $\{|a_k^{(2)}\rangle\}$*.*

Dowód.

Dowód wynika bezpośrednio z Twierdzenia 11. Niech suma dwóch wzajemnych informacji będzie największa dla pewnego stanu maksymalnie splątanego

$$|\Phi_0\rangle_{AB} = id_{d\times d} \otimes V_0 |\Phi^+\rangle_{AB}. \tag{5.78}$$

Ze względu na maksymalizację po wszystkich operatorach unitarnych V mamy oczywiście

$$\max_V \max_{i,j} |\langle b_i^{(1)}|VU^TV^\dagger|b_j^{(2)}\rangle|^2 = \max_{i,j} |\langle b_i^{(1)}|V_0U^TV_0^\dagger|b_j^{(2)}\rangle|^2, \tag{5.79}$$

a zatem ograniczenie (5.76) będzie prawdziwe również w tym przypadku.■

⋆ ⋆ ⋆

Pozostaje pytanie, czy powyższe wyrażenia można uogólnić na przypadek dowolnych stanów niemaksymalnie splątanych. Zauważmy, że gdy wzajemna informacja dla pomiarów wykonywanych na stanie niemaksymalnie splątanym postaci

$$|\Phi\rangle_{AB} = \sum_i \sqrt{\lambda_i}|e_i^A\rangle \otimes |e_i^B\rangle, \tag{5.80}$$

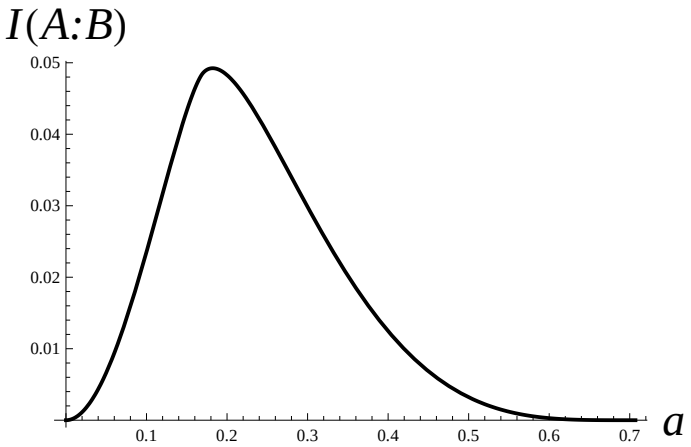
gdzie $|e_i^A\rangle \otimes |e_i^B\rangle$ jest bazą Schmidta, byłaby mniejsza niż wzajemna informacja dla pomiarów wykonywanych na stanie maksymalnie splątanym postaci

$$|\Phi\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_i |e_i^A\rangle \otimes |e_i^B\rangle, \tag{5.81}$$

wtedy Twierdzenia 11 oraz 12 byłyby również prawdziwe dla wszystkich stanów niemaksymalnie splątanych. Poniżej pokażemy jednak, że istnieją pomiary, które wykonywane na stanach maksymalnie splątanych dają w rezultacie mniejszą wartość wzajemnej informacji niż w przypadku, gdyby były wykonywane na stanie niemaksymalnie splątanym.

Przykład. Rozpatrzmy przypadek $d = 2$. Niech Alicja dokonuje pomiaru w bazie $\mathscr{A}$ zdefiniowanej przez wektory

$$\begin{aligned} |a_1\rangle &= \mathscr{N}^-[(1-\sqrt{2})|0\rangle + |1\rangle], \\ |a_2\rangle &= \mathscr{N}^+[(1+\sqrt{2})|0\rangle + |1\rangle], \end{aligned} \tag{5.82}$$



Rys. 5.3: Wzajemna informacja $I(\mathscr{A}:\mathscr{B})$ dla pomiarów w bazach (5.82) i (5.83) liczona na stanie $|\Phi\rangle_{AB}(a) = a|0\rangle\otimes|0\rangle + \sqrt{1-a^2}|1\rangle\otimes|1\rangle$ w zależności od wartości a.

natomiast Bolek niech dokonuje pomiaru w bazie $\mathscr{B}$ zdefiniowanej przez wektory

$$\begin{aligned} |b_1\rangle &= \mathscr{N}^+[(-1-\sqrt{2})|0\rangle + |1\rangle], \\ |b_2\rangle &= \mathscr{N}^-[(-1+\sqrt{2})|0\rangle + |1\rangle], \end{aligned} \tag{5.83}$$

gdzie

$$\mathscr{N}^{\pm} = \left(1 + (\sqrt{2}\pm 1)^2\right)^{-\frac{1}{2}}. \tag{5.84}$$

Zauważmy tutaj, że bazy $\mathscr{A}$ i $\mathscr{B}$ są komplementarne. Niech pomiary będą wykonywane na stanie postaci

$$|\Phi(a)\rangle_{AB} = a|0\rangle\otimes|0\rangle + \sqrt{1-a^2}|1\rangle\otimes|1\rangle. \tag{5.85}$$

Jeśli pomiary w bazach komplementarnych (5.82) oraz (5.83) wykonywane są na stanie maksymalnie splątanym $|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle\otimes|0\rangle + |1\rangle\otimes|1\rangle)$, to wzajemna informacja

$$I(\mathscr{A}:\mathscr{B})_{|\Phi^+\rangle_{AB}} = 0. \tag{5.86}$$

Natomiast dla stanu niemaksymalnie splątanego znajdujemy przykładowo

$$I(\mathscr{A}:\mathscr{B})_{|\Phi(0,18)\rangle_{AB}} \approx 0,049. \tag{5.87}$$

Rys. 5.3 przedstawia zależność wzajemnej informacji $I(\mathscr{A}:\mathscr{B})$ dla stanów (5.85) od parametru a.

Kwestią otwartą jest prawdziwość relacji wykluczania informacji następującej po-

staci:

$$I(\mathscr{A}^{(1)} : \mathscr{B}^{(1)}) + I(\mathscr{A}^{(2)} : \mathscr{B}^{(2)}) \leq \log_2 d + \log_2 c'', \tag{5.88}$$

gdzie d jest wymiarem przestrzeni Hilberta każdego podukładu, a współczynnik c'' zdefiniowany jest jako

$$c'' = \max_V \sum |\langle b_i^{(1)}|VU^TV^\dagger|b_j^{(2)}\rangle|^2, \tag{5.89}$$

przy czym sumowanie jest po d pierwszych co do wielkości współczynnikach $|\langle b_i^{(1)}|VU^TV^\dagger|b_j^{(2)}\rangle|^2$. W powyższym wyrażeniu V jest operatorem unitarnym, a U jest operacją zmiany bazy po stronie Alicji.

Powyższa relacja wykluczania informacji (5.88) jest uogólnieniem relacji (5.8) na przypadek, gdy każdy z użytkowników może dokonywać pomiarów w jednej z dwóch baz. Istotnie, jeśli przyjmiemy, że U jest operatorem identycznościowym, wtedy pomiary Alicji będą ograniczone tylko do jednej bazy $\{|a_k^{(1)}\rangle\}$ i co za tym idzie, współczynnik c'' będzie równy współczynnikowi c zdefiniowanemu poprzednio jako suma d pierwszych co do wielkości współczynników $c_{ij} = |\langle b_i^{(1)}|b_j^{(2)}\rangle|^2$.

Rozdział 6

Równoważność układów PR i kodów swobodnego dostępu

6.1. Wstęp

W bieżącym rozdziale zajmiemy się porównaniem dwóch odmiennych zasobów teorio-informatycznych: układów PR [2,3] z kodami swobodnego dostępu [55]. Wiedząc, że kod swobodnego dostępu może być symulowany przez układ PR z dodatkowym bitem klasycznej komunikacji [52–54], wykażemy równość tych dwóch zasobów poprzez zaprezentowanie możliwości symulowania układów PR przez kody swobodnego dostępu [21]. W szczególności sformułujemy nierówność wiążącą możliwości wykorzystania tych zasobów i pokażemy jak może być ona nasycana.

6.2. Podstawowe pojęcia

Na potrzeby tego rozdziału, zdefiniujmy układ PR w nieco odmienny sposób jak zostało to uczynione w Rozdziale 4. Rozpatrzmy dwuukładowe systemy współdzielone przez dwoje użytkowników, Alicję i Bolka. Niech Alicja dokonuje jednego z dwóch dychotomicznych pomiarów sygnowanych symbolem $x = 0,1$, i analogicznie niech Bolek dokonuje jednego z dwóch dychotomicznych pomiarów sygnowanych symbolem $y = 0,1$. Oznaczmy wyniki odpowiednich pomiarów poprzez X i Y ($X, Y \in \{0,1\}$).

Układem niesygnalizującym nazywamy taki układ, tj. zbiór łącznych rozkładów prawdopodobieństw $p(XY|xy)$, dla którego zachodzą następujące warunki:

$$\forall_{X,x,y,y'} \sum_{Y=0,1} p(XY|xy) = \sum_{Y=0,1} p(XY|xy'), \tag{6.1}$$

$$\forall_{Y,x,y,x'} \sum_{X=0,1} p(XY|xy) = \sum_{X=0,1} p(XY|x'y). \tag{6.2}$$

Powyższy zapis jest stwierdzeniem faktu, że wybór różnych ustawień pomiarowych przez jednego użytkownika nie może mieć wpływu na statystykę wyników dowolnego pomiaru dokonywanego przez drugiego użytkownika. Jeśli układ nie spełnia powyższych warunków, wtedy nazywamy go *układem sygnalizującym* i dalej będziemy go oznaczać wykorzystując symbol „$*$”. Zwróćmy uwagę, że układy niesygnalizujące są szczególnym przypadkiem układów spójnych zdefiniowanych w rozdziale 4.

Układem PR (w skrócie *PR*) będziemy nazywać układ niesygnalizujący, który spełnia warunek

$$X \oplus Y = xy. \tag{6.3}$$

Tak scharakteryzowanemu układowi odpowiada rodzina rozkładów prawdopodobieństwa $\{p(XY|xy)\}$, która spełnia:

$$p(XY|xy) = \begin{cases} \frac{1}{2} & \text{dla } X \oplus Y = xy, \\ 0 & \text{w przeciwnym wypadku.} \end{cases} \tag{6.4}$$

Korelacjami PR będziemy nazywać przypisanie wejść x, y i wyjść X, Y danego układu, dla którego będzie spełniony warunek (6.3). Układ wejść/wyjść dla *PR* przedstawia Rys. 6.1.

Przykład.

Układ *PR* można wyrazić jako mieszaninę dwóch układów sygnalizujących:

$$PR = \frac{1}{2}B^* + \frac{1}{2}\bar{B}^*, \tag{6.5}$$

gdzie sygnalizujący układ B^* jest opisany przez rodzinę rozkładów prawdopodobieństwa:

$$\{p(00|00) = 1\}, \{p(11|01) = 1\}, \{p(00|10) = 1\}, \{p(10|11) = 1\}, \tag{6.6}$$

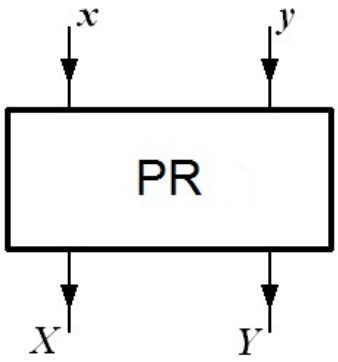
natomiast sygnalizujący układ $\bar{B}^*$ jest opisany przez rodzinę rozkładów prawdopodobieństwa:

$$\{p(11|00) = 1\}, \{p(00|01) = 1\}, \{p(11|10) = 1\}, \{p(01|11) = 1\}. \tag{6.7}$$

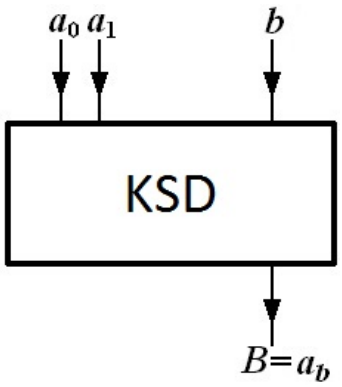
Zauważmy, że układ B^* (analogicznie $\bar{B}^*$):

- nie spełnia warunków niesygnalizacji (6.1); istotnie: niech przykładowo Alicja jako swoje wejście wybierze $x = 0$, wtedy statystyka wyników jej pomiaru (a wręcz sam wynik pomiaru!) jest ściśle skorelowany z wyborem wejścia dokonanym przez Bolka;
- spełnia *PR*-korelacje (6.3);
- łamie nierówność CHSH aż do maksymalnej wartości algebraicznej.

Rozpatrzmy teraz układ, który po stronie Alicji ma dwa wejścia, a_0 oraz a_1, natomiast po stronie Bolka jedno, b ($a_0, a_1, b \in \{0, 1\}$). Dodatkowo niech dany układ ma tylko jedno wyjście po stronie Bolka, B ($B \in \{0, 1\}$). Przypuśćmy, że intencją Bolka



Rys. 6.1: Schematyczne przedstawienie układu PR mającego dwa wejścia x, y oraz dwa wyjścia X, Y. Dla układu PR spełniony jest warunek $X \oplus Y = xy$. Źródło: [21].



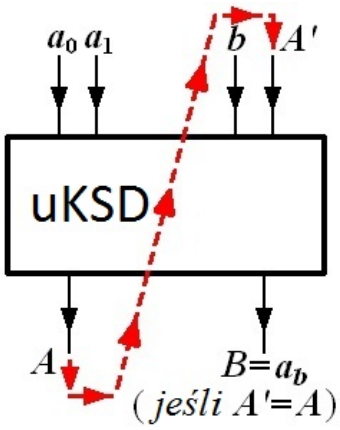
Rys. 6.2: Schematyczne przedstawienie kodu swobodnego dostępu (KSD) mającego dwa wejścia po stronie Alicji a_0, a_1, natomiast po stronie Boba jedno wejście b oraz jedno wyjście B. Dla KSD spełniony jest warunek $p(B = a_b|b) = 1$. Źródło: [21].

jest poznanie *jednego z dwóch* bitów wejścia po stronie Alicji, a_0 lub a_1. Dany układ nazywamy *kodem swobodnego dostępu* (KSD) jeśli spełniony jest warunek $B = a_b$, tj.

$$p(B = a_b|b) = 1, \tag{6.8}$$

dla wszystkich możliwych wejść a_0, a_1, b (por. Rys. 6.2).

Rozpatrzmy teraz układ, który po stronie Alicji ma dwa wejścia, a_0, a_1, oraz po stronie Bolka ma dwa wejścia, b,A' $(a_0, a_1, b, A' \in \{0, 1\})$. Dodatkowo niech dany układ ma tylko po jednym wyjściu po stronie Alicji i Bolka, odpowiednio A oraz B $(A, B \in \{0, 1\})$, gdzie zmienna A generowana jest losowo z prawdopodobieństwem wypadnięcia obu wyników $\frac{1}{2}$. Dany układ nazywamy *układem związanym z kodem swobodnego dostępu* $(uKSD)$ jeśli dla $A = A'$ dany układ działa jak KSD, tj. jeśli dla wszystkich możliwych wejść a_0, a_1, b, A' zachodzi $B = a_b$. Gdy $A \neq A'$ nie nakładamy żadnych warunków na działanie $uKSD$. Zauważmy, że $uKSD$ jest zaprojektowany w taki sposób, by przy dodatkowej komunikacji jednego bitu informacji od Alicji do



Rys. 6.3: Schematyczne przedstawienie układu związanego z kodem swobodnego dostępu ($uKSD$) mającego dwa wejścia po stronie Alicji a_0, a_1, dwa wejścia po stronie Bolka b,A' oraz po jednym wyjściu po obu stronach, odpowiednio A i B. $uKSD$ działa jak KSD, o ile wejście A' jest równe wyjściu A. W szczególności, jeśli Alicja prześle Bolkowi jeden bit informacji (A), to w przypadku gdy Bolek użyje A jako wejście A', otrzymamy $B = a_b$. Źródło: [21].

Bolka działał jak KSD (por. Rys. 6.3). Tym samym KSD nabiera następującego znaczenia: Alicja przesyłając jeden bit informacji do Bolka, daje mu możliwość dokładnego poznania jednego z dwóch bitów informacji, które zakodowała jako wejścia a_0 i a_1. Warto dodać, że taka funkcjonalność nie spełnia zasady informacyjnej przyczynowości [54].

Lemat 14. *Niech Bolek otrzymuje losowo wybrany bit b. Niesygnalizujący uKSD w przypadku $A \neq A'$ działa jak $anty$-KSD, tzn. spełniona jest wtedy relacja*

$$B = a_b \oplus 1. \tag{6.9}$$

Dowód.

W pierwszej kolejności rozpatrzmy prawdopodobieństwo zdarzenia polegającego na tym, że przy wyborze ustalonego b wyjście Bolka B jest równe a_b, tj. $p(B = a_b|b)$. Warunek niesygnalizacji wymaga, aby dla każdego wejścia b:

$$p(B = 0|b, a_b = 0) = p(B = 0|b, a_b = 1). \tag{6.10}$$

Tym samym dla każdego b mamy

$$\begin{aligned} p(B=a_b|b) &= p(B=0,a_b=0|b)+p(B=1,a_b=1|b) \\ &= p(a_b=0|b)p(B=0|b,a_b=0)+p(a_b=1|b)p(B=1|b,a_b=1) \\ &= \frac{1}{2}\left(p(B=0|b,a_b=0)+p(B=1|b,a_b=1)\right) \\ &= \frac{1}{2}\left(p(B=0|b,a_b=1)+1-p(B=0|b,a_b=1)\right) \\ &= \frac{1}{2}, \end{aligned} \tag{6.11}$$

gdzie w trzeciej równości wykorzystaliśmy fakt losowości wejść a_0, a_1 oraz ich niezależności od wejścia b $p(a_b=0|b)=p(a_b=1|b)=\frac{1}{2}$, a w czwartej warunek niesygnalizacji (6.10).

Niech teraz Bolek jako wejście A' wybierze z prawdopodobieństwem $\frac{1}{2}$ dowolną wartość, 0 lub 1. Ze względu na warunek niesygnalizacji, wybór ustawień Bolka nie może mieć wpływu na wyniki otrzymywane po stronie Alicji, a zatem wyjście A musi być niezależne od wejścia A' oraz b. Tym samym mamy:

$$p(A=A'|b)=p(A\neq A'|b)=\frac{1}{2}. \tag{6.12}$$

Możemy teraz napisać w dalszej kolejności

$$\begin{aligned} p(B=a_b|b) &= p(A=A'|b)p(B=a_b|A=A',b)+p(A\neq A'|b)p(B=a_b|A\neq A',b) \\ &= \frac{1}{2}\left(p(B=a_b|A=A',b)+p(B=a_b|A\neq A',b)\right). \end{aligned} \tag{6.13}$$

Jeśli dla $A=A'$ $uKSD$ działa jak KSD, to:

$$p(B=a_b|A=A',b)=1. \tag{6.14}$$

Tym samym, by uniknąć możliwości sygnalizacji, musimy mieć:

$$p(B=a_b|A\neq A',b)=0, \tag{6.15}$$

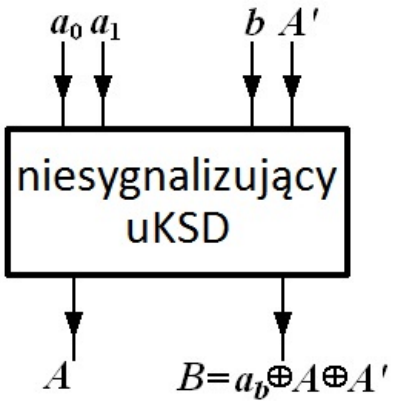
a zatem prawdopodobieństwo zdarzenia przeciwnego $p(B\neq a_b|A\neq A',b)=1$, tj.:

$$B=a_b\oplus 1. \blacksquare \tag{6.16}$$

Zauważmy, że powyższa charakteryzacja niesygnalizującego $uKSD$ pozwala nam zapisać wyjście po stronie Bolka jako:

$$B=a_b\oplus A\oplus A', \tag{6.17}$$

co przedstawia Rys. 6.4



Rys. 6.4: Niesygnalizujący $uKSD$ spełnia warunek $B = a_b \oplus A \oplus A'$: dla $A = A'$ działa jak KSD, natomiast dla $A \neq A'$ działa jak $anty\text{-}KSD$. Źródło: [21].

Rozważmy też szczególny przypadek $uKSD$, który jest zdefiniowany w sposób następujący. Niech dla $A = A'$ dany układ działa jak KSD, tj. jeśli dla wszystkich pozostałych wejść a_0, a_1, b zachodzi $B = a_b$. W przeciwnym przypadku, tj. gdy $A \neq A'$ niech wyjście B będzie losowym bitem nieskorelowanym z żadną inną zmienną. Oczywistym jest, że pierwszy warunek zapewnia, że jest to $uKSD$. Zauważmy jednak, że ze względu na drugi warunek dla ustalonej wartości b, przy założeniu, że Bolek wybiera A' niezależnie od A (oraz niezależnie od b), mamy (przykładowo dla $b = 0$)

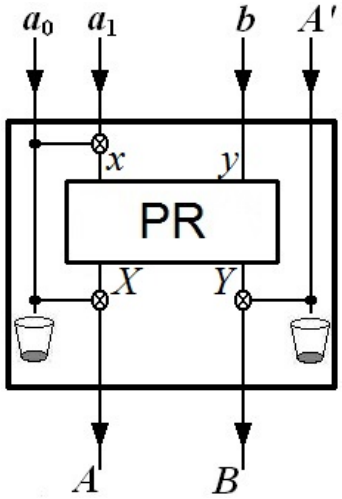
$$\begin{aligned} & p(B = a_0 | b = 0) \\ & = p(A = A' | b = 0) p(B = a_0 | b = 0, A = A') + p(A \neq A' | b = 0) p(B = a_0 | b = 0, A \neq A') \\ & = \frac{1}{2} \times 1 + \frac{1}{2} \times \frac{1}{2} \\ & = \frac{3}{4}, \end{aligned} \tag{6.18}$$

w sprzeczności z warunkiem (6.11) dla niesygnalizującego rozkładu prawdopodobieństwa $p(B = a_b | b)$. Tym samym, tak zdefiniowany $uKSD$ jest *sygnalizujący* i będziemy oznaczać przez $uKSD^*$.

6.3. Równoważność układu PR i niesygnalizującego $uKSD$

Jak pokazano w pracach [52–54], niesygnalizujący $uKSD$ można symulować za pomocą układu PR w sposób przedstawiony na Rys. 6.5. Istotnie, dla układu PR zawsze spełniony jest warunek PR-korelacji

$$X \oplus Y = xy, \tag{6.19}$$



Rys. 6.5: Symulacja niesygnalizującego $uKSD$ przez układ PR. Symbol „⊗" oznacza tutaj wykonanie operacji kontrolowanej negacji odpowiedniego bitu (C-NOT). Źródło: [21].

skąd wyjście układu PR po stronie Bolka spełnia

$$Y = xy \oplus X. \tag{6.20}$$

Stosując zaś wejścia pierwotne układu (por. Rys. 6.5) mamy:

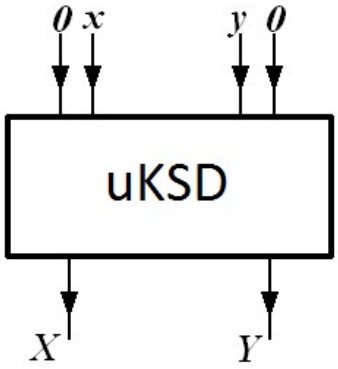
$$Y = (a_0 \oplus a_1)b \oplus X. \tag{6.21}$$

Dla wyjścia wtórnego układu po stronie Bolka dostajemy:

$$\begin{aligned} B &= Y \oplus A' \\ &= (a_0 \oplus a_1)b \oplus X \oplus A' \\ &= a_0 b \oplus a_1 b \oplus X \oplus A' \\ &= a_0(b \oplus 1) \oplus a_1 b \oplus A \oplus A' \\ &= a_b \oplus A \oplus A', \end{aligned} \tag{6.22}$$

gdzie w czwartej linii skorzystaliśmy z przypisania $A = X \oplus a_0$ (C-NOT z Rys. 6.5), natomiast w piątej z tożsamości $a_b \equiv a_0(b \oplus 1) \oplus a_1 b$.

Zastanówmy się teraz w jaki sposób możemy symulować układ PR przez niesygnalizujący $uKSD$ [21]. Zauważmy, że w tym celu możemy wykorzystać schemat symulacji przeciwnej, tj. symulacji niesygnalizującego $uKSD$ przez układ PR przedstawiony na Rys. 6.5. Istotnie, kładąc $a_0 = 0$ oraz $A' = 0$ dla wejść $uKSD$ będących jednocześnie bitami kontrolnymi przy odpowiednich operacjach C-NOT, odzyskujemy poprawne działanie układu PR, przy czym odpowiedni warunek PR-korelacji będzie



Rys. 6.6: Symulacja układu PR przez niesygnalizujący $uKSD$. Odpowiednie wejścia i wyjścia $uKSD$ zostały dobrane w taki sposób $(a_0 = 0, a_1 = x, b = y, A' = 0)$, by spełniony był warunek PR-korelacji w postaci (6.3). Źródło: [21].

miał teraz postać:

$$A \oplus B = a_1 b. \tag{6.23}$$

W tym miejscu musimy sprawdzić, czy powyższy warunek spełniony jest dla dowolnych wartości wejść a_1 oraz b. Zgodnie z Lematem 14, dla niesygnalizującego $uKSD$ możemy napisać:

$$A \oplus a_b \oplus A \oplus A' = a_1 b. \tag{6.24}$$

W przypadku omawianej symulacji (dla ustalonych $A' = 0$ oraz $a_0 = 0$) otrzymujemy warunek:

$$a_b = a_1 b, \tag{6.25}$$

który jest prawdziwy dla dowolnych wartości a_1 oraz b. Tym samym przedstawiony protokół pozwala na symulację układu PR przez niesygnalizujący $uKSD$, co przedstawia Rys. 6.6.

Wykazując możliwość symulacji niesygnalizującego $uKSD$ przez układ PR, oraz przeciwnie, układu PR przez niesygnalizujący $uKSD$, wykazaliśmy tym samym równoważność obu zasobów.

6.4. Nierówność wiążąca zasoby PR i $uKSD$

Jak wynika z poprzedniego rozdziału, zachodzi równoważność układu PR z *niesygnalizującym* $uKSD$. Powstaje naturalne pytanie, czy za pomocą *dowolnej* funkcjonalności działającej jak $uKSD$ (niekoniecznie niesygnalizującej) możemy symulować układ PR. Jak się okazuje, układ PR nie można symulować poprzez wykorzystanie dowolnych sygnalizujących $uKSD$. W ogólności jednak, dla dowolnych układów związa-

nych z kodem swobodnego dostępu zachodzi następująca nierówność wiążąca różne zasoby:

$$uKSD + 1C\text{-bit} + 1L\text{-bit} \geq PR + \mathscr{E}, \tag{6.26}$$

gdzie C-bit oznacza bit klasycznej informacji, L-bit oznacza bit współdzielonej losowości, natomiast $\mathscr{E}$ jest tzw. kanałem wymazującym

$$\mathscr{E}(z) = \varepsilon|0\rangle\langle 0|_f \otimes |\mathbb{I}\rangle\langle\mathbb{I}| + (1-\varepsilon)|1\rangle\langle 1|_f \otimes |z\rangle\langle z|, \tag{6.27}$$

gdzie przez z oznaczyliśmy przesyłaną informację, $\mathbb{I} = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$ oznacza szum, natomiast ε jest prawdopodobieństwem wymazywania. Kanał wymazujący jest sygnowany dwoma etykietami („flagami") $|0\rangle\langle 0|_f$, $|1\rangle\langle 1|_f$ mówiącymi o wyborze wejścia dokonanym przez Bolka.

Udowodnimy nierówność w postaci

$$KSD + 1L\text{-bit} \geq PR + \mathscr{E}, \tag{6.28}$$

która implikuje nierówność (6.26), ponieważ z definicji $uKSD$ mamy, że

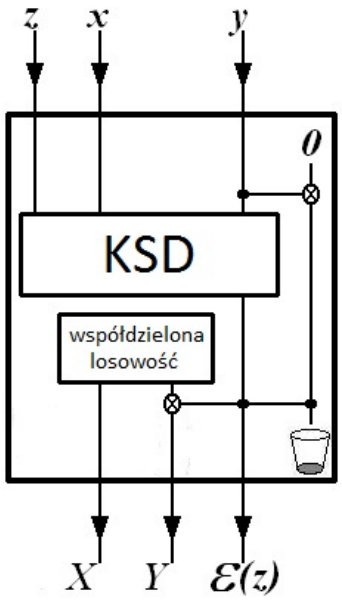
$$uKSD + 1C\text{-bit} \geq KSD. \tag{6.29}$$

Zauważmy, że celem uzyskania PR-korelacji $X \oplus Y = xy$ w przypadku gdy $y = 0$ wystarczy skorzystać ze współdzielonej losowości, ponieważ warunek PR-korelacji wymaga tutaj jedynie by wyjścia po stronie Alicji i Bolka (X i Y) były identyczne. Tym samym wykorzystując L-bit do otrzymania PR-korelacji, KSD możemy wykorzystać celem przesłania informacji z. W przypadku gdy $y = 1$, Bolek musi dodatkowo znać wartość x by wiedzieć, czy wyjścia są zgodne (wtedy mógłby ponownie użyć współdzielonej losowości) czy przeciwne (wtedy powinien zanegować swój bit ze współdzielonej losowości). W celu poznania wartości x Bolek powinien zatem wykorzystać KSD, który nie będzie użyty celem przesłania dodatkowej informacji z.

Rozpatrzmy konkretny protokół, dzięki któremu możemy uzyskać to co opisano powyżej (por. Rys. 6.7). Kodowanie odbywa się w sposób następujący: Dysponując kodem swobodnego dostępu KSD, Alicja koduje dodatkową informację z, którą chce przesłać do wejścia a_0, natomiast wejście układu PR, x, do wejścia a_1, z kolei Bolek koduje wejście układu PR, y, do swojego wejścia b, tj.

$$\begin{aligned} a_0 &= z, \\ a_1 &= x, \\ b &= y. \end{aligned} \tag{6.30}$$

Poza KSD Alicja i Bolek dysponują współdzieloną losowością. W przypadku, gdy $y = 0$ Bolek wykorzystuje L-bit bez żadnych dodatkowych operacji i, co zostało już powiedziane wcześniej, wyjścia X oraz Y będą takie same, tj. PR-korelacje zostaną zachowane. W przypadku, gdy $y = 1$ Bolek dokonuje operacji C-NOT na swoim wyj-

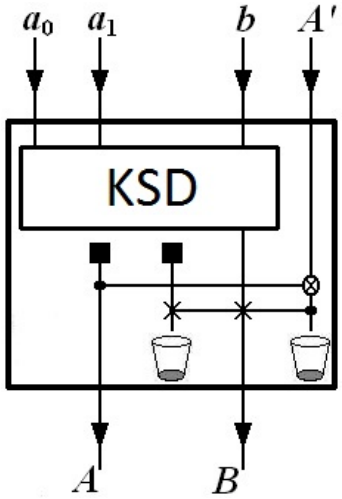


Rys. 6.7: Protokół kodowania wykorzystany do wykazania nierówności zasobowej (6.28). Bit informacji, który ma być przesłany od Alicji do Bolka oznaczony jest jako z. Kanał $\mathcal{E}$ jest kanałem wymazującym z prawdopodobieństwem wymazywania $\varepsilon = p(y = 1)$: z prawdopodobieństwem ε wiadomość z jest tracona, natomiast z prawdopodobieństwem $1-\varepsilon$ wiadomość z jest dostarczana niezmieniona, przy czym odbiorca (Bolek) wie która sytuacja zaszła. Wejścia x, y oraz wyjścia X, Y spełniają warunek PR-korelacji $X \oplus Y = xy$. Źródło: [21].

ściu KSD, B, oraz na bicie pochodzącym ze współdzielonej losowości, gdzie bitem kontrolnym jest B, natomiast bitem docelowym jest bit pochodzący ze współdzielonej losowości. Zobaczmy, że znów PR-korelacje zostaną zachowane. Istotnie, z definicji KSD mamy, że dla $y = 1$ otrzymujemy $B = a_1 \equiv x$. Tym samym, jeśli $x = 0$ to docelowy bit nie ulega zmianie i wyjścia X oraz Y są takie same, natomiast jeśli $x = 1$ to docelowy bit po stronie Bolka zostaje zmieniony na przeciwny, i wtedy wyjścia X oraz Y są różne, czego dokładnie wymagają PR-korelacje. Widzimy zatem, że dysponując kodem swobodnego dostępu oraz współdzieloną losowością możemy z wykorzystaniem tego protokołu zasymulować działanie układu PR.

Sprawdźmy teraz w jakim stopniu przy danym protokole można rownocześnie przesyłać dodatkową informację z od Alicji do Bolka. Kiedy $y = 0$, to wyjście KSD po stronie Bolka, B, jest równe $a_0 = z$, a zatem wiadomość zostaje przesłana dokładnie. Jeśli natomiast $y = 1$, to wyjście B jest równe $a_1 = x$, a zatem wiadomość jest tracona. Tym samym, korzystając z takiego protokołu otrzymujemy dodatkowo kanał wymazujący $\mathcal{E}(z)$ z prawdopodobieństwem wymazywania $\varepsilon = p(y = 1)$, gdzie $p(y = 1)$ to prawdopodobieństwo wybrania przez Bolka wejścia $y = 1$.

Poniżej pokażemy, że niektóre sygnalizujące układy związane z kodem swobodnego dostępu $uKSD^*$ są w pewnym sensie mniej użyteczne (słabsze) niż niesygnali-



Rys. 6.8: Przykład sygnalizującego $uKSD^*$ użytego do wykazania nierówności zasobowej (6.31). Symbol „■" oznacza generację losowego bitu, natomiast bramka z symbolami „×" jest bramką kontrolowanej wymiany odpowiednich bitów. W przypadku gdy $A = A'$, wymiana bitów nie następuje i wyjściem z układu po stronie Bolka jest wyjście KSD, tym samym przedstawiony $uKSD^*$ działa jak KSD. Jeśli natomiast $A \neq A'$, wtedy wyjściem z układu po stronie Bolka jest losowy bit. Źródło: [21].

zujące jeśli chcemy je wykorzystać w celu symulacji układu PR. Okazuje się, że celem symulacji układu PR w tym przypadku potrzebowalibyśmy dodatkowo przynajmniej $\frac{1}{2}$ bitu informacji. Fakt ten możemy sformułować w postaci poniższego twierdzenia:

Twierdzenie 13. *Niech wejścia x i y generowane są z jednorodnym rozkładem prawdopodobieństwa. Załóżmy, że dla $uKSD^*$ zachodzi następująca nierówność:*

$$uKSD^* + 1C\text{-}bit \geq PR + \Lambda, \tag{6.31}$$

gdzie Λ jest dowolnym kanałem możliwym do uzyskania w tej symulacji. Wtedy kanał Λ można otrzymać poprzez przetworzenie kanału wymazującego $\mathscr{E}$ z prawdopodobieństwem wymazywania $\varepsilon = \frac{1}{2}$.

Na Rys. 6.8 schematycznie przedstawiony został sygnalizujący układ $uKSD^*$ służący do wykazania nierówności (6.31). Teza powyższego twierdzenia natomiast implikuje nierówność (6.28), gdyż z definicji mamy, że $uKSD^* + 1C\text{−bit} \geq KSD$.

W celu udowodnienia Twierdzenia 13 rozważymy 2 przypadki, przy czym za każdym razem narzucamy warunek otrzymania PR-korelacji: (1) bit komunikacji nie zostaje użyty do przesłania informacji o wyjściu A po stronie Alicji; (2) bit komunikacji zostaje użyty do przesłania informacji o wyjściu A.

W pierwszej kolejności sformułujmy dwa lematy, które będą wykorzystane w dowodzie.

Lemat 15. *Niech z pewnym ustalonym rozkładem prawdopodobieństwa $p(x, y)$ Alicja otrzymuje bit x a Bolek bit y. Niech dodatkowo Bolek otrzymuje zmienną losową B,*

które może być skorelowane zarówno z x jak i y. Poza tym, Alicja i Bolek nie współdzielą żadnych innych zasobów. Wtedy, chcąc otrzymać PR-korelacje postaci (6.3), dla każdej wartości B prawdopodobieństwo warunkowe $p(xy|B)$ musi wynosić 0 dla pewnej pary xy.

Dowód.

Bez straty ogólności przyjmijmy początkowo, że $B = 0$ (przypadek $B = 1$ dowodzimy analogicznie). Załóżmy przeciwnie, że wszystkie możliwe wyniki (x, y) zachodzą z niezerowym prawdopodobieństwem. Oznaczmy minimalne prawdopodobieństwo w rozkładzie $p(x, y|B = 0)$ przez:

$$p^* = \min_{x,y} p(x, y|B = 0). \tag{6.32}$$

Zauważmy, że rozkład prawdopodobieństwa $p(x, y)$ można zapisać jako kombinację dwóch rozkładów prawdopodobieństwa w następujący sposób:

$$p(x, y) = 4p^* P_{\text{j}} + (1 - 4p^*)P_{\text{nj}}, \tag{6.33}$$

gdzie P_{j} jest rozkładem jednorodnym

$$P_{\text{j}} = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}, \tag{6.34}$$

natomiast rozkład P_{nj} jest zdefiniowany jako

$$P_{\text{nj}} = \{\frac{p(0,0) - p^*}{1 - 4p^*}, \frac{p(0,1) - p^*}{1 - 4p^*}, \frac{p(1,0) - p^*}{1 - 4p^*}, \frac{p(1,1) - p^*}{1 - 4p^*}\}. \tag{6.35}$$

Zwróćmy uwagę, że dla rozkładu jednorodnego P_{j}

$$p(x, y) = p(x)p(y), \tag{6.36}$$

a zatem nie ma żadnych korelacji między zmiennymi x i y.

Prawdopodobieństwo symulacji *PR*-korelacji (a zatem również możliwość maksymalnego łamania nierówności CHSH) nie może zmaleć jeśli Alicja i Bolek otrzymają informacje o rozkładzie prawdopodobieństwa pary zmiennych (x, y). Jeśli jednak z pewnym prawdopodobieństwem (p^*) otrzymują oni jednorodny rozkład prawdopodobieństwa zmiennych (x, y) i wciąż mają możliwość symulacji *PR*-korelacji, to w tym przypadku mogliby maksymalnie złamać nierówność CHSH. Nie jest to jednak możliwe, gdyż Alicja i Bolek nie współdzielą żadnych innych zasobów poza znajomością rozkładu prawdopodobieństwa wejść x, y, które to wielkości w tym przypadku są nieskorelowane ze sobą [98]. ■

Kolejny lemat umożliwia ograniczenie rozważań do strategii deterministycznych.

Lemat 16. *Rozważmy trzy niezależne zmienne x, y oraz z. Załóżmy, że Alicja oraz Bolek mają dostęp do zmiennych ℓ_A i ℓ_B z rozkładem prawdopodobieństwa $p(\ell_A, \ell_B)$, który jest niezależny od x, y, z. Niech Alicja ze zmiennych x i z generuje dwa bity a_0 i a_1, które używa jako wejść do KSD, natomiast Bolek niech ze zmiennej y generuje bit b, który używa jako wejścia do KSD.*

Dowolny kanał $z \to B$ można otrzymać przez kombinację wypukłą kanałów otrzymywanych jako deterministyczne przetwarzanie $(x,z) \to (a_0, a_1)$ dla pewnego ℓ_A oraz $y \to b$ dla pewnego ℓ_B.

Dowód.

Generację pary bitów (a_0, a_1) ze zmiennych x i z przez Alicję możemy rozważyć jako zastosowanie przez nią dwuwejściowego kanału

$$(a_0, a_1) = \Lambda^A_{\ell_A}(x,z) := \sum_i q_i \lambda^A_i(x,z), \tag{6.37}$$

takiego, że jest on kombinacją wypukłą deterministycznych kanałów $\lambda^A_i(x,z)$. Podobnie w przypadku Bolka, generację bitu b ze zmiennej y możemy rozważyć jako zastosowanie przez niego kanału

$$b = \Lambda^B_{\ell_B}(y) := \sum_j r_j \lambda^B_j(y), \tag{6.38}$$

takiego, że jest on kombinacją wypukłą deterministycznych kanałów $\lambda^B_j(y)$. Po wybraniu odpowiednich bitów a_0, a_1, b jako wejście do KSD, wyjściem po stronie Bolka jest

$$\begin{aligned}
B &= \Lambda_{KSD}(a_0, a_1, b) \\
&= \sum_{\ell_A, \ell_B} p(\ell_A, \ell_B) \Lambda_{KSD}\left(\Lambda^A_{\ell_A}(x,z), \Lambda^B_{\ell_B}(y)\right) |\ell_A, \ell_B\rangle\langle \ell_A, \ell_B| \otimes |y\rangle\langle y| \\
&= \sum_{\ell_A, \ell_B, i, j} p(\ell_A, \ell_B) q_i r_j \Lambda_{KSD}\left(\lambda^A_i(x,z), \lambda^B_j(y)\right) |\ell_A, \ell_B\rangle\langle \ell_A, \ell_B| \otimes |y\rangle\langle y| \\
&= \sum_{\ell_A, \ell_B, i, j} s_{\ell_A, \ell_B, i, j} \tilde{\lambda}_{i,j}(x,z,y) |\ell_A, \ell_B\rangle\langle \ell_A, \ell_B| \otimes |y\rangle\langle y| \\
&=: \tilde{\Lambda}(x,z,y),
\end{aligned} \tag{6.39}$$

gdzie wprowadziliśmy oznaczenie $s_{\ell_A, \ell_B, i, j} = p(\ell_A, \ell_B) q_i r_j$.■

Zauważmy przy tym, że możliwość korzystania ze współdzielonej losowości jest szczególnym przypadkiem powyższej strategii przy $\ell_A = \ell_B$. Tym samym dowodząc Twierdzenie 13 możemy pominąć bit współdzielonej losowości $1L-$bit występujący w nierówności (6.28). Dodatkowo, spośród możliwych strategii wystarczy ograniczyć się jedynie do tych, które prowadzą do *PR*-korelacji. Jeśli bowiem dla pewnego ℓ^* odpowiednia strategia nie prowadziłaby do *PR*-korelacji, to również strategia mieszana nie mogłaby ich wygenerować.

Przystąpmy teraz do rozpatrzenia dwóch wspomnianych wcześniej przypadków w celu udowodnienia Twierdzenia 13.

Dowód Twierdzenia 13.

Przypadek (1).

Pokażemy, że jeśli bit komunikacji nie zostaje użyty do przesłania informacji o wyjściu A po stronie Alicji, wtedy kanał Λ w (6.31) jest w najlepszym przypadku kanałem wymazującym z prawdopodobieństwem wymazywania $\varepsilon = \frac{1}{2}$.

Oznaczmy przez m bit informacji o wyniku A przesyłany do Bolka. Celem jest uzyskanie *PR*-korelacji $Y = X \oplus xy$ w każdym przypadku, tj. $m = 0$ oraz $m = 1$. Ponieważ wyjście Y jest generowane w procesie korzystania z KSD, dla każdego m wartość Y w ogólności zależy od wejść i wyjść KSD po stronie Bolka, oraz wszystkich zmiennych, którymi dysponuje:

$$Y = Y(y, b, B). \tag{6.40}$$

Teraz, dla ustalonej wartości $m = m_0$ mamy dwie możliwości: albo $A = A'$, albo $A \neq A'$. W pierwszym przypadku *PR*-korelacje można uzyskać z poprawnie działającego KSD. Natomiast w drugim przypadku, tj. gdy $A \neq A'$, sygnalizujący $uKSD$ zwróci Bolkowi losową wartość B, która nie zależy od działania KSD (por. Rys. 6.8). Z uwagi na to, że b jest otrzymywane z y, wartość Y można otrzymać jedynie poprzez przetwarzanie y:

$$Y = Y(y). \tag{6.41}$$

W przeciwnym przypadku Y byłby z niezerowym prawdopodobieństwem niezależny od (x, X) czyli z pewnym prawdopodobieństwem *PR*-korelacje nie byłyby zachowane. Chcąc uzyskać *PR*-korelacje, Bolek musi mieć pewność, że

$$\begin{cases} Y(y = 0) = X, \\ Y(y = 1) = X \oplus x, \end{cases} \tag{6.42}$$

co pozwala na otrzymanie wartości x z dodania $Y(y = 0)$ oraz $Y(y = 1)$. Widzimy zatem, że w przypadku $A \neq A'$ wartość x musi być znana Bolkowi, by móc otrzymać *PR*-korelacje.

Pokazaliśmy powyżej, że dla ustalonego $m = m_0$ w najgorszym przypadku, by móc odtworzyć *PR*-korelacje, Bolek musi znać albo wartość x, albo A (może też znać wartość obu tych wielkości na raz). Jeśli przez $p_{\text{zg}}(Z)$ oznaczymy *prawdopodobieństwo odgadnięcia* przez Bolka pewnej wielkości Z, to dostaniemy

$$p_{\text{zg}}(x|m = m_0) = 1, \tag{6.43}$$

lub

$$p_{zg}(A|m = m_0) = 1, \tag{6.44}$$

(oba warunki niekoniecznie muszą być wykluczające). Możemy teraz założyć, że obie wartości m Bolek może otrzymać z niezerowym prawdopodobieństwem (w przeciwnym razie, tj. gdyby przykładowo otrzymywał zawsze $m = 0$, wtedy z prawdopodobieństwem równym $\frac{1}{2}$ zachodziłoby $A \neq A'$, co nie pozwalałoby na spełnienie żadnego z powyższych warunków (6.43) i (6.44)). Tym samym, dla dwóch różnych wartości $m = 0, 1$ najprostsza strategia Bolka polegająca na prawidłowym odgadnięciu tylko jednej zmiennej x bądź A sprowadzałaby się do czterech przypadków:

(a) $p_{zg}(A|m = 0) = 1$ oraz $p_{zg}(A|m = 1) = 1$,

(b) $p_{zg}(x|m = 0) = 1$ oraz $p_{zg}(x|m = 1) = 1$,

(c) $p_{zg}(A|m = 0) = 1$ oraz $p_{zg}(x|m = 1) = 1$,

(d) $p_{zg}(x|m = 0) = 1$ oraz $p_{zg}(A|m = 1) = 1$.

(a) Warunki oznaczają łącznie, że Bolek prawidłowo odgaduje wartość A dla każdej otrzymanej wiadomości m od Alicji. Jak pokażemy poniżej, z tego faktu wynika, że wiadomość m musiała zostać użyta do zakomunikowania wartości A. Aby to wykazać, w pierwszej kolejności wprowadzimy pojęcie prawdopodobieństwa zgadnięcia $p_{zg}(A)$ zdefiniowane w następujący sposób:

$$\begin{aligned} p_{zg}(A) &= p(G = A) \\ &= p(A = 0)p(G = 0|A = 0) + p(A = 1)p(G = 1|A = 1), \end{aligned} \tag{6.45}$$

gdzie G określa zmienną skorelowaną z wielkością zgadywaną A. Mówimy, że Bolek zgaduje wartość zmiennej A jeśli istnieje pewna zmienna G, która jest skorelowana ze zmienną A w następujący sposób:

$$p(G = 0|A = 0) = p(G = 1|A = 1) = 1. \tag{6.46}$$

Wtedy też $p_{zg}(A) = p(G = A) = 1$.

Teraz, jeśli Bolek prawidłowo odgaduje wartość A (tj. dysponuje odpowiednią zmienną G), to w ogólności jest to możliwe (i) bez wykorzystania KSD, bądź też (ii) po uprzednim wykorzystaniu KSD.

Rozpatrzmy przypadek (i). Odgadnięcie zmiennej A oznacza, że wzajemna informacja $I(G : A)$ przyjmuje maksymalną wartość. Ze względu na fakt, że zmienna G mogła zostać wygenerowana jedynie ze zmiennych m, b, y, mamy, że

$$I(m, b, y : A) \geq I(G : A), \tag{6.47}$$

co wynika bezpośrednio z nierówności przetwarzania danych [95]. Z reguły łańcuchowej dla wzajemnych informacji mamy natomiast

$$I(m, b, y : A) = I(b, y : A) + I(m : A|b, y) = I(m : A), \tag{6.48}$$

gdzie druga równość wynika z faktu, że zmienne b i y nie są skorelowane z A. Widzimy tym samym, że skoro $I(G : A)$ przyjmuje maksymalną wartość to tym bardziej również $I(m : A)$ przyjmuje maksymalną wartość, co sprawia, że m musi być maksymalnie skorelowane z A.

Rozpatrzmy przypadek (ii). Załóżmy, że zgadnięcie A dokonane było z wykorzystaniem wiadomości m i wyjścia B z KSD, tzn. zmienna G może zależeć od zmiennych m, b, y, B. Ze względu na fakt, że KSD z pewnym niezerowym prawdopodobieństwem zwraca losową wartość B (co wynika z braku uprzedniej pełnej wiedzy o wartości A) wartość G nie może zależeć od zmiennej B. W przeciwnym razie zmienna G z pewnym prawdopodobieństwem nie byłaby skorelowana z A. Tym samym zmienna G może zależeć jedynie od m, b, y, który to przypadek został rozpatrzony w (i).

Rozpatrzywszy powyższe przypadki wnioskujemy, że wiadomość m została użyta do przesłania wartości A. To z kolei pozwala na poprawne działanie KSD równocześnie uniemożliwiając poznanie wartości x potrzebnej do zapewnienia warunku PR-korelacji.

(b) Warunki oznaczają łącznie, że Bolek prawidłowo odgaduje wartość x dla każdej otrzymanej wiadomości m od Alicji, tj. dysponuje odpowiednią zmienną G skorelowaną ze zmienną x. Jak pokażemy niżej, z tego faktu wynika, że wiadomość m musiała zostać użyta do przesłania wartości x. Aby to wykazać, podobnie jak w punkcie (a) musimy rozpatrzyć dwa przypadki, gdzie zakładamy, że w ogólności jest to możliwe (i) bez wykorzystania KSD, bądź też (ii) po uprzednim wykorzystaniu KSD.

Przypadek (i) rozpatrujemy analogicznie jak w punkcie (a), przy czym tutaj rozpatrujemy zmienną x zamiast A. Rozpatrzmy zatem przypadek (ii). Gdy wartość A nie jest w pełni znana, argumentacja jest analogiczna jak w punkcie (a). Pozostaje jedynie możliwość, gdy znana jest wartość zarówno zmiennej A, jak również x, tzn. Bolek dysponuje zmiennymi odpowiednio G_1 i G_2 zależnymi w ogólności od m, b, y, B. Zauważmy, że ze względu na fakt, że KSD z pewnym niezerowym prawdopodobieństwem zwraca losową wartość B (co wynika z braku uprzedniej pełnej wiedzy o wartości A) wartość G_1 (G_2) nie może zależeć od zmiennej B. W przeciwnym razie zmienna G_1 (G_2) z pewnym prawdopodobieństwem nie byłaby skorelowana z A (x). Tym samym zmienna G_1 (G_2) może zależeć jedynie od m, y (gdzie wykorzystaliśmy fakt, że zmienna b zależy wyłącznie od y). Zauważmy teraz, że skoro wartość x jest znana,

to zachodzi $I(G_2 : x) = 1$. Biorąc dodatkowo pod uwagę $I(G_1 : A)$, mamy:

$$\begin{aligned} I(G_1 : A) + I(G_2 : x) &\leq I(m, y : A) + I(m, y : x) \\ &\leq I(y : A) + I(m : A|y) + I(y : x) + I(m : x|y) \\ &= I(m : A|y) + I(m : x|y) \\ &\leq I(m : A, x|y) + I(x : A|y) \\ &= I(m : A, x|y) \\ &\leq \min\{H(m|y), H(x, A|y)\} \\ &\leq H(m|y) \\ &\leq 1, \end{aligned} \tag{6.49}$$

gdzie w pierwszej linii skorzystaliśmy z nierówności przetwarzania danych, w trzeciej linii z niezależności x oraz A od y, w czwartej linii wykorzystaliśmy silną subaddytywność entropii, a w piątej fakt, że zmienne x oraz A są niezależne od siebie. Z nierówności (6.49) wynika, że $I(G_1 : A) = 0$, tj. pełna znajomość x implikuje pełną nieznajomość A. Zatem, jakkolwiek można symulować *PR*-korelacje, nie można otrzymać równocześnie poprawnie działającego KSD. W tym przypadku otrzymywany kanał Λ jest kanałem wymazującym z prawdopodobieństwem wymazywania $\varepsilon = \frac{1}{2}$ (ze względu na działanie wyłącznie $uKSD^*$).

(c) W zależności od wartości m Bolek poprawnie odgaduje wartość A ($p_{zg}(A|m = 0) = 1$) bądź x ($p_{zg}(A|m = 0) = 1$). Zauważmy jednak, że dana sytuacja nie może mieć miejsca, o ile rozkład prawdopodobieństwa $p(A, x)$ miałby być jednorodny, jak w przypadku $uKSD$, gdzie wartości x i A generowane są losowo. Niech przykładowo Bolek z prawdopodobieństwem równym 1 zgaduje wartość $A = 1$ w przypadku $m = 0$ oraz z prawdopodobieństwem równym 1 zgaduje wartość $x = 1$ w przypadku $m = 1$. Widzimy, że prawdopodobieństwo $p(A = 0, x = 0) = 0$, ponieważ dla dowolnego m jedna z wielkości x i A musi być równa 1. Tym samym nie istnieje łączny rozkład prawdopodobieństwa dla trzech zmiennych $p(A, x, m)$, który spełniałby obydwa warunki na raz.

(d) Przypadek ten rozpatrujemy analogicznie jak przypadek (c).

Podsumujmy powyższe przypadki. Widzimy, że: albo (a) niemożność poznania x uniemożliwia uzyskanie *PR*-korelacji, albo (b) otrzymujemy *PR*-korelacje z dodatkowym kanałem wymazującym (z prawdopodobieństwem wymazywania $\varepsilon = p(y = 1)$) uzyskanym z częściowo działającego KSD.

Przypadek (2).

Pokażemy, że jeśli bit komunikacji zostaje użyty do przesłania informacji o wyjściu A po stronie Alicji, możliwe powstałe kanały Λ można otrzymać z kanału wymazującego z prawdopodobieństwem wymazywania $\varepsilon = \frac{1}{2}$.

Musimy rozważyć dwie możliwe sytuacje: (2a) wartość b nie zależy od y; (2b) wartość b zależy od y. Z Lematu 16 wiemy, że obie sytuacje możemy rozpatrywać całkowicie oddzielnie, tzn. z pominięciem możliwych strategii mieszanych, gdzie b może zależeć od y tylko częściowo (przykładowo z prawdopodobieństwem $\frac{1}{2}$ $b = \bar{y}$, a z prawdopodobieństwem $\frac{1}{2}$ b jest całkowicie losowe).

(2a). Zauważmy, że jeśli b jest niezależne od y, to możemy przyjąć, że $b = \text{const}$, dzięki czemu KSD funkcjonuje jak kanał binarny transmitując odpowiednią informację a_b. Jeśli jednak wymuszamy otrzymanie warunku PR-korelacji, wtedy Bolek musi znać wartość x [98]. Z drugiej strony, jeśli Alicja użyje kanału binarnego w celu przesłania wartości x, wtedy nie jest możliwe równoczesne przesłanie informacji z za pomocą tego samego kanału. Tym samym, możliwy kanał Λ musi mieć zerową pojemność (por. (6.49)).

(2b). Bez straty ogólności możemy przyjąć, że $b = y$ (przypadek $b = \bar{y}$ może być rozpatrzony analogicznie). Skorzystajmy z Lematu 15. Zakładając, że dla każdej wartości B można otrzymać PR-korelacje, mamy że dla $B = 0$ jedynie trzy z czterech możliwych zdarzeń $xyB = 000, 010, 100, 110$ mogą mieć miejsce i podobnie, dla $B = 1$ jedynie trzy z czterech możliwych zdarzeń $xyB = 001, 011, 101, 111$ mogą mieć miejsce. Poniżej przedstawione zostały wszystkie 16 możliwości tych zdarzeń:

$$1)\begin{pmatrix} y & x & B \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad 2)\begin{pmatrix} y & x & B \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad 3)\begin{pmatrix} y & x & B \\ 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad 4)\begin{pmatrix} y & x & B \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

$$5)\begin{pmatrix} y & x & B \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad 6)\begin{pmatrix} y & x & B \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad 7)\begin{pmatrix} y & x & B \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad 8)\begin{pmatrix} y & x & B \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix},$$

$$9)\begin{pmatrix} y\ x\ B \\ 0\ 0\ 0 \\ 0\ 1\ 0 \\ 0\ 1\ 1 \\ 1\ 0\ 0 \\ 1\ 0\ 1 \\ 1\ 1\ 1 \end{pmatrix},\quad 10)\begin{pmatrix} y\ x\ B \\ 0\ 0\ 0 \\ 0\ 0\ 1 \\ 0\ 1\ 0 \\ 1\ 0\ 1 \\ 1\ 1\ 0 \\ 1\ 1\ 1 \end{pmatrix},\quad 11)\begin{pmatrix} y\ x\ B \\ 0\ 0\ 0 \\ 0\ 0\ 1 \\ 0\ 1\ 1 \\ 1\ 0\ 0 \\ 1\ 1\ 0 \\ 1\ 1\ 1 \end{pmatrix},\quad 12)\begin{pmatrix} y\ x\ B \\ 0\ 0\ 1 \\ 0\ 1\ 0 \\ 0\ 1\ 1 \\ 1\ 0\ 0 \\ 1\ 0\ 1 \\ 1\ 1\ 0 \end{pmatrix},$$

$$13)\begin{pmatrix} y\ x\ B \\ 0\ 0\ 0 \\ 0\ 0\ 1 \\ 0\ 1\ 0 \\ 0\ 1\ 1 \\ 1\ 0\ 0 \\ 1\ 0\ 1 \end{pmatrix},\quad 14)\begin{pmatrix} y\ x\ B \\ 0\ 0\ 0 \\ 0\ 0\ 1 \\ 0\ 1\ 0 \\ 0\ 1\ 1 \\ 1\ 1\ 0 \\ 1\ 1\ 1 \end{pmatrix},\quad 15)\begin{pmatrix} y\ x\ B \\ 0\ 0\ 0 \\ 0\ 0\ 1 \\ 1\ 0\ 0 \\ 1\ 0\ 1 \\ 1\ 1\ 0 \\ 1\ 1\ 1 \end{pmatrix},\quad 16)\begin{pmatrix} y\ x\ B \\ 0\ 1\ 0 \\ 0\ 1\ 1 \\ 1\ 0\ 0 \\ 1\ 0\ 1 \\ 1\ 1\ 0 \\ 1\ 1\ 1 \end{pmatrix}.$$

Rozważmy jakie kanały z x do B implikują te zdarzenia. Zauważmy, że możemy wyróżnić następujące klasy:

- dla $y = 1$, B jest deterministyczną funkcją x (przypadki 1 i 2);
- dla $y = 0$, B jest deterministyczną funkcją x (przypadki 3 i 4);
- przynajmniej jedna wartość x (konkretna wartość może zależeć od y) jest dokładnie przekazywana do B (przypadki 5-12).

Pozostałe przypadki (13-16) nie mogą być zrealizowane, gdyż zdarzenia zredukowane do xy nie oddają wszystkich czterech możliwości $00, 01, 10, 11$, czego wymagamy jeśli prawdopodobieństwa $p(x, y)$ mają być niezerowe. Zauważmy dalej, że spośród 12 przypadków możemy wyszczególnić 3 przypadki reprezentatywne, z których pozostałe 9 otrzymać można poprzez zastosowanie negacji bitów dla zmiennych x, y, B:

(i) dla $y = 0$ mamy $B = x$ (przypadek 3 jako reprezentatywny dla przypadków 1-4);

(ii) dla $y = 0$, $x = 0$ implikuje $B = 0$, natomiast dla $y = 1$, $x = 0$ implikuje $B = 1$ (przypadek 6 jako reprezentatywny dla przypadków 5-8);

(iii) dla $y = 0$, $x = 0$ implikuje $B = 0$, natomiast dla $y = 1$, $x = 1$ implikuje $B = 1$ (przypadek 9 jako reprezentatywny dla przypadków 9-12).

(i). Z warunku działania KSD dla $y = 0$ mamy $B = a_0$. Tym samym, chcąc wiernie odtworzyć wartość x musimy dokonać kodowania $x \to a_0$. Bit z kodujemy natomiast

xyB	xBy	kanał $x \to B$	kodowanie dla $x=0$	kodowanie dla $x=1$	kanał $z \to B$
000					
010	000	$0 \to 0$	$x=0$	$x=1$	
~~100~~	110	$1 \to 1$	$\Downarrow$	$\Downarrow$	$z \to$ szum
110			$a_0=0$	$a_0=1$	
~~001~~	001				
011	011	$0 \to 0$	$x=0$	$x=1$	
101	101	$\diagdown\!\!\diagup$	$\Downarrow$	$\Downarrow$	$z \to z$
111	111	$1 \to 1$	$a_1=z$	$a_1=z$	

Tab. 6.1: Schemat kodowania dla przypadku 3. Dla bitu z otrzymujemy kanał wymazujący z prawdopodobieństwem wymazywania $p(y=0)=\frac{1}{2}$

do drugiego wejścia KSD $z \to a_1$. Przedstawiony schemat kodowania dla przypadku 3 zobrazowany został w Tabeli 6.1. Powyższy protokół kodowania sprowadza się do protokołu przedstawionego na Rys. 6.7. Otrzymany w ten sposób kanał Λ jest kanałem wymazującym z prawdopodobieństwem wymazywania równym $p(y=0)=\frac{1}{2}$

$$\mathscr{E}(z)=\frac{1}{2}|0\rangle\langle 0|_f \otimes |\mathbb{I}\rangle\langle\mathbb{I}| + \frac{1}{2}|1\rangle\langle 1|_f \otimes |z\rangle\langle z|. \tag{6.50}$$

(ii). W celu wiernego odtworzenia wartości x musimy dokonać następującego kodowania: dla $x=0$ kodujemy 0 do wejścia a_0 oraz 1 do wejścia a_1; dla $x=1$ do wejść a_0 oraz a_1 kodujemy bit z. Dla bitu z otrzymujemy kanał, będący kanałem tłumiącym amplitudę. Dla $y=0$ jest to kanał Λ_0: wartość 0 jest transmitowana bez zmian $0 \to 0$ z prawdopodobieństwem równym 1, natomiast wartość 1 jest tłumiona $1 \to 0,1$ z prawdopodobieństwem równym $p(x=0)=\frac{1}{2}$. Działanie kanału Λ_0 można przedstawić w formie

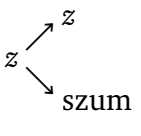
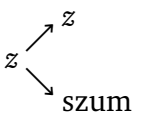
$$\Lambda_0(z)=\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|z\rangle\langle z|. \tag{6.51}$$

Podobnie, dla $y=1$ jest to kanał Λ_1: wartość 1 jest transmitowana bez zmian $1 \to 1$ z prawdopodobieństwem równym 1, natomiast wartość 0 jest tłumiona $0 \to 0,1$ z prawdopodobieństwem równym $p(x=0)=\frac{1}{2}$. Działanie kanału Λ_1 można przedstawić w formie

$$\Lambda_1(z)=\frac{1}{2}|1\rangle\langle 1| + \frac{1}{2}|z\rangle\langle z|. \tag{6.52}$$

Tym samym łącznie uzyskujemy kanał Λ, którego działanie jest dane wyrażeniem

$$\Lambda(z)=\frac{1}{2}|0\rangle\langle 0|_f \otimes \Lambda_0(z) + \frac{1}{2}|1\rangle\langle 1|_f \otimes \Lambda_1(z), \tag{6.53}$$

xyB	xBy	kanał $x \to B$	kodowanie dla $x=0$	kodowanie dla $x=1$	kanał $z \to B$
000 ~~010~~ 100 110	000 100 110	$0 \to 0$ $\nearrow$ $1 \to 1$	$x=0$ $\Downarrow$ $a_0=0$	$x=1$ $\Downarrow$ $a_0=z$	
~~001~~ 011 101 111	101 011 111	$0 \quad 0$ $\times$ $1 \to 1$	$x=0$ $\Downarrow$ $a_1=1$	$x=1$ $\Downarrow$ $a_1=z$	

Tab. 6.2: Schemat kodowania dla przypadku 6. Dla bitu z otrzymujemy kanał tłumiący amplitudę z prawdopodobieństwem tłumienia $p(x=0)=\frac{1}{2}$

gdzie flagi $|0\rangle\langle 0|_f$ oraz $|1\rangle\langle 1|_f$ odpowiadają wartości y wybranej przez Bolka.

Przedstawiony schemat kodowania dla przypadku 6 zobrazowany został w Tabeli 6.2.

(iii). W celu wiernego odtworzenia wartości x musimy dokonać następującego kodowania: dla $x=0$ kodujemy 0 do wejścia a_0 oraz bit z do wejścia a_1; dla $x=1$ kodujemy bit z do wejścia a_0 oraz 1 do wejścia a_1. Dla bitu z otrzymujemy kanał, będący kanałem tłumiącym amplitudę. Dla $y=0$ jest to kanał Λ_0: wartość 0 jest transmitowana bez zmian $0 \to 0$ z prawdopodobieństwem równym 1, natomiast wartość 1 jest tłumiona $1 \to 0,1$ z prawdopodobieństwem równym $p(xy=00)+p(xy=11)=\frac{1}{2}$. Podobnie, dla $y=1$ jest to kanał Λ_1: wartość 1 jest transmitowana bez zmian $1 \to 1$ z prawdopodobieństwem równym 1, natomiast wartość 0 jest tłumiona $0 \to 0,1$ z prawdopodobieństwem równym $p(xy=00)+p(xy=11)=\frac{1}{2}$. Podobnie jak w poprzednim przypadku, łącznie uzyskujemy kanał Λ, którego działanie przedstawić można jako

$$\Lambda(z)=\frac{1}{2}|0\rangle\langle 0|_f \otimes \Lambda_0(z)+\frac{1}{2}|1\rangle\langle 1|_f \otimes \Lambda_1(z). \tag{6.54}$$

Przedstawiony schemat kodowania dla przypadku 9 zobrazowany został w Tabeli 6.3.

Podsumowując powyższe przypadki widzimy, że dla bitu z kanał Λ jest: (i) kanałem wymazującym z prawdopodobieństwem wymazywania $\varepsilon=p(y=0)=\frac{1}{2}$, (ii) kanałem tłumiącym amplitudę z prawdopodobieństwem tłumienia $p(x=0)=\frac{1}{2}$, (iii) kanałem tłumiącym amplitudę z prawdopodobieństwem tłumienia $p(xy=00)+p(xy=11)=\frac{1}{2}$. Jak pokażemy poniżej, jeśli wartości x i y były generowane losowo niezależnie od siebie (każdy wynik 0 lub 1 z prawdopodobieństwem $\frac{1}{2}$), to odpowiednie kanały tłumiące amplitudę można otrzymać z kanału wymazującego. W tym celu zapiszmy działanie kanału wymazującego z prawdopodobieństwem $\frac{1}{2}$ w następującej

xyB	xBy	kanał $x \to B$	kodowanie dla $x=0$	kodowanie dla $x=1$	kanał $z \to B$
000 010 100 ~~110~~	000 100 110	$0 \to 0$ $\nearrow$ $1 \to 1$	$x=0$ $\Downarrow$ $a_0=0$	$x=1$ $\Downarrow$ $a_0=z$	z $\nearrow$ z $\searrow$ szum
~~001~~ 011 101 111	001 011 111	$0 \to 0$ $\searrow$ $1 \to 1$	$x=0$ $\Downarrow$ $a_1=z$	$x=1$ $\Downarrow$ $a_1=1$	z $\nearrow$ z $\searrow$ szum

Tab. 6.3: Schemat kodowania dla przypadku 9. Dla bitu z otrzymujemy kanał tłumiący amplitudę z prawdopodobieństwem tłumienia $p(xy=00)+p(xy=11)=\frac{1}{2}$

formie:

$$\mathcal{E}(z)=\frac{1}{2}|0\rangle\langle 0|_f\otimes|\mathbb{I}\rangle\langle\mathbb{I}|+\frac{1}{2}|1\rangle\langle 1|_f\otimes|z\rangle\langle z|, \tag{6.55}$$

gdzie stany $|0\rangle\langle 0|_f$ oraz $|1\rangle\langle 1|_f$ są odpowiednią informacją („flagą") o wyborze wejścia do KSD dokonanym przez Bolka. Niech teraz Bolek dokona następującego przetworzenia informacji:

- losowa zmiana flagi $|1\rangle\langle 1|_f$ na $|0\rangle\langle 0|_f$ lub $|1\rangle\langle 1|_f$ przy zachowaniu wyniku $|z\rangle\langle z|$:

$$|1\rangle\langle 1|_f\otimes|z\rangle\langle z|\to\frac{1}{2}(|0\rangle\langle 0|_f+|1\rangle\langle 1|_f)\otimes|z\rangle\langle z|; \tag{6.56}$$

- losowa zmiana flagi $|0\rangle\langle 0|_f$ na $|0\rangle\langle 0|_f$ lub $|1\rangle\langle 1|_f$ przy zmianie wyniku $|\mathbb{I}\rangle\langle\mathbb{I}|$ na $|0\rangle\langle 0|$ lub $|1\rangle\langle 1|$ odpowiednio do nowej flagi:

$$|0\rangle\langle 0|_f\otimes|\mathbb{I}\rangle\langle\mathbb{I}|\to\frac{1}{2}|0\rangle\langle 0|_f\otimes|0\rangle\langle 0|+\frac{1}{2}|1\rangle\langle 1|_f\otimes|1\rangle\langle 1|. \tag{6.57}$$

Po tak zastosowanym przetworzeniu informacji otrzymamy

$$\begin{aligned}
&\frac{1}{2}\left(\frac{1}{2}|0\rangle\langle 0|_f\otimes|0\rangle\langle 0|+\frac{1}{2}|1\rangle\langle 1|_f\otimes|1\rangle\langle 1|\right)+\frac{1}{2}\left(\frac{1}{2}|0\rangle\langle 0|_f+\frac{1}{2}|1\rangle\langle 1|_f\right)\otimes|z\rangle\langle z|\\
&=\frac{1}{2}|0\rangle\langle 0|_f\otimes\left(\frac{1}{2}|0\rangle\langle 0|+\frac{1}{2}|z\rangle\langle z|\right)+\frac{1}{2}|1\rangle\langle 1|_f\otimes\left(\frac{1}{2}|1\rangle\langle 1|+\frac{1}{2}|z\rangle\langle z|\right)\\
&=\frac{1}{2}|0\rangle\langle 0|_f\otimes\Lambda_0(z)+\frac{1}{2}|1\rangle\langle 1|_f\otimes\Lambda_1(z),
\end{aligned} \tag{6.58}$$

gdzie przez $\Lambda_0(z)$ i $\Lambda_1(z)$ oznaczyliśmy odpowiednie kanały tłumiące amplitudę. Oczywiście, przy odpowiednio dobranej zmianie szumu na poszczególne stany $|0\rangle\langle 0|$ lub

$|1\rangle\langle 1|$ można otrzymać dowolną kombinację kanałów $\Lambda_i(z)$. Widzimy zatem, że każdy z rozważanych kanałów tłumiących amplitudę można otrzymać z kanału wymazującego, o ile prawdopodobieństwo wymazywania wynosi $\varepsilon = \frac{1}{2}$.

Zwróćmy uwagę, że kodowanie wartości z w przypadkach (i)-(iii) mogłoby w ogólności wykorzystywać dowolną funkcję z, co jednak nie doprowadzi do zwiększenia pojemności kanału $\Lambda(z)$, co wynika z nierówności przetwarzania danych [94].

Rozpatrzone przypadki (1) oraz (2) dowodzą tezy twierdzenia.■

Rozdział 7

Podsumowanie

Niniejsza rozprawa została poświęcona zbadaniu wybranych własności korelacji występujących w mechanice kwantowej oraz ogólnych teoriach probabilistycznych.

Rozdział 2 stanowił przegląd najważniejszych wiadomości teoretycznych dotyczących matematycznego ujęcia zagadnień poruszanych w tej pracy.

W Rozdziale 3 rozważyliśmy metodę aktywacji łamania nierówności CHSH opartą na wykorzystaniu wielokrotnej wymiany splątania kwantowego w łańcuchu dwuqubitowych stanów kwantowych. Pokazaliśmy, że po dokonaniu dostatecznie dużej liczby wymian splątania przy odpowiednich wynikach pomiarów Bella na stanach niełamiących nierówności CHSH możliwe jest otrzymanie stanu kwantowego łamiącego tę nierówność. Minimalna liczba wymian splątania kwantowego potrzebna do aktywacji nielokalności zależy od początkowych stanów, przy czym aktywację można otrzymać nawet jeśli stan centralny ρ_0 posiadał dowolnie małą wagę stanu splątanego. Należy zwrócić uwagę, że pokazana tutaj aktywacja łamania nierówności CHSH różni się jakościowo od wyniku uzyskanego w [99] gdzie aktywację łamania nierówności CHSH pokazano w inny sposób w układzie dwóch użytkowników.

W Rozdziale 4 wprowadziliśmy dwie miary kontekstualności: wzajemną informację kontekstualności oraz względną entropię kontekstualności. Pomimo, że definicje obu miar bazują na odmiennych podejściach ilościowego ujęcia kontekstualności, ich wartości dla danych układów mają taką samą wartość, a tym samym są one równoważne. W dalszej kolejności skupiliśmy się na wyznaczeniu wartości względnej entropii kontekstualności dla następujących układów: kwadratu Peresa-Mermina, gwiazdy Mermina [6–8], układu $KCBS$ [9], oraz ogólnych układów łańcuchowych (w tym m.in. układ Popescu-Rohrlicha [2, 3]). Należy zwrócić uwagę, że wprowadzona miara kontekstualności pozwala na uniwersalny sposób porównywania kontekstualności różnych układów. Dzięki temu mogliśmy stwierdzić, że wartość miary kontekstualności dla układu $KCBS$ oraz układu kwantowego maksymalnie łamiącego nierówność CHSH jest o rząd wielkości mniejsza niż wartość miary dla układu Popescu-Rohrlicha czy kwadratu Peresa-Mermina oraz gwiazdy Mermina.

W Rozdziale 5 wprowadziliśmy relacje wykluczania informacji ograniczające sumę dwóch wzajemnych informacji dla odpowiednich par obserwabli mierzonych przez odległych użytkowników na współdzielonych stanach kwantowych. Pierwsza z nich ogranicza sumę dwóch wzajemnych informacji w przypadku, gdy jedna osoba mierzy zawsze tylko jedną obserwablę, natomiast druga mierzy jedną z dwóch obserwabli. Pokazaliśmy, że ta relacja jest silniejsza niż relacja wykluczania informacji Halla [4], która wynika bezpośrednio z zasady nieoznaczoności Maassena-Uffinka [5]. Podali-

śmy dowód tej relacji dla pewnych szczególnych przypadków, oraz dla pełności prezentacji przytoczyliśmy dowód w ogólnym przypadku przedstawiony w pracy autorstwa Colesa i Pianiego [20]. Druga relacja wykluczania informacji ogranicza sumę dwóch wzajemnych informacji w przypadku, gdy obie osoby mierzą jedną z dwóch obserwabli. Podaliśmy dowód relacji dla pewnej szczególnej klasy stanów kwantowych, jednak pozostaje pytanie, czy dana relacja jest słuszna dla dowolnych stanów.

Rozdział 6 poświęciliśmy porównywaniu różnych zasobów. Wprowadziliśmy *układ związany z kodem swobodnego dostępu* ($uKSD$), który wraz dodatkowym bitem informacji działa jak kod swobodnego dostępu. Następnie pokazaliśmy, że jeśli układ ten spełnia warunek niesygnalizowania to jest on równoważny układowi Popescu-Rohrlicha. W dalszej części sformułowaliśmy nierówność wiążącą różne zasoby, przy czym wykazaliśmy, że dysponując dowolnym $uKSD$, bitem komunikacji oraz bitem współdzielonej losowości możemy zasymulować układ Popescu-Rohrlicha otrzymując jednocześnie kanał wymazujący.

Bibliografia

[1] J.F. Clauser, M.A. Horne, A. Shimony, R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[2] P. Rastall, Found. Phys. **15**, 963 (1985).

[3] S. Popescu, D. Rohrlich, Found. Phys. **24**, 379 (1994).

[4] M.J.W. Hall, Phys. Rev. Lett. **74**, 3307 (1995).

[5] H. Maassen, J. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).

[6] A. Peres, Phys. Lett. A **151**, 107 (1990).

[7] D.N. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).

[8] D.N. Mermin, Rev. Mod. Phys. **65**, 803815 (1993).

[9] A.A. Klyachko, M.A. Can, S. Binicioğlu, A.S. Shumovsky, Phys. Rev. Lett. **101**, 020403 (2008).

[10] M. Żukowski, A. Zeilinger, M.A. Horne, A.K. Ekert, Phys. Rev. Lett. **71**, 4287 (1993).

[11] M. Żukowski, A. Zeilinger, H. Weinfurter, Ann. N.Y. Acad. Sci. **755**, 91 (1995).

[12] A. Wójcik, J. Modławska, A. Grudka, M. Czechlewski, Phys. Lett. A **374**, 4831 (2010).

[13] W. Kłobus, W. Laskowski, M. Markiewicz, A. Grudka, Phys. Rev. A **86**, 020302(R) (2012).

[14] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, A. Wójcik, Phys. Rev. Lett. **112**, 120401 (2014).

[15] R.F. Werner, Phys. Rev. A **40**, 4277 (1989).

[16] V. Vedral, M.B. Plenio, Phys. Rev. A **57**, 1619 (1998).

[17] E.M. Rains, Phys. Rev. A **60**, 179 (1999).

[18] W. van Dam, R.D. Gill, P.D. Grünwald, IEEE Trans. Inf. Theory **51**, 2812 (2005).

[19] A. Grudka, M. Horodecki, P. Horodecki, R. Horodecki, W. Kłobus, Ł. Pankowski, Phys. Rev. A **88**, 032106 (2013).

[20] P.J. Coles, M. Piani, arXiv:quant-ph/1307.4265.

[21] A. Grudka, K. Horodecki, M. Horodecki, W. Kłobus, M. Pawłowski, arXiv:1307.7904.

[22] J.S. Bell, Physics **1**, 195 (1964).

[23] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, S. Wehner, arXiv:quant-ph/1303.2849.

[24] R. Horodecki, P. Horodecki, M. Horodecki, Phys. Lett. A **200**, 340 (1995).

[25] T. Vértesi, Phys. Rev. A **78**, 032112 (2008).

[26] L.J. Landau, Phys. Lett. A **123**, 115 (1987).

[27] B.S. Tsirelson, Lett. Math. Phys. **4**, 93 (1980).

[28] M. Froissard, Nuovo Cimento B **64**, 241 (1981).

[29] D. Collins, N. Gisin, J. Phys. A **37**, 1775 (2004).

[30] J.-D. Bancal, N. Gisin, S. Pironio, J. Phys. A **43**, 385303 (2010).

[31] N. Brunner, N. Gisin, Phys. Lett. A **372**, 3162 (2008).

[32] K.F. Pál, T. Vértesi, Phys. Rev. A **79**, 022120 (2009).

[33] D. Avis, H. Imai, T. Ito, Y. Sasaki, arXiv:quant-ph/0404014.

[34] D. Collins, N. Gisin, N. Linden, S. Massar, S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).

[35] J. Barrett, A. Kent, S. Pironio, Phys. Rev. Lett. **97**, 170409 (2007).

[36] R.F. Werner, M.M. Wolf, Phys. Rev. A **64**, 032112 (2002).

[37] M. Żukowski, C. Brukner, **88**, 210401 (2002).

[38] M. Ardehali, Phys. Rev. A **46**, 5375 (1992).

[39] A. Belinskii, D. Klyshko, Phys. Usp. **36**, 653 (1993).

[40] D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).

[41] C. Śliwa, Phys. Lett. A **317**, 165 (2003).

[42] W. Laskowski, T. Paterek, C. Brukner, M. Żukowski, Phys. Rev. Lett. **93**, 200401 (2004).

[43] E. Cavalcanti, C. Foster, M. Reid, P. Drummond, Phys. Rev. Lett. **99**, 210405 (2007).

[44] Q. He, E. Cavalcanti, M. Reid, P. Drummond, Phys. Rev. Lett. **103**, 180402 (2009).

[45] A. Salles, D. Cavalcanti, A. Acin, D. Perez-Garcia, M. Wolf, Quant. Inf. and Comp. **10**, 0703 (2010).

[46] B. Grandjean, Y.-C. Liang, J.-D. Bancal, N. Brunner, N. Gisin, Phys. Rev. A **85**, 052113 (2012).

[47] J.-L. Chen, C. Wu, L. Kwek, C. Oh, Phys. Rev. A **78**, 032107 (2008).

[48] S.L. Braunstein, C.M. Caves, Ann. Phys. **202**, 22 (1990).

[49] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, D. Roberts, Phys. Rev. A **71**, 022101 (2005).

[50] J. Barrett, arXiv:quant-ph/0508211.

[51] Ll. Masanes, A. Acin, N. Gisin, Phys. Rev. A **73**, 012112 (2006).

[52] W. van Dam, arXiv:quant-ph/0501159.

[53] S. Wolf, J. Wullschleger, arXiv:quant-ph/0502030.

[54] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, M. Żukowski, Nature **461**, 1101 (2009).

[55] M. Pawłowski, M. Żukowski, Phys. Rev. A **81**, 042326 (2010).

[56] H. Robertson, Phys. Rev. **34**, 163 (1929).

[57] S. Wehner, A. Winter, New J. Phys. **12**, 025009 (2010).

[58] W. Heisenberg, Zeitschrift für Physik **43**, 172 (1927).

[59] C.E. Shannon, Bell System Technical Journal **27**, 379 (1948).

[60] I.I. Hirschmann, Am. J. Math. **79**, 152 (1957).

[61] W. Beckner, Ann. Math. **102**, 1 (159).

[62] I. Białynicki-Birula, J. Mycielski, Phys. Lett. A **108**, 384 (1985).

[63] I. Białynicki-Birula, Ł. Rudnicki, Statistical Complexity, Ed. K.D. Sen, Springer (2011)

[64] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).

[65] K. Kraus, Phys. Rev. D **35**, 3070 (1987).

[66] M. Berta, M. Christandl, R. Colbeck, J.M. Renes, R. Renner, Nature Phys. **6**, 659 (2010).

[67] R. Prevedel, D.R. Hamel, R.Colbeck, K.Fisher, K.J. Resch, Nature Phys. **7**, 757 (2011).

[68] T. Pramanik, P. Chowdhury, A.S. Majumdar, Phys. Rev. Lett. **110**, 020402 (2013).

[69] M. Berta, M. Christandl, F. Furrer, V.B. Scholz, M. Tomamichel, arXiv:quant-ph/1308.4527.

[70] M.-L. Hu, H. Fan, Phys. Rev. A **87**, 022314 (2013).

[71] M.J.W. Hall, Phys. Rev. A **55**, 100 (1997).

[72] S. Kochen, E.P. Specker, J. Math. Mech. **17**, 59 (1967).

[73] J.S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[74] M. Pavičić, J.-P. Merlet, B.D. McKay, N.D. Megill, J. Phys. A **38**, 1577 (2005).

[75] A. Cabello, Int. J. Quant. Info. **4**, 55 (2006).

[76] A. Cabello, J.M. Estebaranz, G. Garcia-Alcaine, Phys. Lett. A **212**, 183 (1996).

[77] A. Cabello, J.M. Estebaranz, G. Garcia-Alcaine, Phys. Lett. A **339**, 425 (2005).

[78] M. Kernaghan, A. Peres, Phys. Lett. A **198**, 1 (1995).

[79] C. Simon, C. Brukner, A. Zeilinger, Phys. Rev. Lett. **86**, 4427 (2001).

[80] J.Å. Larsson, Europhys. Lett. **58**, 799 (2002).

[81] A. Cabello, Phys. Rev. Lett. **101**, 210401 (2008).

[82] P. Badziąg, I. Bengtsson, A. Cabello, I. Pitowsky, Phys. Rev. Lett. **103**, 050401 (2009).

[83] R. Lapkiewicz, P. Li, C. Schaeff, N.K. Langford, S. Ramelow, M. Wieśniak, A. Zeilinger, Nature **474**, 490 (2011).

[84] P. Kurzyński, R. Ramanathan, D. Kaszlikowski, Phys. Rev. Lett. **109**, 020404 (2012).

[85] S.L. Braunstein, C.M. Caves, Phys. Rev. Lett. **61**, 662 (1988).

[86] T. Vidick, S. Wehner, Phys. Rev. Lett. **107**, 030402 (2011).

[87] T. Fritz, R. Chaves, Phys. Rev. A **85**, 032113 (2012).

[88] N.J. Cerf, C. Adami, Phys. Rev. Lett. **79**, 5194 (1997).

[89] A. Grudka, P. Kurzyński, Phys. Rev. Lett. **100**, 160401 (2008).

[90] M. Horodecki, P. Horodecki, R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[91] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[92] M. Araújo, M.T. Quintino, C. Budroni, M.T. Cunha, A. Cabello, Phys. Rev. A **88**, 022118 (2013).

[93] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[94] T.M. Cover, J.A. Thomas, *Elements of Information Theory* (John Wiley & Sons, Inc., New York 1991), pp. 1–509.

[95] M.A. Nielsen, I.L. Chuang, *Quantum computation and Quantum Information* (Cambridge University Press, Cambridge 2000), pp. 1–695.

[96] F. Topsøe, Entropy **3**, 162 (2001).

[97] A. Rényi, Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability, 547 (1960).

[98] M. Pawłowski, J. Kofler, T. Paterek, M. Seevinck, C. Brukner, New J. Phys. **12**, 083051 (2010).

[99] M. Navascués, T. Vértesi, Phys. Rev. Lett. **106**, 060403 (2011).