

量子系の統計的推測と量子相関¹

— 推定における量子効果の検証に向けて —

科学技術振興事業団 ERATO 今井量子計算機構プロジェクト 林 正人²

(2003 年 5 月 12 日受理)

量子力学では量子状態の量子相関が非局所性を表すものとして扱われているが、量子系の統計的推測においては量子状態の量子相関だけでなく、量子測定における量子相関も重要な役割を果たすことを解説する。さらに、現在の技術で十分に可能な測定で量子相関を用いることによる効果が明確なものを量子ガウス状態や量子 2 準位系を例にとって提案する。

1 量子系の統計的推測の意義

近年、情報機器から医療機器に至るまで様々な分野で微細化が進みつつある。そのため、従来は古典的な記述で十分であったにも関わらず、その微細化の結果、量子力学的記述が必要となることは少なくない。例えば、現在のコンピュータ機器は古典的な bit を基本としていていると考えられる。個々のデバイスは量子力学に従うわけであるが、1つの古典的 bit を表す粒子数が従来は十分多かったので、巨視的にはこれを古典的に扱うことが許された。しかしながら、微細化の結果、1つの古典的 bit を表す粒子（電子）の数がこのような扱いができる程度には多くはないという状況になりつつある。事実、現状では1ビットに対応する電子の数が数百単位となっている状況もあると言われており、これが数十単位になる微細化が進むと何らかの量子的な情報処理を必然的に考えざるを得なくなる。その他、ナノスケールの微細な系を取り扱う際にも、同様のことが問題になると思われる。

この問題の1つの定式化が量子系での統計的推測である。このような研究は、1967年に光通信を量子的に扱う研究に関連して Helstrom[1] により始められ、70年代を通じて精力的に研究された[2, 3, 4]。80年代に入ると研究者の数が激減し停滞したが、90年代後半から、70年代には無かった視点から研究されるようになり、急速に進展した[5, 6]。この問題設定では、測定者に有用な情報が、量子的な信号を決定するパラメータとして与えられていると考える。そして、信号に対する測定を通じてのみ、そのパラメータに関する情報が得られると考える。一般に量子系では測定を行なうと系は乱され、系がもともと持っていた情報が失われる。そのため、それぞれの状況の下で、最も多くの情報が獲得できる測定を選ぶ必要がある。一般には量子力学的な非可換性のために、複数のパラメータに対して、同時に最適な測定は存在しない。場合によっては、推定したい

¹本稿は、編集部の方から特にお願ひして執筆していただいた記事である。

²E-mail: masahito@qci.jst.go.jp

パラメータが複数あるときには、片方のパラメータに対して最適な測定を行うと、他のパラメータに対する情報が全く得られないことが起きてしまう。故に、この問題設定では、測定の選択が本質的な部分になる。ただ、困ったことに、実際に現在の技術でそのような量子力学的な信号に対して可能な測定の範囲と、量子力学の原理から許される測定の範囲とのギャップは大きい。特に、現在の技術で可能な測定は問題となる信号の性質にもかなり依存するなどの理由もあり、一般的な定式化は困難である。このような事情から、しばしば、理論的な立場として、量子力学的に許される全ての測定の範囲で、推測のために最適な測定を選ぶという問題が採用される。

このような問題設定は、微細化の結果として得られた量子系だけでなく、量子計算や量子暗号などのように（微細化の結果ではなく）意図的に構成された量子系を扱う際にも有効である。量子計算や量子暗号の部品を作ったとしてもそれらが設計図通りに動作するとは限らない。それゆえ、それらが設計図通りの性能を示すか否か推測する必要があるわけであるが、もちろん、この推測は量子系に対する測定を通じてなされるものであり、量子力学の確率解釈から必然的に統計的なものになる。その統計的推測をなるべく効率良く行うための手段として、量子系の統計的推測は有効な手段である。

応用例の一つとして人体内部の情報を得る技術である Positron Emission Tomography(PET) や Single Photon Emission Computed Tomography(SPECT) の性能向上が考えられる。ここで、これらを例として量子系の統計的推測がどのような形で適用されるか簡単に説明しよう。PET や SPECT ではともに、放射性同位体を体内に投与して、外部からその分布状態を測定することで、人体内部の情報を推測する。PET では、陽電子を放出する放射性同位体を投与することで、体内での崩壊が起き陽電子が放出される。その陽電子はすぐに電子と衝突し、1 対の γ 線を放出する。この 1 対の（運動量和が 0 となる 2 つの） γ 線を測定することで体内の情報を得る。一方、SPECT では γ 線を放出する放射性同位体を投与し、同様に γ 線を測定することで体内の情報を得る。しかし、放射性同位体をある一定量以上、体内に投与することは危険である。そのため、より少ない投与で精度の高い情報を得ることが課題となる。もちろん、投与する放射性同位体を減らせば、体内での崩壊のときに放出される γ 線の量は少なくなり、より緻密な情報処理が要求される。 γ 線は量子力学的対象で、放出された γ 線から、その発生源を特定するには、 γ 線を検出した場所と、衝突時の運動量の方向を同時に推定することが要求される。しかし、量子力学では良く知られているように、位置と運動量を同時に推定することは困難であり、また、これらは完全には確定していない。そのような状況の下では、発生源を未知パラメータとする、量子状態の推定問題として問題を定式化するのが適切であると思われる。ここでのポイントはデータ処理過程だけでなく、測定過程までも最適化しようとする点にある。

その他、素粒子実験や重力波測定など、量子系から精密に情報を読み出す必要のある分野への応用も期待できる。このように、量子系の推測は量子系を用いた様々な処理を実際に実行する際の基礎となる部分であり、今後、人類がさらなる微細化の結果、量子系の操作に取り組むに当たり、その必要性は益々高まると思われる。さらに信頼性のある量子状態発生のためにも、推定論的手法が重要になることは間違い無い。既に、実際に生成されたエンタングルメント状態の検証に量子状態推定は用いられている [7, 8]。今後この分野の発展には、物理学、統計学の双方の分野から

のアプローチが期待される。なお、その他の歴史的なことや参考文献については参考文献 [5] や [9] を参照されたい。

2 量子相関を用いた推定

量子計算や量子暗号などをはじめとして様々な量子情報処理が提案されている。これらの情報処理ではエンタングルした状態（量子力学的に相関を持つ状態）が重要な役割を果たすことが少なくない。エンタングルした状態は量子系にのみ存在するもので、量子力学的な効果を有効に発揮するにはこのエンタングルした状態を利用することが必要であることが多い。一方、量子情報処理の1つである量子系での統計的推測では複数の状態が独立に準備されている状況を想定することが多く、エンタングルした状態を想定することは少ない。しかし、状態がエンタングルしていない場合でも、すなわち、状態が量子力学的に相関を持っていない場合でも、測定として量子力学的な相関（量子相関）を持つ測定を行った方が、量子相関を持たない測定を行う場合よりも推定精度が良くなる場合がある。量子相関を持つ測定は量子系特有のものであるので、それを持たない測定との推定精度の差が量子効果を表していると考えることができる。本稿では、現実の物理系でしばしば表れる量子ガウス状態の推定問題に注目し、実現可能な量子相関を持つ測定で明確に量子効果が表れるものを提案する。もちろん、量子相関を持つ測定が重要な役割を果たすのは量子系の統計的推測に限ったことではない。量子系での通信路符号化でも量子相関を用いることで通信路容量が向上することが理論的には確かめられている。それゆえ、量子相関を持つ量子測定の効果に関する研究は量子情報理論での1つの重要なテーマである。本稿では統計的推測における推定論でこのような量子効果を実験的に検証する枠組みを提案する。

最初に3節で量子系の数学的記述と推定の定式化について述べ、次に4節で量子 i.i.d. 条件と量子相関を用いた推定と、用いない推定の定義を与える。そして、5節では量子系で重要な状態である量子ガウス状態の推定について述べる。この量子ガウス状態の推定はデータ数が少ない場合での議論が可能であるが、一般のモデルではデータが少ない場合には適用できない。6節ではそのような一般のモデルでも展開できる、データが多い場合の議論（漸近論）を扱う。さらに7節では、もう一つの量子効果の検証となる実験として、量子2準位系の固有値推定を扱った。なお、確率分布のパラメータ推定は統計学で古くから議論されてきたテーマであり、特に漸近論については蓄積が豊富である。これらについて、より詳しく述べるために、付録Bを準備した。はじめにB.1節では未知のパラメータが1つの場合の漸近論について述べ、B.2節では未知パラメータが複数ある場合について述べた。そして、B.3節では2次のオーダーの漸近論について簡単に扱った。

3 量子系の数学的記述と推定の定式化

ここでは量子系を記述するために基礎的な事項を説明する。以下に述べられることは極めて基本的であるが、標準的な量子力学の教科書には必ずしも十分に記述されていないので注意を要する。量子力学では注目する物理系（量子系）は Hilbert 空間（系の表現空間、以下 $\mathcal{H}$ と記す。）で

記述され、状態は Hilbert 空間 $\mathcal{H}$ 上の密度作用素で記述される。 $\mathcal{H}$ 上の密度作用素とは $\mathcal{H}$ から $\mathcal{H}$ への self-adjoint (Hermite) な線型作用素で positive semi-definite かつその Trace が 1 となるものである。このような密度作用素の全体を $\mathcal{S}(\mathcal{H})$ で表す。

測定は $\mathcal{H}$ 上の恒等作用素 I の positive semi-definite 作用素による分解 $\{M_i\}$ で記述される。この分解 $\{M_i\}$ は POVM (正作用素値測定 Positive Operator-Valued Measure) と呼ばれる。系が密度作用素 $\rho \in \mathcal{S}(\mathcal{H})$ に対応する状態にあるとき POVM $\{M_i\}_{i=1}^k$ に対応する測定を行うと、その測定値 $i = 1, \dots, k$ は確率分布

$$P_\rho^M(i) := \text{Tr } \rho M_i \quad (1)$$

に従う。これが確率の条件 ($P_\rho^M(i) \geq 0, \sum_i P_\rho^M(i) = 1$) を満たすために、 ρ 及び M が positive semi-definite であることなどが効いていることに注意されたい。この定式化では状態と測定の再現性が暗に仮定されている。そうでなければ、式 (1) を実験により検証することは不可能である。

よく知られているように、物理量が自己共役作用素 X で記述され、そのスペクトル分解が $X = \sum_i x_i E_i$ (x_i は固有値で E_i はその固有空間への射影) で与えられるとき、分解 $\{E_i\}$ を物理量 X の測定とよぶことがある。しかしながら、より広い意味で

$$X = \sum_i x_i M_i \quad (2)$$

を満たす POVM $\{M_i\}$ を X の測定とよぶこともある。実際多くの粒子を一度に測る際にはその期待値 $\text{Tr } X\rho = \sum_i x_i \text{Tr } M_i \rho$ が直接得られるため、そのような場合には POVM の形ではなく対応する自己共役作用素 X の形が問題になる。しかしながら、粒子を1つずつ測定する場合には直接問題になるのは測定値 i が得られる確率 $P_\rho^M(i)$ であり、どの自己共役作用素に対応するかを議論するだけでは不十分である。推定の平均2乗誤差を評価するには (2) の意味で対応する作用素ではなく測定値に従う確率分布 P_ρ^M に注目する必要性があり、この事実は6節で具体的に例を用いて示される。今後混乱のおそれがあるので、本稿では物理量（または自己共役作用素） X の測定という表現は用いない。

また、多くの場合、物理量 X を測定しようとして、そのスペクトル分解 $\{E_i\}$ に対応する測定を行なったつもりでも、実際には測定過程においてノイズがありスペクトル分解 $\{E_i\}$ から若干ずれた測定を行っていることが多い。このようなずれを表現するには測定を POVM で表現することが不可欠であり、それゆえ、測定を POVM で表現することは応用上、極めて有効である。

2つの量子系 $\mathcal{H}_1, \mathcal{H}_2$ の合成系はテンソル積空間 $\mathcal{H}_1 \otimes \mathcal{H}_2$ で記述される。特に各系の状態の密度作用素がそれぞれ ρ_1, ρ_2 で与えられており、相互作用が無いときには、合成系上の状態の密度作用素は $\rho_1 \otimes \rho_2$ で与えられる。量子力学系におけるパラメータ推定では、状態に対応する未知の密度作用素が密度作用素族 $\mathcal{S} = \{\rho_\theta \in \mathcal{S}(\mathcal{H}) | \theta \in \Theta \subset \mathbb{R}\}$ に含まれていることが既知の上で、測定値から未知のパラメータ θ を推定することになる。これを先の PET や SPECT の例に合わせて説明すると、人体内部の情報（崩壊が起きた位置など）がパラメータ $\theta \in \Theta$ に対応し、崩壊によって放出される γ 線が ρ_θ に対応する。

4 量子 i.i.d. 条件と 2 種の定式化

上述の方法を用いて状態の密度作用素を推測するためには、その状態にある系を複数準備する必要がある。そのため n 個のシステムにわたって同一の状態 ρ_θ が準備されることになる。このとき合成系 $\mathcal{H}^{\otimes n} := \underbrace{\mathcal{H} \otimes \cdots \otimes \mathcal{H}}_n$ 上の状態は密度作用素 $\rho^{\otimes n} := \underbrace{\rho \otimes \cdots \otimes \rho}_n$ で記述される (量子 n -i.i.d.³ 条件)。

ここで導入した量子 n -i.i.d. 条件を仮定すると、推定量は量子系 $\mathcal{H}^{\otimes n}$ に対する測定を表す (測定値集合が Ω_n である) POVM M^n と Ω_n からパラメータ空間への写像 $\hat{\theta}_n$ の組 $(M^n, \hat{\theta}_n)$ で記述される。多くの古典系と同様、平均 2 乗誤差 (MSE) を誤差の基準にとることとする。すなわち、パラメータ空間が 1 次元のときは単純にその平均 2 乗誤差

$$v_\theta(M^n, \hat{\theta}_n) := \int_{\Omega_n} (\hat{\theta}_n(\omega) - \theta)^2 \text{Tr} \rho_\theta^{\otimes n} M^n(d\omega)$$

を最小化することを考える。ただ、推定量に何も条件を課さないのであれば、例えば、推定値 $\hat{\theta}$ を常に 0 にとるという推定量を考えると、真のパラメータが 0 のとき上記の値は 0 になるが、他の値のときには良い性能を示さない。それゆえ、推定量に何らかの条件を課しないと、意味のある問題設定にならない。簡単な条件として推定量の期待値が真のパラメータに一致するという不偏性条件、すなわち

$$\int_{\Omega_n} \hat{\theta}_n \text{Tr} \rho_\theta^{\otimes n} M^n(d\omega) = \theta, \quad \forall \theta \in \Theta$$

を課することが多い。パラメータ空間 Θ が多次元 d の場合には i 番目のパラメータ θ^i に対してそれぞれ、平均 2 乗誤差 (MSE)

$$\int_{\Omega_n} (\hat{\theta}_n^i(\omega) - \theta^i)^2 \text{Tr} \rho_\theta^{\otimes n} M^n(d\omega) \quad (3)$$

を考えることになり、不偏性条件は

$$\int_{\Omega_n} \hat{\theta}_n^i(\omega) \text{Tr} \rho_\theta^{\otimes n} M^n(d\omega) = \theta^i, \quad \forall \theta \in \Theta, \quad 1 \leq i \leq d$$

で与えられる。量子推定の場合には、一般に各パラメータに対して最適な測定が異なり、それらが、非可換性のため同時に実現できないという深刻な問題がしばしば起きる。

この論文では、推定方式として、 n 個の系に対して個別的な測定を行う場合と、量子相関 (量子力学的な相互作用) を用いた複数の系にまたがる測定を行った場合を比較し、どの程度の推定精度の差が生じるか考えることにする。これにより、推定という問題の下での量子効果の影響を正確に見積もることができる。例えば、図 1 のように、各系を個別に測定して、古典的データ処理を行って、推定するという方式では、量子相関は用いる必要はない。もちろん、この場合、各系に個別に行う測定は、それまでに行った測定の測定値に応じて適応的に選択できることとする。

³i.i.d. は identically independent distribution の略

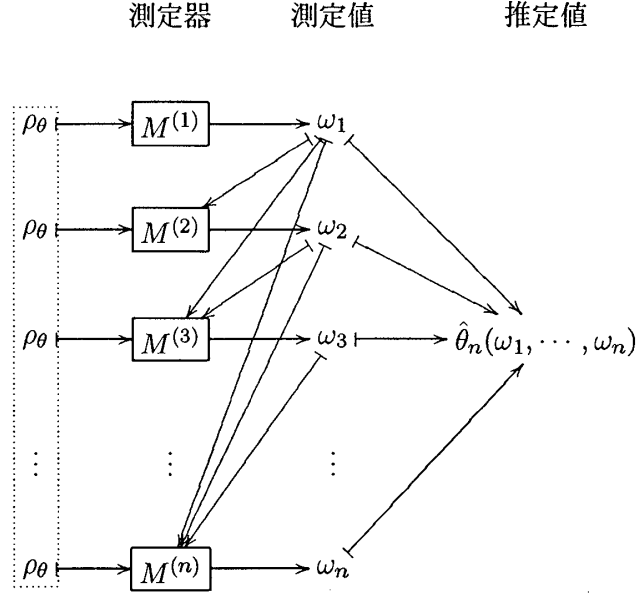


図 1: 量子相関を用いない推定

このような測定は以下の形の POVM で記述される.

- $M^n = \{M_{\omega_1, \dots, \omega_{n-1}, \omega_n}^{(1)} \otimes \dots \otimes M_{\omega_1, \dots, \omega_{n-1}, \omega_n}^{(n)}\}_{(\omega_1, \dots, \omega_n) \in \Omega^n},$
- $M_{\omega_1, \dots, \omega_{j-1}, \omega_j}^{(j)} = (M_{\omega_1, \dots, \omega_{j-1}, \omega_j}^{(j)})^\dagger \geq 0,$
- $\int_{\Omega} M_{\omega_1, \dots, \omega_{j-1}, \omega_j}^{(j)} d\omega_j = I \quad (j = 1, \dots, n).$

上記の条件を満たす POVM に限った場合の (3) の最小値を C_θ^{NQ} と書くことにする.

逆に, 図 1 で表した測定に含まれない測定を量子相関を用いた測定とよぶことにする. このような測定は図 2 のように表すことができる.

図 2 で表されるタイプの POVM での (3) の最小値を C_θ^Q とする. C_θ^Q と C_θ^{NQ} を比較することによって, 各系間の相関を用いることがどれだけ推定精度を向上させるか議論することができる. 本稿では, 後に定義を与える量子ガウス状態族に対して, 不偏性条件の下で上記の比較を行う. その次に, 不偏性条件とは異なる一般的な設定である, 1 次漸近理論を解説し, その枠内で量子ガウス状態族と他のモデルに対して同じ比較を行う.

5 量子ガウス状態の推定

この節では量子系での重要な状態である量子ガウス状態の推定について議論する. 5.1 節ではこの系で重要な測定方法について述べ, 5.2 節では量子相関を用いない場合の限界及びそのときの最適な推定方法を与える. そして 5.3 節では, 量子相関を用いた場合についても考え, そのときの限界及び, 限界を達成する測定を与える. 5.4 節では現実の不完全な測定で, 量子相関を用いることにより, 5.2 節で与えた量子相関の無いときの限界を超えるような推定方法を提案する.

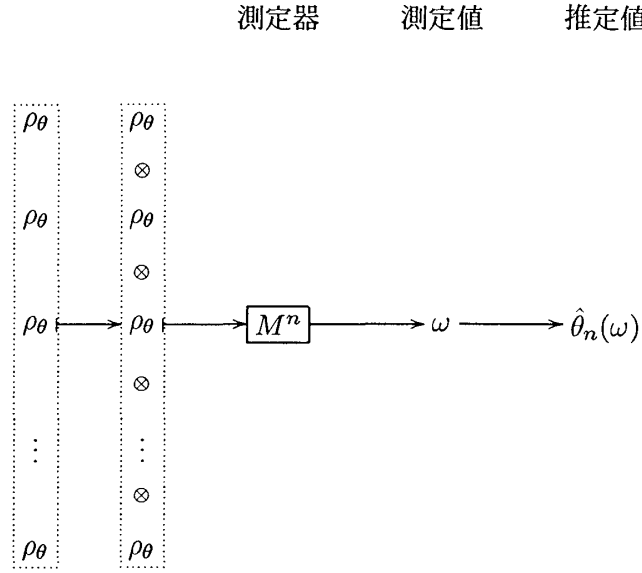


図 2: 量子相関を用いた推定

5.1 光子と量子ガウス状態

振動数が ν の光子からなる物理系は $|0\rangle, |1\rangle, \dots, |k\rangle, \dots$ を正規直交基底とする Hilbert 空間 $L^2(\mathbb{R})$ で記述される. ($|k\rangle$ に k 次の Hermite 関数に対応させると $\mathbb{R}$ 上の 2 乗可積分空間 $L^2(\mathbb{R})$ になる.) このとき状態 $|k\rangle\langle k|$ を光子の個数が k の状態であると見なすことができる. このように粒子の個数 (0 を含む正の整数) に対応するベクトルが表現空間の正規直交基底になる粒子のことを Bose 粒子 (boson) とよぶ. この系では, coherent ベクトル $|\alpha\rangle_a := \sum_{k=0}^{\infty} e^{-\frac{|\alpha|^2}{2}} \frac{\alpha^k}{\sqrt{k!}} |k\rangle$ ($\alpha \in \mathbb{C}$) が重要な役割を果たし, それに対応する密度作用素 $\rho_{\alpha,0} := |\alpha\rangle_a {}_a\langle\alpha|$ は coherent 状態とよばれる. この系において重要な POVM として

$$\boxed{\mathbf{N}}: k \mapsto |k\rangle\langle k| \quad (4)$$

$$\boxed{\mathbf{H}}: \alpha \mapsto |\alpha\rangle_a {}_a\langle\alpha| \quad (5)$$

が知られており, それぞれ個数測定, ヘテロダイン測定とよばれる. 前者では n は離散値である整数に値を取るのに対し, 後者では連続値である複素数に値を取るので, $\sum$ が積分に変わることに注意を要する. coherent 状態は比較的安定な状態で量子光学で重要な役割を果たす. 以下, 未知状態が coherent 状態をガウス積分で重ね合わせた量子ガウス状態 $\rho_{\zeta,N} := \frac{1}{\pi N} \int e^{-\frac{|\alpha-\zeta|^2}{N}} |\alpha\rangle_a {}_a\langle\alpha| d\alpha$ である場合の推定理論の概略を述べる. 例えば, 熱浴中で緩和する振動数 ν の光子からなる量子系の従う Master 方程式

$$\frac{d\rho}{dt} = -i[\nu a^\dagger a, \rho] - \frac{c\bar{n}}{2}(aa^\dagger\rho - 2a^\dagger\rho a + \rho aa^\dagger) - \frac{c(\bar{n}+1)}{2}(aa^\dagger\rho - 2a\rho a^\dagger + \rho a^\dagger a)$$

の下では coherent 状態 $|\zeta_0\rangle_a {}_a\langle\zeta_0|$ は

$$|\zeta_0\rangle_a {}_a\langle\zeta_0| \mapsto \rho_{\zeta_0 e^{-\frac{ct}{2}} - i\nu t, \bar{n}(1-e^{-ct})}$$

と時間発展する [10]. (ここで $\bar{n}$ 平均光子数で c は熱浴との結合係数である.) それゆえ, 量子ガウス状態の推定を $\bar{n}, \zeta_0$ が既知である場合では, t, ν, c を推定するために用いることが可能である. その他, $\bar{n}, c, \nu$ や $\bar{n}, \nu, t$ が既知である場合に, ζ_0, t や ζ_0, c を推定するためにも利用することができる. それゆえ, 量子ガウス状態からなる状態族を仮定することは十分な適用範囲を持つ. 以下では, 未知の状態が量子ガウス状態であると仮定した場合の推定について考える.

5.2 量子相関なしの推定

はじめに, ζ が 0 であることが既知で, N のみを推定する場合, すなわち状態族 $\left\{ \rho_{0,N} = \sum_{k=0}^{\infty} \frac{1}{N+1} \left(\frac{N}{N+1} \right)^k |k\rangle\langle k| \mid N \geq 0 \right\}$ の推定を考える. この場合の不偏性条件の下での最適な測定は, 個数測定 (4) を行い, そのデータ $k_1, \dots, k_n$ について平均値 $\hat{k}$ を推定量とするものに限られ, その平均 2 乗誤差は

$$\frac{1}{n} \sum_{k=0}^{\infty} (k - N)^2 \text{Tr} \rho_{0,N} |k\rangle\langle k| = \frac{N(N+1)}{n} \quad (6)$$

で与えられ, 量子相関を用いることのメリットが無いことが示せる. (この場合は状態族に含まれる全ての状態が互いに可換であるので, 古典的な統計的推測の理論に帰着できる. 特にこの場合は, 幾何分布のパラメータ推定に帰着される.)

次に ζ のみを推定する場合を考えることにする. このときには推定量に不偏性を課すと, 図 2 で示した個別的な測定の範囲で最適な推定方法は, 各系に (5) で与えられるヘテロダイン測定を行い, 得られたデータである $\alpha_1, \dots, \alpha_n$ についての平均値を ζ の推定値 $\hat{\zeta}$ とするものに限られる [2]. その平均 2 乗誤差は

$$\frac{1}{n} \int_{\mathbb{C}} |\zeta - \alpha|^2 \text{Tr} \rho_{\zeta,N} |\alpha\rangle_a \langle \alpha| d\alpha = \frac{1}{n} \int_{\mathbb{C}} |\zeta|^2 {}_a \langle \alpha | \rho_{0,N} | \alpha \rangle_a d\alpha = \frac{N+1}{n} \quad (7)$$

で与えられる. 不偏性条件の下でこれを最小化する最適測定は N には依存しない. この場合も同様に, 量子相関を用いることのメリットが無いことも示せる [5].

では ζ のみではなく N まで推定するのであればどうかであろうか. この場合は N の MSE と ζ の MSE の両方を同時に最小化する必要が生じるわけである. 多くの場合は両者を重み付けして和を取ったものを最小化する. しかし, ここでは問題を簡単化するために, ζ の MSE については最適な値を達成するという条件を課し, それを満たす不偏推定量の中で N の MSE を最小化するという基準を採用することとする. その理由は, 多くの場合 ζ に関する情報の方が重要であると考えられるからである.

先の議論から個別的な測定に限ると, 上記の条件を満たす推定量のための測定は個別にヘテロダイン測定を行うものに限られる. そのゆえ, この問題設定では測定を工夫するという余地が無い. 以下, N を推定する推定量の 2 つの候補 $\hat{N}_{1,n}, \hat{N}_{2,n}$ を挙げ, その性能を比較する. 個々のデータ α_i は平均値 ζ , 分散 $N+1$ の正規分布に従う. すなわち, その確率密度関数は $\frac{1}{\pi(N+1)} e^{-\frac{|\alpha-\zeta|^2}{N+1}}$ で与えられる. それゆえ, ζ の推定値 $\hat{\zeta}$ は得られたデータについての平均値で与えられる. 正規分布の分散の推定量は統計学で良く知られており, それを適用すると, N の推定量で不偏性を満た

すものの中では $\hat{N}_{1,n} := \frac{1}{n-1} \sum_{i=1}^n |\alpha_i - \hat{\zeta}|^2 - 1$ が最適である．もちろん， $n = 1$ の場合には，その分散を推定することは不可能なので， $n \geq 2$ であることが必要である．真のパラメータが ζ, N である場合の確率変数 X についての期待値を $E_{\zeta,N} X$ とすると，その MSE $E_{\zeta,N}(\hat{N}_{1,n} - N)^2$ は

$$\begin{aligned} E_{\zeta,N} \left(\frac{1}{n-1} \sum_{i=1}^n |\alpha_i - \hat{\zeta}|^2 - (N+1) \right)^2 &= E_{0,N} \left(\frac{1}{n-1} \sum_{i=1}^n |\alpha_i - \hat{\zeta}|^2 - (N+1) \right)^2 \\ &= E_{0,N} \left(\frac{1}{n-1} \sum_{i=1}^n |\alpha_i - \hat{\zeta}|^2 \right)^2 - (N+1)^2 = \frac{n^2}{(n-1)^2} E_{0,N} \left(\frac{1}{n} \sum_{i=1}^n |\alpha_i - \hat{\zeta}|^2 \right)^2 - (N+1)^2 \\ &= \frac{n^2}{(n-1)^2} E_{0,N} \left(\frac{1}{n} \sum_{i=1}^n |\alpha_i|^2 - |\hat{\zeta}|^2 \right)^2 - (N+1)^2 \\ &= \frac{n^2}{(n-1)^2} E_{0,N} \left(\left(\frac{1}{n} \sum_{i=1}^n |\alpha_i|^2 \right)^2 - 2 \frac{1}{n} \sum_{i=1}^n |\alpha_i|^2 |\hat{\zeta}|^2 + |\hat{\zeta}|^4 \right) - (N+1)^2 = \frac{(N+1)^2}{n-1} \end{aligned} \quad (8)$$

となる．もちろん， $(N+1)^2 > N(N+1)$ であるから，この値は (6) より大きい．

しかし，推定量は必ずしも不偏性条件を満たす必要は無く，代わりに推定量 $\hat{N}_{2,n} := \frac{1}{n} \sum_{i=1}^n |\alpha_i - \hat{\zeta}|^2 - 1$ を用いることもある．この推定量の期待値は $E_{\zeta,N}(\hat{N}_{2,n}) = \frac{n}{n-1}(N+1) - 1$ となるため，不偏性は満たさないが，MSE は

$$E_{\zeta,N}(\hat{N}_{2,n} - N)^2 = \frac{(N+1)^2}{n}$$

となり， $\hat{N}_{1,n}$ の MSE より一様に小さい．しかし，普通，統計では不偏性を満たす推定量 $\hat{N}_{1,n}$ を用いることが多い．このように ζ と N の同時推定では量子相関を用いない場合では，個々の系にヘテロダイン測定を行ない，その後適切なデータ処理を行なうことが有効な方法であり，そのときの ζ についての MSE は $\frac{N+1}{n}$ となり， N についてのそれぞれの MSE $E_{\zeta,N}(\hat{N}_{1,n} - N)^2$ と $E_{\zeta,N}(\hat{N}_{2,n} - N)^2$ はともに上記のように計算できる．

5.3 量子相関を用いた推定

次に量子相関を使う場合はどうなるであろうか？この場合以下のような方法により，量子相関を用いない場合よりも有効な推定が可能である．まず，各系同士をうまく相互作用させて，以下の様に $\rho_{\zeta,N}^{\otimes n}$ を時間発展させる．

$$\rho_{\zeta,N}^{\otimes n} \mapsto \rho_{\sqrt{n}\zeta,N} \otimes \rho_{0,N}^{\otimes n-1}. \quad (9)$$

そして，1 番目の系に対してはヘテロダイン測定を行い，その測定値をそのまま ζ の推定値とし，残りの $n-1$ 個の系に対しては個数測定を行い，サンプル平均を N の推定値とする．このように行くと， ζ に対しては先のものと同じ精度で推定することができ， N に関しては 1 個分の情報を失うだけで他の $n-1$ 個に対しては最適な測定を行うことができる [11]．この推定量は不偏性を満たし， ζ 及び N の MSE はそれぞれ， $\frac{N+1}{n}$ ， $\frac{N(N+1)}{n-1}$ となる． ζ の MSE については (7) と一

致し, N についてはほとんど (6) と一致して⁴, 量子相関を使わない場合の N の MSE(8) より小さい.

このような測定は $n = 2^m$ の時は以下のようにすると実現可能である. 例えば $n = 4$ のときには次の図 3 の様にビームスプリッター (ハーフミラー) を組み合わせた測定を行うことにより実現できる.

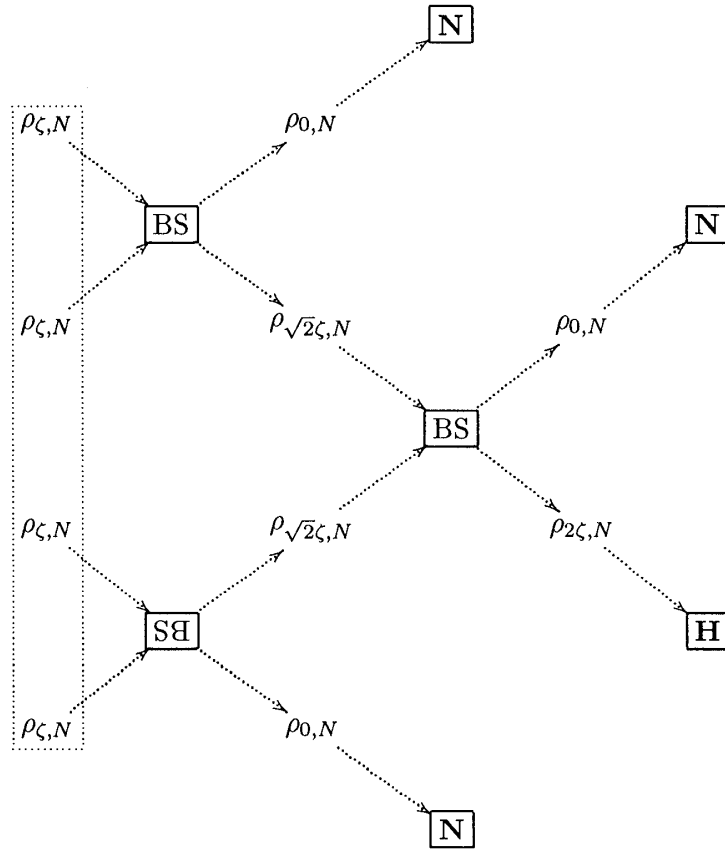


図 3: トーナメント方式による推定

図 3 では個数測定, ヘテロダイン測定をそれぞれ $|N\rangle$, $|H\rangle$ で表し, ビームスプリッターを $|BS\rangle$ で表した. ビームスプリッターとは上からの入射光の反射光は反射の際に位相が反転する (固定端反射) のに対し下からの入射光に対しては反射の際に位相がずれない (自由端反射) ものである. ビームスプリッター $|BS\rangle$ の上下を逆にしたものを $|SB\rangle$ で表した. ここで構成した相互作用の結果, 式 (9) のような時間発展が実現される. このように量子ガウス状態の 2 つのパラメータ ζ と N の同時推定では原理的には量子相関を使うことで量子相関を使わない最適な場合の精度を上回ることが可能である. 特にこの方法は, 個別に ζ や N を推定する場合の最適な精度を同時に実現している.

⁴ヘテロダイン測定の測定値が 1 つのみでは N に関しては全く情報を与えないことを考えると, N の推定に $\rho_{\sqrt{n}\zeta, N}$ の情報を用いないことは不思議ではない.

5.4 量子効果が表れる推定の実現に向けて（その1）

しかし、上述のようなトーナメント方式を用いると、多くの回数の相互作用を行わねばならず、実現はやや困難である。以下では比較的实现が容易な N の推定量 $\hat{N}_{7,n}, \hat{N}_{8,n}$ を提案し、先に与えた相関を用いない推定量 $\hat{N}_{2,n}$ とこれらの推定量を比較する。入力状態の数が $n = 2m$ の場合で2個の状態間の相関のみ用いる方法であれば実現が比較的容易であると考えられる。すなわち、上の実験で $n = 2$ の場合を m 回繰り返すわけである。当然、ヘテロダイン測定が m 回で、個数測定が m 回となる。ただ、残念なことに、ヘテロダイン測定はかなり理想的なものが実現できるが、個数測定の方は理想的なものを実現することは困難である。つまり、個数測定で実際にできるものはその効率（量子効率） t が1未満のものである。さらに、多くの場合 $k = 0$ とそれ以外の区別しかできず、（測定値の数が）2値の測定しかできない。付録Eで述べるように、そのような量子効率が t の個数測定の2値の POVM は

$$\left\{ \sum_{k=0}^{\infty} |k\rangle\langle k| (1-t)^k, I - \sum_{k=0}^{\infty} |k\rangle\langle k| (1-t)^k \right\} \quad (10)$$

で与えられる [12].

そこで以下では、ヘテロダイン測定と上述の不完全な個数測定を使った推定方法を考え、これらの性能について議論する。この場合 ζ の推定量はヘテロダイン測定の測定値 $\alpha_i (i = 1 \dots m)$ の平均値を用いて $\hat{\zeta} := \frac{1}{\sqrt{2}} \sum_{i=1}^m \alpha_i$ で与えられ、その MSE は $\frac{N+1}{2m}$ となる。また、 N の推定量であるが、ヘテロダイン測定の測定値のみを使うのであれば、その推定値は $\hat{N}_{1,m} := \frac{1}{m-1} \sum_{i=1}^m |\alpha_i - \sqrt{2}\hat{\zeta}|^2 - 1$ で与えられ、その MSE は $\frac{(N+1)^2}{m-1}$ となる。

一方、先の量子効率が t の不完全な個数測定を用いると、

$$\text{Tr} \sum_{k=0}^{\infty} |k\rangle\langle k| (1-t)^k \rho_{0,N} = \sum_{k=0}^{\infty} \frac{1}{N+1} \left(\frac{N}{N+1} \right)^k (1-t)^k = \frac{1}{1+Nt}$$

となることから、その測定値は2項分布

$$\left\{ p_N(0) = \frac{1}{1+Nt}, p_N(1) = \frac{Nt}{1+Nt} \right\} \quad (11)$$

に従う。すなわち、分布族 $\{p_N | N > 0\}$ についての推定を扱うことになる。それゆえ、この分布に関する m 回の独立同一試行の下で0が l 回得られた場合は、 $p_{\hat{N}_4}(0) = \frac{1}{1+\hat{N}_{4,2m}t} = \frac{l}{m}$ となるように推定値 $\hat{N}_4$ を $\hat{N}_{4,2m} := \frac{m-l}{lt}$ と決定することが自然に思える。しかし、この推定量を用いると $l = 0$ のときには推定値 $\hat{N}_{4,2m}$ が無限大に発散してしまう。そのために、

$$\hat{N}_{5,2m} := \begin{cases} \frac{m}{t} & l = 0 \text{ のとき} \\ \frac{m-l}{lt} & \text{上記以外} \end{cases}$$

を用いることとする。その他に、

$$\hat{N}_{6,2m} := \frac{m-l}{(l+1)t}$$

という推定量を考えることもできる. しかし, 2つの推定量 $\hat{N}_{5,2m}, \hat{N}_{6,2m}$ はともに, 不偏推定量ではない. それどころか, 変換 $N \mapsto \frac{1}{1+Nt}$ が非線型であるために, 分布族 $\{p_N | N > 0\}$ の不偏推定量を構成すること自体が困難である. しかし, 推定量が良いか悪いかは必ずしも, 不偏性を満たすか否かではない. むしろ, 平均2乗誤差 $v_N(\hat{N}_{5,2m}), v_N(\hat{N}_{6,2m})$ の大小で比較すべきである.

なお, のちに詳しく述べるが, 推定量 $\hat{N}_{6,2m}$ の方が, 推定量 $\hat{N}_{5,2m}$ よりも, その期待値が真のパラメータに近い. 今の場合, ヘテロダイン測定と個数測定の双方のデータが使えるので, それぞれの測定による推定量の凸結合によって作られる推定量 $\hat{N}_{7,2m} := \lambda \hat{N}_{1,m} + (1-\lambda) \hat{N}_{5,2m}$ ($1 > \lambda > 0$) または $\hat{N}_{8,2m} := \lambda \hat{N}_{1,m} + (1-\lambda) \hat{N}_{6,2m}$ ($1 > \lambda > 0$) を考え, これを量子相関を用いない推定量と比較する. (以下の議論では凸結合を取る片方の推定量が不偏性を満たすことが必要なので, $\hat{N}_{2,m}$ ではなく $\hat{N}_{1,m}$ を用いた.) 一般に2つの独立な確率変数 X_1, X_2 があり, 片方の期待値が0であるとき, その凸結合 $\lambda X_1 + (1-\lambda) X_2$ ($1 > \lambda > 0$) の2乗の期待値は $\lambda^2 v_1 + (1-\lambda)^2 v_2$ となる. ここで X_1, X_2 の2乗の期待値を v_1, v_2 とおいた. その最小値は $\lambda = \frac{v_2}{v_1+v_2}$ のときに得られ, $1/(\frac{1}{v_1} + \frac{1}{v_2})$ となる. それゆえ, $\hat{N}_{7,2m} - N$ 及び $\hat{N}_{8,2m} - N$ に対して, 最小となる λ を選ぶと, 推定量 $\hat{N}_{7,2m}, \hat{N}_{8,2m}$ の平均2乗誤差はそれぞれ, $V_2 = v_N(\hat{N}_{8,2m}) = \left(\frac{m-1}{(N+1)^2} + v_N(\hat{N}_{6,2m})^{-1} \right)^{-1}$, $V_3 = v_N(\hat{N}_{7,2m}) = \left(\frac{m-1}{(N+1)^2} + v_N(\hat{N}_{5,2m})^{-1} \right)^{-1}$ で与えられる.

以下の結果を繰返し回数 $m = 10$, 量子効率 $t = 0.5$ の場合に量子相関を使わないときの平均2乗誤差 $V_1 = v_N(\hat{N}_{2,2m}) = \frac{(N+1)^2}{2m}$ と比較すると図4のようになる. 参考のためサンプル数 n が大きいときの漸近的な場合の挙動を表すものとして $V_4 = (\lim_{m \rightarrow \infty} m v_N(\hat{N}_{7,2m}))/10$ との比較も行なった.

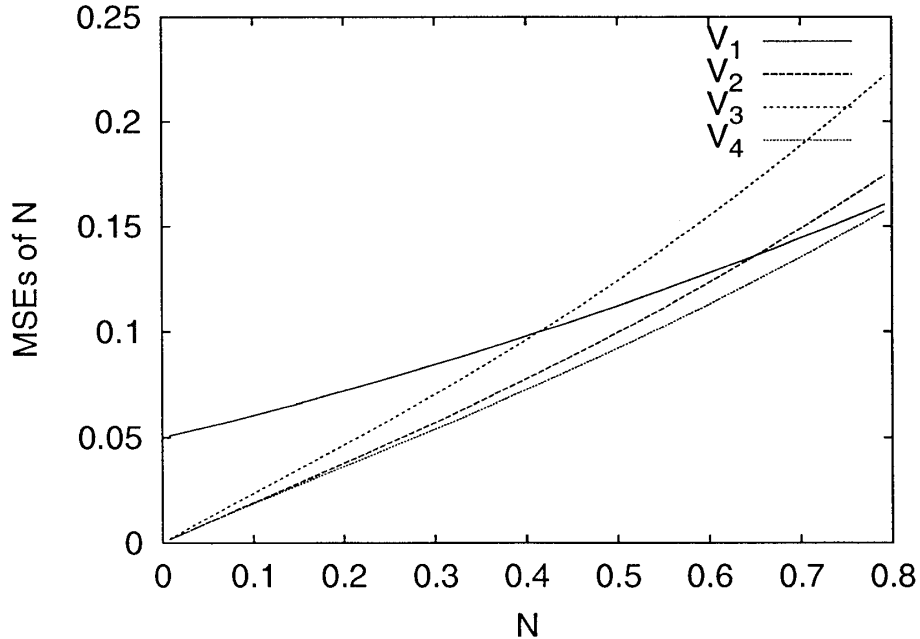


図 4: パラメータ N の平均2乗誤差 (MSE) の比較

ほぼ 0.65 より N が小さい場合では, 量子相関を使った推定量 $\hat{N}_{8,2m}$ は量子相関を使わなかった

推定量 $\hat{N}_{2,2m}$ より良い精度を与えている。また、同じ実験の下でも2つの異なるデータ処理に対応する推定量 $\hat{N}_{8,2m}$ と $\hat{N}_{7,2m}$ はかなり MSE が異なる。このことはこのような場合ではデータ処理の選択も無視できない要素であることを示している。

以上のように、現実の不完全な測定を考慮しても量子相関を用いることにより推定精度が向上する可能性が示された。可視光の場合には最も良い場合で量子効率 $t = 0.83$ という数字が記録されており [13], 市販の計測機器でも $t = 0.5$ ぐらいは可能である。確かに、ビームスプリッターで干渉させる実験は決して容易でなないが、現代の技術で十分可能である。今後、この測定を実装し、量子効果が状態推定にも有効であることが実験的に実証されることが期待される。

6 漸近論

先の節では、実際の実験に結び付く議論が必要であったために、十分にデータが多くない場合を扱った。このような場合では解析的に綺麗な形で、個々の推定量の MSE を与えることは困難であった。しかし、データ数が十分大きいと、比較的多くの場合について、 $\frac{1}{m^2}$ のオーダーでの推定量の MSE の振る舞いを解析的に扱うことが可能である。この節では一般のモデルでも展開できる、データが多い場合の議論（漸近論）を扱う。はじめに 6.1 節では、量子ガウス状態の場合での漸近論を扱い、6.2 節では古典系（確率分布族）のパラメータ推定での漸近論について簡単に紹介する。さらに 6.3 節では一般の量子状態のパラメータ推定の場合の漸近論について簡単に紹介する。

6.1 量子ガウス状態での漸近論

量子系 $\mathcal{H}$ の状態族 $\{\rho_\theta \in \mathcal{S}(\mathcal{H}) | \theta \in \Theta \subset \mathbb{R}\}$ の推定では、そこで準備される系の数 n が十分大きいときの精度を議論することが多い。このような場合は MSE についての $\frac{1}{m}$ で巾展開を行ない、その係数で推定量の優劣を比較することになる。例えば、量子ガウス状態について先に扱った推定量の MSE について $\frac{1}{m}$ で巾展開すると

$$\begin{aligned} v_N(\hat{N}_{1,2m}) &= (N+1)^2 \frac{1}{2m} \\ v_N(\hat{N}_{2,2m}) &\cong (N+1)^2 \frac{1}{2m} + (N+1)^2 \frac{1}{(2m)^2} \\ v_N(\hat{N}_{7,2m}) &\cong 2C_{N,t} \frac{1}{2m} + 4C_{N,t}^2 \left(\frac{t(7Nt+2)}{N(Nt+1)^2} + \frac{1}{(N+1)^2} \right) \frac{1}{(2m)^2} \\ v_N(\hat{N}_{8,2m}) &\cong 2C_{N,t} \frac{1}{2m} + 4C_{N,t}^2 \left(\frac{2t^2}{(Nt+1)^2} + \frac{1}{(N+1)^2} \right) \frac{1}{(2m)^2} \\ C_{N,t} &:= \left(\frac{t}{N(1+Nt)^2} + \frac{1}{(N+1)^2} \right)^{-1} \end{aligned}$$

と計算できる。これらの推定量の優劣を大まかに議論する場合は、はじめに第一項の係数 ($\frac{1}{2m}$ の係数) に注目し、これが同じ場合、第二項の係数 ($\frac{1}{(2m)^2}$ の係数) に注目する。それゆえ、量子相関を用いないときの MSE $v_N(\hat{N}_{1,2m})$, $v_N(\hat{N}_{2,2m})$ 及び量子相関を用いたときの MSE $v_N(\hat{N}_{7,2m})$,

$v_N(\hat{N}_{8,2m})$ を比較するには, $(N+1)^2$ と $2C_{N,t}$ の大小関係が問題となる. $(N+1)^2 > 2C_{N,t}$ となるための必要十分条件は逆数を考えることにより, $\frac{1}{(N+1)} < \frac{t}{N(1+Nt)^2}$ となる. これは, $N < 1$ かつ $t > \frac{-N^2+2N+1-(N+1)\sqrt{1+2N-3N^2}}{2N^3}$ で与えられる. この条件を満たす領域は図5のグラフより上の領域であり, この領域ではここで述べたタイプの量子相関を用いることで推定精度が上昇する.

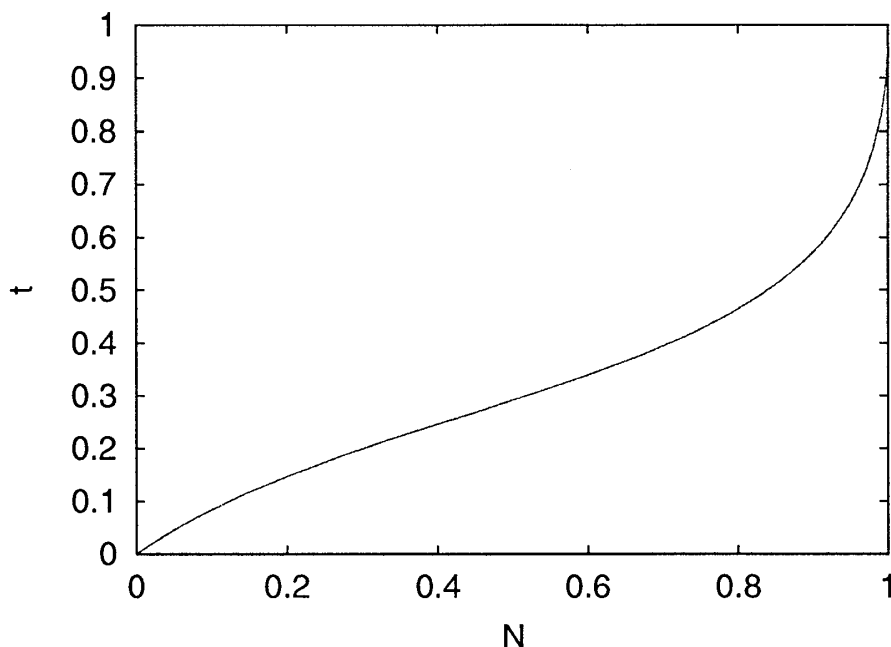


図 5: 漸近的に相関を用いた効果が現れる条件

次に関心のあるのが, $v_N(\hat{N}_{7,2m})$ と $v_N(\hat{N}_{8,2m})$ の第二項の係数である. これらの大小関係は $\frac{t(7Nt+2)}{N(Nt+1)^2}$ と $\frac{2t^2}{(Nt+1)^2}$ の大小関係で決まるが, $\frac{7Nt+2}{Nt} > 2$ なので, 前者の方が常に大きい. それゆえ, $\frac{1}{m^2}$ のオーダーまで注目した場合, $\hat{N}_{8,2m}$ の方が優れていると言える. 詳しくは付録 B.3 を参照のこと.

さらに, 一般の量子系 $\mathcal{H}$ の状態族 $\{\rho_\theta \in \mathcal{S}(\mathcal{H}) | \theta \in \Theta \subset \mathbb{R}\}$ の推定において, そこで準備される系の数 n が十分大きい場合では $\frac{1}{m}$ の巾展開での第一項の係数 (上の例では $2C_{N,t}$ や $(N+1)^2$) すなわち,

$$\lim_{n \rightarrow \infty} n \int (\hat{\theta}_n(\omega) - \theta)^2 \text{Tr} \rho_\theta^{\otimes n} M^n(d\omega)$$

を用いて議論することになる. パラメータの数が複数のときは代わりに以下の極限值

$$\lim_{n \rightarrow \infty} n \int \sum_i (\hat{\theta}_n^i(\omega) - \theta^i)^2 \text{Tr} \rho_\theta^{\otimes n} M^n(d\omega)$$

を議論することが多い. そして統計的推測では, 単に個々の推定量の列について上記の極限を考えるだけではなく, その限界, すなわち, 上記の極限がどこまで小さくできるかということを問題にする. ガウス状態族の ζ, N の推定で量子相関を用いる場合ではその限界値は $(N+1)^2$ となることが示せる. 逆に, 量子相関を用いない場合はこれを達成することはできないことも示せる.

6.2 古典系での推定の漸近論

古典的な場合には推定量の MSE についての $\frac{1}{n}$ の巾展開での第 1 項の係数の最適化（1 次漸近論）についての一般論が既に確立している．パラメータが 1 つの確率分布族 $\{p_\theta(\omega)|\theta \in \Theta \subset \mathbb{R}\}$ のパラメータ推定で，未知の分布 p_θ に独立に従う n 個のデータ $\vec{\omega}_n := (\omega_1, \dots, \omega_n)$ から未知パラメータ θ を決定するという設定では，最尤推定 $\hat{\theta}_{ML,n}(\vec{\omega}_n) := \operatorname{argmax}_{\theta \in \Theta} p_\theta(\omega_1) \cdots p_\theta(\omega_n)$ を用いることが多い⁵．データ数 n が大きいとき，最尤推定 $\hat{\theta}_{ML,n}$ の MSE は漸近的に 0 に収束するが，その早さが問題となり，次の漸近的な近似式

$$v_\theta(\hat{\theta}_{ML,n}) \cong J_\theta^{-1} \frac{1}{n}, \quad J_\theta := \int \left(\frac{d}{d\theta} \log p_\theta(\omega) \right)^2 p_\theta(\omega) d\omega \quad (12)$$

が成り立つ．なお， J_θ は Fisher 情報量とよばれる．逆に任意の推定量の列 $\{\hat{\theta}_n(\vec{\omega}_n)\}$ に対して，極限值 $\lim_{n \rightarrow \infty} n v_\theta(\hat{\theta}_n)$ が $\theta \in \Theta$ について一様収束するのであれば，

$$\lim_{n \rightarrow \infty} n v_\theta(\hat{\theta}_n) \geq J_\theta^{-1} \quad (13)$$

となる．それゆえ，漸近一次の議論の範囲では最尤推定量が最適な推定量となり，推定量として最適なものを選んだ場合の推定量の精度を表す量として Fisher 情報量の値に注目することが多い．

ここで確率空間 Ω から確率空間 Ω' への写像 f を考えると，自然に確率空間 Ω' 上の分布 p'_θ を $p'_\theta(B) := p_\theta(f^{-1}(B))$ (B は Ω' の部分集合) で定義できる．この分布 p'_θ のなす分布族の Fisher 情報量を J'_θ で表すと，以下の Fisher 情報量の単調性とよばれる不等式が成り立つ．

$$J'_\theta \leq J_\theta \quad (14)$$

この不等式が成り立つことは，Fisher 情報量のもつ意味（式 (12) 及び不等式 (13)）を考えれば次に述べるように容易に確認できる．はじめに，確率分布族 $\{p'_\theta\}$ に対する推定量 $\hat{\theta}'$ は推定量 $\hat{\theta}(\omega) := \hat{\theta}'(f(\omega))$ を考える．すると，自然に確率分布族 $\{p_\theta\}$ に対する推定量と見なすことができる．2 つの推定量 $\hat{\theta}$ と $\hat{\theta}'$ の平均 2 乗誤差は等しいので，分布族 $\{p'_\theta\}$ の最適な場合の平均 2 乗誤差は分布族 $\{p_\theta\}$ の最適な場合の平均 2 乗誤差より大きいことがわかる．これより，不等式 (14) が確認できる．しかし，極限值 $\lim_{n \rightarrow \infty} n v_\theta(\hat{\theta}_n)$ が J_θ^{-1} と一致する推定量は最尤推定量だけではない，漸近的に最適な推定量は他にも複数存在する．それらの中で優劣を扱うには，その次のオーダーの係数を議論することが多い．これについて詳しくは付録 B.3 を参照のこと．

6.3 量子系での推定の漸近論

次にパラメータが 1 つの量子状態族 $\{\rho_\theta \in \mathcal{S}(\mathcal{H})|\theta \in \Theta \subset \mathbb{R}\}$ のパラメータ推定について考えてみよう．この場合は測定に対応する POVM $M = \{M_i\}$ の選択の自由度があるわけである．それゆえ， M に応じて決まる確率分布族 $\{P_{\rho_\theta}^M|\theta \in \Theta\}$ のパラメータ推定に帰着できる．以下ではこの分布族の Fisher 情報量を J_θ^M で表す．先と同様に，未知のパラメータ θ に応じて決まる同一の状態 ρ_θ が n 個独立に準備されている状況を考えるわけであるが，この場合その表現空間は

⁵ $\operatorname{argmax}_x f(x)$ は $f(x)$ の最大値を与える x を表す．

$\mathcal{H}^{\otimes n}$ になる。例えば個々の系に対して、同一の測定 M を行い、確率分布族 $\{P_{\rho_\theta}^M | \theta \in \Theta\}$ に対する最尤推定量 $\hat{\theta}_{ML,n}$ を適用した場合の MSE は $v_\theta(M \times n, \hat{\theta}_{ML,n}) \cong (J_\theta^M)^{-1} \frac{1}{n}$ となる。 ($M \times n$ で $\mathcal{H}^{\otimes n}$ の各量子系 $\mathcal{H}$ に対して独立に POVM M を行うという測定を表すこととする。) しかも、先と同様に推定量の列 $\{(M^n, \hat{\theta}_n)\}$ に対して、極限值 $\lim_{n \rightarrow \infty} n v_\theta(M^n, \hat{\theta}_n)$ が $\theta \in \Theta$ について一様収束するのであれば、 $\lim_{n \rightarrow \infty} n v_\theta(\hat{\theta}_{ML,n}) \geq \min_M (J_\theta^M)^{-1}$ となる。それゆえ、 J_θ^M を最大にする測定 M を行い、最尤推定量を用いる方法が漸近一次の議論の範囲では最適となる。しかし、この戦略には J_θ^M を最大化する最適な測定 M が一般には未知のパラメータ θ に依存するという致命的な欠陥がある。

この問題については以下の戦略を適用すると、漸近的に限界 $\min_M (J_\theta^M)^{-1}$ を達成できることが知られている。同一状態にある n 個の量子系が準備されたとき、まずはじめの $\sqrt{n}$ 個の量子系に対して任意の θ に対して $J_\theta^{M_0} > 0$ となる M_0 を行い、それによって得られる確率分布族 $\{P_{\rho_\theta}^{M_0} | \theta \in \Theta\}$ に対する最尤推定を行いその推定値を θ_0 とする。次に、残りの $n - \sqrt{n}$ 個の量子系に $M_1 := \operatorname{argmax}_M J_{\theta_0}^M$ となる測定を行い、確率分布族 $\{P_{\rho_\theta}^{M_1} | \theta \in \Theta\}$ に対する最尤推定を最終的な推定値とする。このようにして構成された推定量は漸近的に限界 $\min_M (J_\theta^M)^{-1}$ を達成する [6, 14]。さらに、この値は以下に定義する SLD Fisher 情報量 J_θ の逆数と一致する：

$$J_\theta := \operatorname{Tr} L_\theta^2 \rho_\theta, \quad \frac{d\rho_\theta}{d\theta} = \frac{1}{2} (L_\theta \rho_\theta + \rho_\theta L_\theta).$$

ここで L_θ は対称対数微分 (SLD) とよばれ、エルミート行列（または自己共役作用素）である。また、 $J_\theta = J_\theta^M$ となる測定は対称対数微分 L_θ のスペクトル分解に対応する測定で与えられる。すなわち、 L_θ の対角化が $L_\theta = \sum_k l_k E_k$ (l_k は固有値、 E_k は射影) で与えられるとすると、 M が測定 $\{E_k\}$ のとき、 $J_\theta = J_\theta^M$ となる。それゆえ、このような測定 $\{E_k\}$ が未知パラメータに依存しないモデルであれば、解析が極めてやさしい。

7 量子2準位系の固有値推定

この節では、もう一つの量子効果の検証となる実験として、量子2準位系の固有値推定を扱う。7.2 節では量子相関を用いない場合での測定の最適化について議論し、そして、7.3 節では量子相関を用いた場合の測定について議論する。7.4 節では中間的な場合として古典的な相関を用いた場合について議論する。最後に7.5 節では現実の測定を扱うときに必要な注意や、そのような現実の不完全な測定でも量子相関を用いることで量子相関を用いない推定を部分的に上回る性能が得られることを示す。

量子系 $\mathbb{C}^2$ は量子2準位系とよばれ、最も基本的な量子系である。この系の密度行列は 2×2 のエルミート行列で表されるが、以下のようなパラメータを取ることが可能である。

$$\rho_{r,\theta} := \frac{1}{2} \begin{pmatrix} 1 + r \cos \theta_1 & r \sin \theta_1 e^{i\theta_2} \\ r \sin \theta_1 e^{-i\theta_2} & 1 - r \cos \theta_1 \end{pmatrix}, \quad 0 \leq \theta_1 \leq \frac{\pi}{2}, 0 \leq \theta_2 \leq 2\pi.$$

この場合、 r を決めることで $\rho_{r,\theta}$ の固有値がきまり、 $\theta = (\theta_1, \theta_2)$ を決めることで $\rho_{r,\theta}$ の固有ベクトルが決まる。以下では、パラメータ r のみを推定する問題を考える。この問題では θ を推定す

る必要が無いが、これが未知であれば、MSE が SLD Fisher 情報量の逆数となる推定量を構成する測定は θ に依存する。そのため、そのような精度を持つ推定のためには測定を適応的に改良する必要がある。実際の実験では、適応的に測定を改良することが困難である場合が多い。それゆえ、以下では未知状態 $\rho_{r,\theta}$ にある量子系 $\mathbb{C}^2$ を $2m$ 個、独立に準備し、各系に独立に同一測定 M を行い r を推定する場合と、未知状態 $\rho_{r,\theta}^{\otimes 2}$ にある量子系 $(\mathbb{C}^2)^{\otimes 2}$ を m 個、独立に準備し、各系に独立に同一測定 M を行い r を推定する場合を比較する。前者で最適な場合の MSE と、後者で最適な場合の MSE の差が、量子相関を用いることによるメリットとみなすことができる。

7.1 基底が分かっている場合の固有値の推定

はじめに $|\theta\rangle := \cos \theta_1 |+\rangle + e^{i\theta_2} \sin \theta_1 |-\rangle$, $|\theta^\perp\rangle := \sin \theta_1 |+\rangle - e^{i\theta_2} \cos \theta_1 |-\rangle$, $|+\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $|-\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ と置くと, $\rho_{r,\theta} = \frac{1+r}{2} |\theta\rangle\langle\theta| + \frac{1-r}{2} |\theta^\perp\rangle\langle\theta^\perp|$ となる。それゆえ, θ に対応する基底 $|\theta\rangle, |\theta^\perp\rangle$ を用いて測定すると, 固有値に対応する二項分布 $\frac{1+r}{2}, \frac{1-r}{2}$ が得られる。実際, θ を固定したときに得られる量子状態族 $\{\rho_{r,\theta} | 0 < r < 1\}$ の SLD Fisher 情報量は $\frac{1}{1-r^2}$ であり, n 個の同一状態 $\rho_{r,\theta}$ にある量子系にこの測定を行い, $|\theta\rangle$ に対応する測定値が k 個得られた場合は, $\frac{k}{n} = \frac{1+r}{2}$ と推定値 $\hat{r}_n$ を定めると, その MSE は $\frac{1-r^2}{n}$ となる。

7.2 量子相関を用いない場合の最適化

以下では θ が未知である場合ので r の推定で, 量子相関を用いない場合を考え, 1 次漸近論の枠組で最適化を行う。測定 M を用いて推定する場合を考えよう。すると, 収束に関する一様性を満たす推定量 $\hat{r}_n$ を用いると仮定すると, その MSE は

$$\lim_{n \rightarrow \infty} n v_{r,\theta}(M \times n, \hat{r}_n) \geq (J_{r,\theta,r}^M)^{-1} \quad (15)$$

となる。ここで, $J_{r,\theta,r}^M$ は θ と測定 M に依存した r が未知パラメータである確率分布族 $\{\rho_{r,\theta}^M | 0 < r < 1\}$ の Fisher 情報量である。不等式 (15) は θ が未知である場合は 6.2 節の議論から直接導かれる。それゆえ, θ が既知である場合も成り立つ。さらに, 付録 B.2 の最後で述べた事実から, 適切な古典的な推定量 $\hat{r}_n$ を選ぶと (15) で等号が成立する。

このモデルではパラメータ r を固定すると, パラメーター θ のなす空間は 3 次元球面になり, ユニタリ群の作用は球面上の回転に相当する。すなわち, $SU(2)$ の要素 U に対して, $U \rho_{r,\theta} U^\dagger = \rho_{r,O(U)\theta}$ となる 3 次元空間上の回転 $O(U)$ が対応する。そのような θ に関する対称な構造に注目し, θ についての最悪な場合の MSE すなわち $\max_\theta \lim_{n \rightarrow \infty} n v_\theta(M \times n, \hat{r}_n)$ に注目し, これを最適化する問題を考える。この手法はミニマックス法とよばれ, 統計学などでもしばしば用いられる。

ここでパラメーター θ のなす空間に全測度が 1 となる回転に関して不変な測度 $\mu(d\theta)$ を考える。

$$\mu(d\theta) := \frac{\sin \theta_1 \cos \theta_1}{\pi} d\theta_1 d\theta_2$$

すると、付録 C.2 で示すように、不等式

$$\max_{\theta} \lim_{n \rightarrow \infty} n v_{r,\theta}(M \times n, \hat{r}_n) \geq \left(\int J_{r,\theta;r}^M \mu(d\theta) \right)^{-1} \quad (16)$$

が成立し、測定 M がある種の対称性を持ち、 $\hat{r}_n$ が最尤推定であるとき、等号が成り立つ。

この問題は付録 C.2 で示すように、より一般的な設定の下で、群共変的測定とよばれるある種の対称性を満たす測定 M に限った場合の $\int J_{r,\theta;r}^M \mu(d\theta)$ の最大化問題に帰着させることができる。(群共変的測定では $J_{r,\theta;r}^M$ は θ に依存しないので、 $J_{r,\theta;r}^M$ を最大化させる問題とみることもできる。) しかも、付録 C.3 で示すように、今扱っている問題では群の作用により全ての純粋状態が互いに移り合うことから、測定 $M_{opt}(d\theta) := 2|\theta\rangle\langle\theta|\mu(d\theta)$ を考えると、

$$\int J_{r,\theta;r}^M \mu(d\theta) \leq J_{r,\theta;r}^{M_{opt}} \quad (17)$$

が成り立つ。 M_{opt} の定義での係数 2 は空間の次元に対応するものである。これを掛けないと全測度が単位行列にならない。 $J_{r,\theta;r}^{M_{opt}}$ は後に計算で確かめるように θ に依存しない。この測定 M_{opt} は群共変性を満たすので、付録 C.1 の補題 1 より、最尤推定量 $\hat{r}_{ML,n}$ について、

$$\max_{\theta} \lim_{n \rightarrow \infty} n v_{r,\theta}(M_{opt} \times n, \hat{r}_n) = \lim_{n \rightarrow \infty} n v_{r,\theta_0}(M_{opt} \times n, \hat{r}_n) = \left(J_{r,\theta_0;r}^{M_{opt}} \right)^{-1} \quad (18)$$

が成り立つ。この式は、今考えている設定では測定 M_{opt} を行い、最尤推定を行う推定方法が漸近的に最適であることを示している。なお、測定 M_{opt} は確率 $\mu(\theta)$ で θ 方向のスピンを測定すなわち、 $\{|\theta\rangle\langle\theta|, |\theta^\perp\rangle\langle\theta^\perp|\}$ を行なうことで実現できる。この測定は純粋状態からなる状態族において真の状態を推定する問題でも漸的に最適測定であることが示されている [15, 16]。

この他に統計学では未知パラメータに事前分布を仮定し、その事前分布についての平均値に注目するという問題設定があり、これが Bayes 法とよばれ、しばしば採用される。この問題に先の不変測度 $\mu(d\theta)$ を事前分布として Bayes 法を適用すると、

$$\lim_{n \rightarrow \infty} n \int v_{r,\theta}(M \times n, \hat{r}_n) \mu(d\theta)$$

に注目することになる。先の議論から積分と極限を入れ替えると、不等式

$$\lim_{n \rightarrow \infty} n \int v_{r,\theta}(M \times n, \hat{r}_n) \mu(d\theta) \geq \int (J_{r,\theta}^M)^{-1} \mu(d\theta) \geq \left(\int J_{r,\theta}^M \mu(d\theta) \right)^{-1} \geq \left(J_{r,\theta;r}^{M_{opt}} \right)^{-1}$$

が成り立つ。なお、上式を示す際に 2 つ目の不等式では関数 $x \mapsto 1/x$ が下に凸であることに注目し付録 A で示す Jensen の不等式 (29) を用いた。

このように Bayes の問題設定でも $J_{r,\theta;r}^{M_{opt}}$ が最適値を与える。それゆえ、ミニマックス法と Bayes 法の双方の問題設定で測定 M_{opt} を行い、最尤推定を用いる方法が最適となる。さらに、 $\langle\theta|\rho_{r,0}|\theta\rangle = \frac{1}{2}(1 + r \cos 2\theta_1)$ と計算できることから、そのときの Fisher 情報量は

$$\begin{aligned} J_{r,\theta;r}^{M_{opt}} &= \int_0^{2\pi} \int_0^{\pi/2} \left(\frac{d}{dr} \log(1 + r \cos 2\theta_1) \right)^2 \frac{(1 + r \cos 2\theta_1) 2 \sin \theta_1 \cos \theta_1}{2\pi} d\theta_1 d\theta_2 \\ &= \frac{1}{2} \left(-\frac{2}{r^2} + \frac{1}{r^3} \log \left(\frac{1+r}{1-r} \right) \right) \end{aligned}$$

と計算できる。

7.3 量子相関を用いる場合の最適化

5節で議論したように、1回だけ量子相関を持つような相互作用をさせる場合との比較を行う。量子系 $\mathbb{C}^4 = (\mathbb{C}^2)^{\otimes 2}$ 上の状態族 $\{\rho_{r,\theta}^{\otimes 2}\}$ におけるパラメータ r を推定する問題を考える。この場合、群 $SU(2)$ の要素 U による作用 $U^{\otimes 2}$ の下で、この状態族は閉じていることが示せる。それゆえ、先の節と同種の議論により、群共変的測定の中で、 $J_{r,\theta;r}^M$ を最大化する問題に帰着できる。詳しくは付録 C.2 を参照せよ。なお、先の場合と異なり、 $J_{r,\theta;r}^M$ を最大化する共変的測定 M は r に依存すると思われる。

$(\mathbb{C}^2)^{\otimes 2}$ は $\{|+, +\rangle, \frac{1}{\sqrt{2}}(|+, -\rangle + |-, +\rangle), |-, -\rangle\}$ で張られる空間 $\mathcal{H}_s$ と $\frac{1}{\sqrt{2}}(|+, -\rangle - |-, +\rangle)$ で張られる 1次元空間 $\mathcal{H}_a$ の直和 $\mathcal{H}_s \oplus \mathcal{H}_a$ に既約分解できる。例えば、極めて簡単な群共変的測定として、 $\mathcal{H}_s$ と $\mathcal{H}_a$ への射影 P_s, P_a からなる 2 値測定 M_1 がある。この測定の場合、 P_s に対応する測定値が得られる確率は $\frac{3+r^2}{4}$ であり、 P_a に対応する測定値が得られる確率は $\frac{1-r^2}{4}$ である。従って、Fisher 情報量 $J_{r,\theta}^{M_1}$ は $\frac{4r^2}{(1-r^2)(3+r^2)}$ である。

もちろん、この問題での推定精度の限界は、群共変的測定の中での $J_{r,\theta}^M$ の最大値である。この最大化問題を議論するには、 $\mathcal{H}_a$ は 1次元であるので付録 C.3 の (43) より、以下に述べる群共変的な測定に限って最適化すれば十分であることが確認できる。すなわち、測定値集合が $SU(2)$ と $\{a\}$ の和集合 $SU(2) \amalg \{a\}$ であり、 $\mathcal{H}_s$ 上の状態 ρ_s について決まる測定 M^{ρ_s} :

$$M^{\rho_s}(dg) = 3g^{\otimes 2}\rho_s(g^{\otimes 2})^\dagger \nu(dg), \quad M^{\rho_s}(a) = P_a$$

について、 $J_{r,\theta}^{M^{\rho_s}}$ の最大値を議論することになる。この測定は、はじめに測定 $\{P_a, P_s\}$ を行い、 P_s に対応する測定値が得られたときは、測定 $3g^{\otimes 2}\rho_s(g^{\otimes 2})^\dagger \nu(dg)$ を行うことを意味する。

この最大化問題は難しいので、以下では ρ_s が $|+, +\rangle$ の場合 (M_2)、 ρ_s が $\frac{1}{\sqrt{2}}(|+, -\rangle + |-, +\rangle)$ の場合 (M_3)、 ρ_s が $|-, -\rangle$ の場合 (M_4)、の場合、さらに、 $M_5 = \frac{1}{3}(M_2 + M_3 + M_4)$ の場合を比較する。ここで扱う測定のうち M_2, M_3, M_4 は、それぞれ対応する状態が群の表現に関して不変部分群 S_1 (1次元円周) を持つので、測定値集合は $SU(2) = S^3$ (3次元球面) を S_1 で割って得られる集合 $S^3/S^1 = S^2$ (2次元球面) となる。しかし、多くの場合では ρ_s についての不変部分群が単位元だけとなり、この場合、 ρ_s から生成される測定の測定値集合は群 $SU(2)$ となる。なお、 M_5 は混合状態から生成される測定、(すなわち ρ_s が混合状態の場合の測定) ではなく、純粋状態から生成される測定 M_2, M_3, M_4 の確率的混合である。一般に、このような M_2, M_3, M_4 などの群共変的な POVM に対応する測定の実現方法は必ずしも明確ではないが、その混合である M_5 の実現方法は以下に示すように明確である。先の定義では M_5 は M_2, M_3, M_4 の確率的混合として与えられたが、別の解釈も存在する。 M_5 は θ に対して、直交基底系 $|\theta, \theta\rangle, \frac{1}{\sqrt{2}}(|\theta, \theta^\perp\rangle + |\theta^\perp, \theta\rangle), \frac{1}{\sqrt{2}}(|\theta, \theta^\perp\rangle - |\theta^\perp, \theta\rangle), |\theta^\perp, \theta^\perp\rangle$ に対応する測定を考え、それを確率 $\mu(\theta)$ で混合した測定でもある。それゆえ、後者の確率的混合として M_5 を実現方法を採用することで、 M_5 は M_1 の次に容易に実現できることと思われる。

M_2 の場合は $\langle \theta, \theta | \rho_{r,0} | \theta, \theta \rangle = \frac{1}{4} (1 + r \cos 2\theta_1)^2$ となるので, Fisher 情報量は

$$\begin{aligned} J_{r,\theta;r}^{M_2} &= \int_0^{2\pi} \int_0^{\pi/2} \left(\frac{d}{dr} \log (1 + r \cos 2\theta_1)^2 \right)^2 \frac{(1 + r \cos 2\theta_1)^2 3 \sin \theta_1 \cos \theta_1}{4\pi} d\theta_1 d\theta_2 \\ &\quad + \left(\frac{d}{dr} \log \left(\frac{1-r^2}{4} \right) \right)^2 \frac{1-r^2}{4} \\ &= \frac{1}{1-r^2} \end{aligned}$$

と計算できる. ちなみに, 7.1 節で議論したパラメータ θ が既知である場合の Fisher 情報量は, 考えている量子系が $(\mathbb{C}^2)^{\otimes 2}$ なので, この値の 2 倍の $\frac{2}{1-r^2}$ となる. なお, M_4 については M_2 と実質的に同じである.

一方, M_3 の場合は $\frac{1}{\sqrt{2}} (\langle \theta, \theta^\perp | + \langle \theta^\perp, \theta |) \rho_{r,0} \frac{1}{\sqrt{2}} (|\theta, \theta^\perp\rangle + |\theta^\perp, \theta\rangle) = \frac{1}{4} (1 + r^2(-1 + 2 \sin^2 2\theta_1))$ となるので,

$$\begin{aligned} J_{r,\theta;r}^{M_3} &= \int_0^{2\pi} \int_0^{\pi/2} \left(\frac{d}{dr} \log (1 + r^2(-1 + 2 \sin^2 2\theta_1)) \right)^2 \frac{(1 + r^2(-1 + 2 \sin^2 2\theta_1)) 3 \sin \theta_1 \cos \theta_1}{4\pi} d\theta_1 d\theta_2 \\ &\quad + \left(\frac{d}{dr} \log \left(\frac{1-r^2}{4} \right) \right)^2 \frac{1-r^2}{4} \\ &= \frac{1}{1-r^2} - \frac{3}{r^2} + \frac{3}{2\sqrt{2}r^3\sqrt{1+r^2}} \log \left(\frac{\sqrt{1+r^2} + \sqrt{2}r}{\sqrt{1+r^2} - \sqrt{2}r} \right) \end{aligned}$$

と計算できる. したがって, M_5 の Fisher 情報量 $J_{r,\theta;r}^{M_5}$ は以下ようになる.

$$J_{r,\theta;r}^{M_5} = \frac{1}{1-r^2} - \frac{1}{r^2} + \frac{1}{2\sqrt{2}r^3\sqrt{1+r^2}} \log \left(\frac{\sqrt{1+r^2} + \sqrt{2}r}{\sqrt{1+r^2} - \sqrt{2}r} \right).$$

以上のように, M_1 から M_5 までの 5 種類の量子相関を用いた測定による精度が得られた.

ここで M_0 を先の相関を用いない測定を 2 度行うことによる測定として, 上記の測定による Fisher 情報量と比較したものが図 6 である. すなわち, 図 6 では M_0 は $J_{r,\theta;r}^{M_{opt}}$ の 2 倍を表す. なお, これらを直接比較しても, r が 1 に近いときはどれも発散するので, これらに $1-r$ を掛けたものを示した. その結果, 個別的に M_0 を行なう場合と比較的実現が容易である量子相関を用いた測定である M_1 を比較すると, r が 0.7 より大きい場合, すなわち非常に純粋状態に近い場合には後者の方が推定精度が良くなることが分かった. M_1 の次に実現が容易と考えられる M_5 は一様に M_0, M_1 より優れた推定精度を与えることも確認できる. しかし, これら M_0, M_1, M_2, M_3, M_5 の間で他のどれよりも全ての点で小さい推定精度与えるものは無く, その優劣は未知のパラメータ r に依存する.

ここまでは, 用いる量子相関の数を 2 に限定していたが, 原理的には準備された全系にまたがる相関を用いて推定することが可能である. このような場合の推定を Keyl & Werner[17] や松本[18] は $SU(2)$ の表現の既約成分に射影する測定で行うことを考えた. 正確には, 彼らは一般の次元での固有値推定問題を扱ったので, $SU(n)$ の表現論を用いた. 彼らは, そのときの裾確率に注目し, その漸近的な減少率を評価した. 彼らが導出した減少率は漸近的には最適である [19, 20, 21]. 松本[18] は同時に MSE についても議論した.

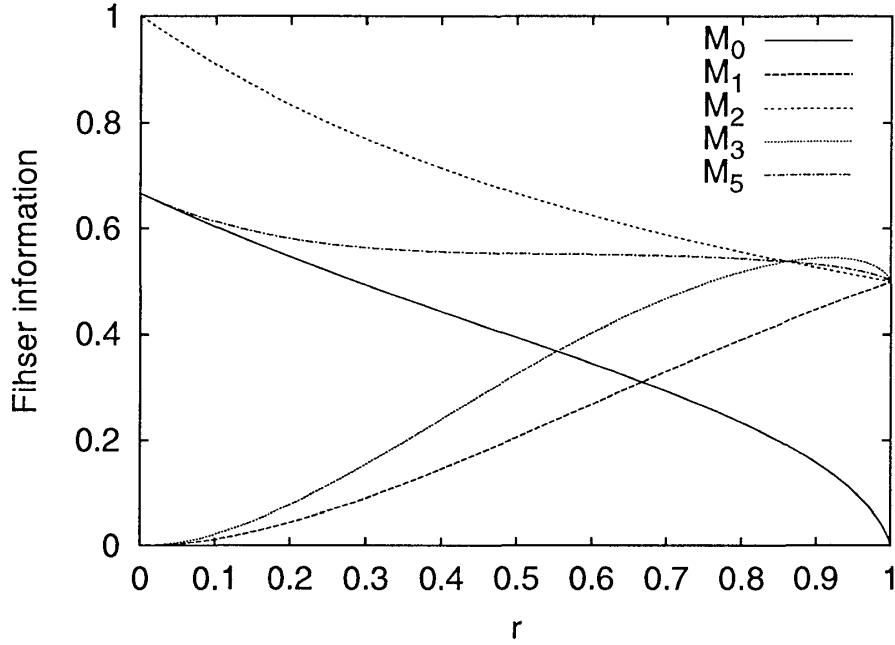


図 6: Fisher 情報量についての M_0, M_1, M_2, M_3, M_5 の比較

7.4 古典的な相関を用いた場合

7.3 節では $(\mathbb{C}^2)^{\otimes 2}$ 上の測定で量子的な相関を用いた場合と全く相関を用いなかった場合とを、群共変的測定の Fisher 情報量 $J_{r,\theta;r}^M$ を基準にして比較した。だが、古典的な相関を用いることで、全く相関を用いなかった場合の最適値 $J_{r,\theta;r}^{M_{opt}}$ を上回る可能性は否定できない。ところが、以下に示すように、古典的な相関を用いてもこの問題設定の下では $J_{r,\theta;r}^{M_{opt}}$ を上回ることはない。ここで古典的な相関という言葉の定義を明らかにせずに用いたが、図 1 で与えた適応的な POVM などがそれにあたる。ただ、以下ではもう少し広いクラスとして separable な POVM の範囲で考える。 $\mathcal{H}^{\otimes 2}$ 上の POVM $M = \{M_\omega\}$ が

$$M_\omega = M_\omega^1 \otimes M_\omega^2 \quad (19)$$

と書けるとき separable とよばれる。 $(\mathbb{C}^2)^{\otimes 2}$ の場合は考えている群が $SU(2)$ であり、群共変的な測定 M はいくつかの純粋なテンソル積状態 $\rho_{1,k} \otimes \rho_{2,k}$ と k についての確率 p_k を用いて

$$M(k, g) = 4p_k g^{\otimes 2} (\rho_{1,k} \otimes \rho_{2,k}) (g^{\otimes 2})^\dagger \quad (20)$$

で与えられる。これが POVM であるには

$$I = \sum_k \int_{SU(2)} 4p_k g^{\otimes 2} (\rho_{1,k} \otimes \rho_{2,k}) (g^{\otimes 2})^\dagger \nu(dg) \quad (21)$$

である必要がある。もちろん、 k が連続である場合も考えられ、その場合は $\sum_k p_k$ を適当な確率測度に置き換えることになる。 $(\mathbb{C}^2)^{\otimes 2}$ での射影 P_s, P_a が不変であることから、条件 (21) は

$$\begin{aligned}\mathrm{Tr} P_s \sum_k \int_{\mathrm{SU}(2)} 4p_k g^{\otimes 2} (\rho_{1,k} \otimes \rho_{2,k}) (g^{\otimes 2})^\dagger \nu(dg) &= 3 \\ \mathrm{Tr} P_a \sum_k \int_{\mathrm{SU}(2)} 4p_k g^{\otimes 2} (\rho_{1,k} \otimes \rho_{2,k}) (g^{\otimes 2})^\dagger \nu(dg) &= 1\end{aligned}$$

と同値であり、これは

$$\mathrm{Tr} P_s \sum_k p_k (\rho_{1,k} \otimes \rho_{2,k}) = \frac{3}{4}, \quad \mathrm{Tr} P_a \sum_k p_k (\rho_{1,k} \otimes \rho_{2,k}) = \frac{1}{4}$$

と同値である。 $\rho_{1,k} \otimes \rho_{2,k}$ は純粋状態であるので、 $\beta_k := \mathrm{Tr} \rho_{1,k} \rho_{2,k}$ とすると、

$$\sum_k p_k \beta_k = \frac{1}{2} \quad (22)$$

となることと同値となる。そして、その Fisher 情報量は

$$\begin{aligned}J_{r,\theta;r}^M &= \sum_k 4p_k \int_{\mathrm{SU}(2)} \left(\frac{d \log (\mathrm{Tr} \rho_{r,\theta} g \rho_{1,k} g^\dagger \mathrm{Tr} \rho_{r,\theta} g \rho_{2,k} g^\dagger)}{dr} \right)^2 \left(\mathrm{Tr} \rho_{r,\theta} g \rho_{1,k} g^\dagger \mathrm{Tr} \rho_{r,\theta} g \rho_{2,k} g^\dagger \right) \nu(dg)\end{aligned}$$

となる。付録 F で示すように、

$$\begin{aligned}& \int_{\mathrm{SU}(2)} \left(\frac{d \log (\mathrm{Tr} \rho_{r,\theta} g \rho_{1,k} g^\dagger \mathrm{Tr} \rho_{r,\theta} g \rho_{2,k} g^\dagger)}{dr} \right)^2 \left(\mathrm{Tr} \rho_{r,\theta} g \rho_{1,k} g^\dagger \mathrm{Tr} \rho_{r,\theta} g \rho_{2,k} g^\dagger \right) \nu(dg) \\ &= \frac{2\beta_k - 1}{3} + \frac{\beta_k - 1}{r^2} + \frac{1 - \beta_k}{2r^3} \log \frac{1+r}{1-r}\end{aligned} \quad (23)$$

となるので、条件 (22) から

$$J_{r,\theta;r}^M = 2J_{r,\theta;r}^{M_{opt}} = -\frac{2}{r^2} + \frac{1}{r^3} \log \left(\frac{1+r}{1-r} \right)$$

を得る。つまり、2つの系の間で古典的な相関を用いても、個別的に同じ測定を行なった場合よりも、推定精度を改良することが無いことが分かる。このことは、推定のときに、量子的な相関を用いることの重要性を意味している。

7.5 量子効果が表れる推定の実現に向けて（その2）

現実の測定で量子相関を用いた推定効果を検証するには、不完全な測定について、きちんと議論する必要がある。先の定式化では n 個の粒子を独立に準備することを前提にしていたが、実際の粒子発生器から粒子が n 個の状態を正確に発生させることはきわめて困難であり、粒子の個数を制御できる変数とみることはできない。むしろ現実の系では、粒子を発生させる時間が制御可能な

変数だと見るべきである。この場合、個々の粒子の発生が独立であるとみなすと、一定時間に発生される粒子の個数 n はポアソン分布

$$e^{-t} \frac{t^n}{n!} \quad (24)$$

に従うと考えられる。発生時間を長くするに従い、それに比例して発生する粒子の個数の期待値 t が大きくなる。

さらに、測定に対応する POVM $M = \{M_i\}$ は必ずしも $\sum_i M_i = I$ を満たすだけでなく、むしろ一般には

$$\sum_i M_i \leq I \quad (25)$$

で表されると考える方が自然である。なぜなら、測定器でカウントできない場合があるからである。例えば密度作用素 $\rho \in \mathcal{S}(\mathcal{H})$ で表される粒子を n 個独立に同一の測定器 $M = \{M_i\}_{i=1}^l$ で測定すると、その測定値 i を k_i 個カウントする確率は 2 項分布 $\binom{n}{k_i} (\text{Tr } \rho M_i)^{k_i} (1 - \text{Tr } \rho M_i)^{n-k_i}$ で与えられ、 $(k_1, \dots, k_l)$ は多項分布

$$\frac{n!}{(n - \sum_i k_i)!} (1 - \text{Tr } \rho \sum_i M_i)^{n - \sum_i k_i} \prod_{i=1}^l \frac{(\text{Tr } \rho M_i)^{k_i}}{k_i!}$$

に従う。 (l) はカウントされる測定値の種類である。すなわち、 $\{M_i\}_{i=1}^l$ の要素の数に対応する。) 一般には総粒子数 n が不明であるので、(25) で考えることが本質的である。粒子数 n がポアソン分布 (24) に従うことを考慮に入れると $M = \{M_i\}$ で測定した場合、個々の測定値 i のカウント数 k_i はポアソン分布

$$\sum_{n \geq k_i} \left(e^{-t} \frac{t^n}{n!} \right) \left(\frac{n!}{(n - k_i)! k_i!} (1 - \text{Tr } \rho M_i)^{n - k_i} (\text{Tr } \rho M_i)^{k_i} \right) = e^{-t(\text{Tr } \rho M_i)} \frac{(t(\text{Tr } \rho M_i))^{k_i}}{k_i!} \quad (26)$$

に従い、 $(k_1, \dots, k_l)$ はポアソン分布

$$\begin{aligned} \sum_{n \geq \sum_i k_i} \left(e^{-t} \frac{t^n}{n!} \right) \left(\frac{n!}{(n - \sum_i k_i)!} (1 - \text{Tr } \rho \sum_i M_i)^{n - \sum_i k_i} \prod_{i=1}^l \frac{(\text{Tr } \rho M_i)^{k_i}}{k_i!} \right) \\ = \prod_{i=1}^l e^{-t(\text{Tr } \rho M_i)} \frac{(t(\text{Tr } \rho M_i))^{k_i}}{k_i!} \end{aligned} \quad (27)$$

に従う。従って、状態族 $\{\rho_\theta | \theta \in \Theta\}$ の未知パラメータ θ を測定 $M = \{M_i\}$ を用いての推定は、かく乱母数 t を持つ確率分布族

$$\left\{ \prod_{i=1}^l e^{-t(\text{Tr } \rho_\theta M_i)} \frac{(t(\text{Tr } \rho_\theta M_i))^{k_i}}{k_i!} \middle| \theta \in \Theta, t > 0 \right\} \quad (28)$$

での θ の推定とみなすことができる。一般に用いられる最尤法を用いると、推定値 $\hat{\theta}$ は $\hat{\theta} := \arg\max_{\theta \in \Theta} \sum_i k_i \log \text{Tr } \rho_\theta M_i$ となり、データの経験分布が $\{\frac{k_i}{\sum_j k_j}\}$ となるときの確率分布族 $\{p_\theta | \theta \in \Theta\}$ ($p_\theta(i) := \text{Tr } \rho_\theta M_i / (\sum_j \text{Tr } \rho_\theta M_j)$) での最尤法と一致する。

特に, $\sum_i M_i = I$ であれば, n 個のデータのときの MSE がほぼ, $\frac{1}{J_\theta^M n} \cong \frac{1}{J_\theta^M (n+1)}$ であるので, (24) で平均すると,

$$\sum_{n=0}^{\infty} e^{-t} \frac{t^n}{n!} \frac{1}{J_\theta^M n} \cong \sum_{n=1}^{\infty} e^{-t} \frac{t^n}{(n+1)!} \frac{1}{J_\theta^M} = \frac{1}{t J_\theta^M}$$

となる. また Fisher 情報量も $t J_\theta^M$ となることが確認できる. 例えば, 7.2 節で扱った量子相関を用いない群共変的測定の場合は, Fisher 情報量は $\frac{t}{2} \left(-\frac{2}{r} + \frac{1}{r^3} \log \left(\frac{1+r}{1-r} \right) \right)$ となり, 最尤法を用いたときの MSE はほぼその逆数になる. 従って, このような問題設定では, MSE の t 倍, または (Fisher 情報量)/ t に注目することになる. 前者は単位時間当りの誤差で, 後者は単位時間当りの情報量と見ることができる.

ここで, 量子 2 準位系の固有値推定で 7.3 節で議論した $(\mathbb{C}^2)^{\otimes 2}$ での測定 $\{P_a, P_s\}$ を, この議論に適用してみよう. この場合, 最も実験が容易なのは, 片方の空間のみ粒子が検出できる測定である. 例えば, 空間 $\mathcal{H}_a$ に来た粒子を確率 α_a で検出する測定器は $\{\alpha_a P_a\}$ で記述され, k 個の粒子を検出する確率は,

$$e^{-t\alpha_a \frac{1-r^2}{4}} \frac{\left(t\alpha_a \frac{1-r^2}{4} \right)^k}{k!}$$

で与えられる. 従って, 最尤推定量は式

$$k = t\alpha_a \frac{1 - \hat{r}_{ML}^2}{4}$$

を解くことで, $\hat{r}_{ML} = \sqrt{1 - \frac{4k}{t\alpha_a}}$ となる. もちろん, ここでは t, α_a は既知である必要がある. そのときの Fisher 情報量は $\frac{t\alpha_a r^2}{1-r^2}$ となり, MSE はほぼその逆数の $\frac{1-r^2}{t\alpha_a r^2}$ になる.

同様に, 空間 $\mathcal{H}_s$ に来た粒子を確率 α_s で検出する測定器であれば $\{\alpha_s P_s\}$ で記述され, k 個の粒子を検出する確率は,

$$e^{-t\alpha_s \frac{3+r^2}{4}} \frac{\left(t\alpha_s \frac{3+r^2}{4} \right)^k}{k!}$$

で与えられる. 従って, 最尤推定量は式

$$k = t\alpha_s \frac{3 + \hat{r}_{ML}^2}{4}$$

を解くことで, $\hat{r}_{ML} = \sqrt{3 + \frac{4k}{t\alpha_s}}$ となる. もちろん, ここでも t, α_s は既知である必要がある. そのときの Fisher 情報量は $\frac{t\alpha_s r^2}{3+r^2}$ となり, MSE はほぼその逆数の $\frac{3+r^2}{t\alpha_s r^2}$ になる. したがって, $\alpha_s = \alpha_a$ であれば, 測定 $\{\alpha_a P_a\}$ の方が優れていることになる. 一般に測定 $\{\alpha_a P_a, \alpha_s P_s\}$ を用いた場合は, その Fisher 情報量は $\frac{t\alpha_a r^2}{1-r^2} + \frac{t\alpha_s r^2}{3+r^2}$ となり, MSE はほぼ, その逆数になる.

例えば, 測定 $\{0.5 P_a\}$ を用いた場合 (以下 M'_1 と記す) と, 7.2 節で扱った量子相関を用いない共変的測定を完全に行った場合 (以下 M_0 と記す) を比較しよう. そのためには, 両者の Fisher 情報量 $\frac{0.5r^2}{1-r^2}$ と $2J_{r,\theta;r}^{M_{opt}} = \left(-\frac{2}{r^2} + \frac{1}{r^3} \log \left(\frac{1+r}{1-r} \right) \right)$ を比較することになる. これを図 7 に示した. こ

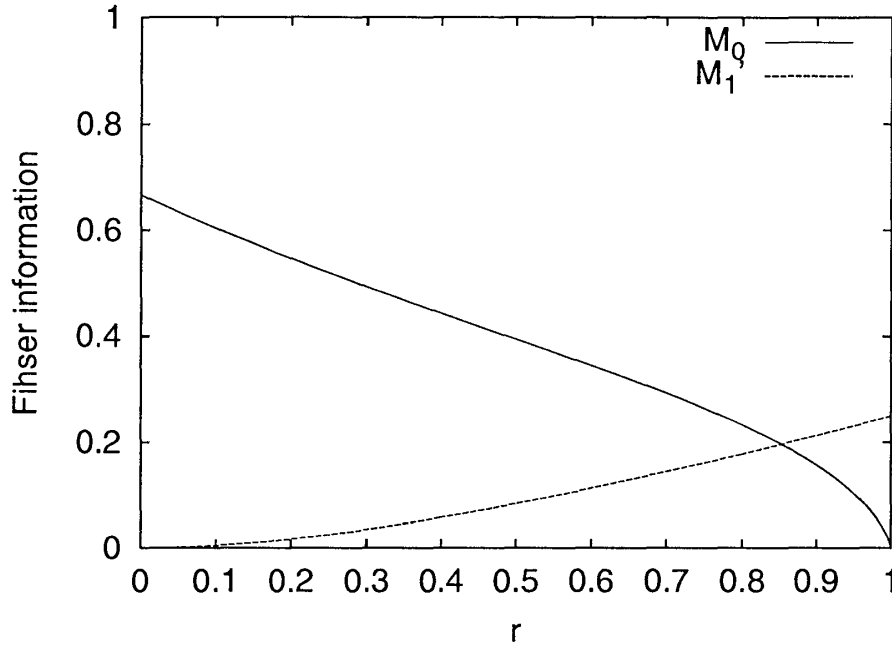


図 7: Fisher 情報量についての M_0 と M_1' 比較

でも図 6 と同様に, r が 1 に近い場合は無限大になるので, これらに $1-r$ を掛けたものをグラフに表した. この場合, $r \geq 0.85$ のとき, M_1' の方が優れている.

量子効率 $t = 0.5$ という数字は決して不可能な数字ではないことを考えると, この実験は十分可能であると考えられる. 事実, 既に Hendrych, Dusek, Filip, Fiurasek [22] によって測定 $\{\alpha P_a, \alpha P_s\}$ を用いて固有値を推定する実験はなされている. しかし, 彼らの実験では, このような測定でも固有値が推定可能であるということに主眼が置かれており, どの程度の精度が保証される推定であるかは十分には議論されていない. もちろん, 本稿で述べたような量子相関を用いることのメリットも全く議論されていない. 今後, この測定を実現することで, 量子相関を用いることにより, 用いない場合よりも, 推定精度が向上することが実証されることが期待される.

8 結論

量子相関を用いた場合と量子相関を用いない場合の推定限界を緻密に評価し, 量子効果が表れる推定が現実の実験で実現可能であることを述べてきた. 本稿を読んだ読者のうち, 一人でも多くの人が, このような実験に興味を持ち, 推定においても測定の量子効果が重要であることが実証されることを期待する.

謝辞

総合研究大学院大学の井元信之先生には文献 [12] を, ERATO 今井量子計算機構プロジェクト/NEC の富田彰久博士には文献 [13] を, 東京工業大学総合理工学研究科の宇佐見康二氏には文

献 [22] を、京都大学理学研究科の中尾裕也博士には文献 [10] をそれぞれご教示頂きました。さらに、富田彰久博士、中尾裕也博士には原稿を丁寧にチェックして頂き、大変有益なコメントを頂きました。

Appendix

A Jensen の不等式

正の実数から正の実数への関数 f が任意の $1 > \lambda > 0$ となる λ と $x_1 \neq x_2 > 0$ に対して、

$$\lambda f(x_1) + (1 - \lambda)f(x_2) > f(\lambda x_1 + (1 - \lambda)x_2)$$

となるとき、 f は凸であるといわれる。確率空間 Ω 上の正の実数に値を持つ確率変数 X と Ω 上の確率分布 p に対して、Jensen の不等式:

$$\int_{\Omega} f(X(\omega))p(d\omega) \geq f\left(\int_{\Omega} X(\omega)p(d\omega)\right) \quad (29)$$

が成り立つ。特に、等号が成立するのは X が確率 1 である一定値を取る場合に限られる。

この不等式を用いると以下の事実を示すことができる。正定値なエルミート行列（固有値が全て正の実数の行列） A とノルムが 1 のベクトル v に対して、

$$\langle v|A|v \rangle^{-1} \leq \langle v|A^{-1}|v \rangle \quad (30)$$

が成り立ち、かつ、等号が成立するのは v が A の固有ベクトルである場合に限られる。

A をスペクトル分解 $A = \sum_i a_i E_i$ とすると (a_i は固有値で E_i はそれに対応する射影), $\langle v|E_i|v \rangle$ は確率分布になる。関数 $x \mapsto 1/x$ は凸なので、Jensen の不等式が使える、

$$\langle v|A|v \rangle^{-1} = \left(\sum_i a_i \langle v|E_i|v \rangle \right)^{-1} \leq \sum_i a_i^{-1} \langle v|E_i|v \rangle = \langle v|A^{-1}|v \rangle$$

となる。また、上式の等号成立条件も Jensen の等号成立条件から容易に得られる。

B 確率分布族の推定における漸近論

6 節では、推定したい未知パラメータが 1 つの場合のみ扱ったが、非漸近的な議論が可能な指数型分布族についてまとめ、その後に、パラメータが複数の場合、及び、推定したいパラメータに加えて、推定する必要のない未知パラメータがある場合の議論を簡単にまとめる。その後に、推定に関する 2 次の漸近理論をまとめる。

B.1 未知のパラメータが 1 つの場合

パラメータが 1 次元の確率分布族 $\{p_{\theta}|\theta \in \Theta \subset \mathbb{R}\}$ の推定量 $\hat{\theta}_n$ の推定精度を議論する場合、平均 2 乗誤差 (MSE)

$$v_{\theta}(\hat{\theta}_n) := \int_{\Omega^n} (\hat{\theta}_n(\vec{\omega}_n) - \theta)^2 p_{\theta}^n(d\vec{\omega}_n)$$

を議論することが多い。このとき、推定量にはしばしば不偏性条件:

$$\theta = \int_{\Omega^n} \hat{\theta}_n(\vec{\omega}_n) p_{\theta}^n(d\vec{\omega}_n)$$

が課される。推定量がこの不偏性条件を満たすとき、 $\text{MSE } v_{\theta}(\hat{\theta}_n)$ について以下の Cramér-Rao 不等式:

$$v_{\theta}(\hat{\theta}_n) \geq \frac{J_{\theta}^{-1}}{n} \quad (31)$$

が成り立つ。特に全ての θ で (31) で等号を達成する不偏推定量は有効推定量とよばれる。有効推定量が存在するか否かは、確率分布のなす集合 $\{p_{\theta}\}_{\theta \in \Theta}$ に依存するだけでなく、そのパラメータ θ の取り方にも依存する。後に述べるようにパラメータ θ の下で有効推定量が存在しても、 θ からある非線型な変換を施して得られる異なったパラメータ θ' の下では有効推定量は存在するとは限らない。

以下、有効推定量が存在するための必要十分条件を述べるために、指数型分布族について簡単にまとめる。確率変数 X を用いて定義される確率分布

$$p_{\theta}(\omega) = p_0(\omega) \exp(\theta X(\omega) - \psi(\theta)) \quad (32)$$

$[\psi(\theta) := \log \int_{\Omega} \exp(\theta X(\omega)) p_0(d\omega)]$ からなる確率分布族 $\{p_{\theta} | \theta \in \mathbb{R}\}$ は指数型分布族とよばれ、精緻な統計的推測の理論が整備されている。上記の指数型分布族には (32) で与えられるパラメータ (自然パラメータ) θ の他に、

$$\eta := \int_{\Omega} X(\omega) p_{\theta}(d\omega)$$

で与えられるパラメータ (期待値パラメータ) がある。 $X_i(\vec{\omega}_n) := X(\omega_i)$ としたとき、 $X^n(\vec{\omega}_n) := \sum_{i=1}^n \frac{1}{n} X_i$ は期待値パラメータ η の不偏推定量で、その分散は期待値パラメータ η の下での Fisher 情報量 J_{η} を用いて $\frac{J_{\eta}^{-1}}{n}$ で与えられる。すなわち、推定量 $X^n(\vec{\omega}_n)$ は指数型分布族の期待値パラメータ η に関する有効推定量である。この推定量は最尤推定量でもある。逆に、有効推定量が存在する場合は、指数型分布族の期待値パラメータの場合に限られる。

しかし、指数型分布族でない場合でも、不等式 (31) に相当する不等式は成り立つ。推定量の無限列 $\{\hat{\theta}_n\}$ が真のパラメーターに確率収束することに加え、この収束に関するある種の一意性を課すと、(31) に相当する不等式

$$\lim_{n \rightarrow \infty} n v_{\theta}(\hat{\theta}_n) \geq J_{\theta}^{-1} \quad (33)$$

が成り立つ。特に $\hat{\theta}_n$ が最尤推定量 $\hat{\theta}_{ML}(\vec{\omega}_n) := \arg\max_{\theta} \prod_{i=1}^n p_{\theta}(\omega_i)$ であるときに、(33) で等号が成り立つ。確率分布には以下に定義する Hellinger 距離

$$H(p, q) := \int_{\Omega} \left(\sqrt{p(\omega)} - \sqrt{q(\omega)} \right)^2 d\omega$$

が知られているが、以下のようにこの距離の極限として、Fisher 情報量 J_{θ} は特徴付けられる。

$$\frac{J_{\theta}}{4} = \lim_{\epsilon \rightarrow 0} \frac{H(p_{\theta}, p_{\theta+\epsilon})}{\epsilon^2}.$$

B.2 未知のパラメータが複数ある場合

次に、多次元のパラメータを持つ確率分布族 $\{p_\theta(\omega) | \theta \in \Theta \subset \mathbb{R}^d\}$ に対するパラメータ θ の推定を考える。この場合は、推定量は Ω^n (n は得られたデータの数) から Θ または $\mathbb{R}^d$ への関数 $\hat{\theta}_n$ として与えられる。その精度を各パラメータの MSE の和 $\sum_i \int_{\Omega^n} (\hat{\theta}_n^i(\vec{\omega}_n) - \theta^i)^2 p_\theta^n(d\vec{\omega}_n)$ ($p_\theta^n(d\vec{\omega}_n) := \prod_{i=1}^n p_\theta(\omega_i)$) を用いて評価すること多いが、MSE 行列とよばれる $d \times d$ の対称行列 $v_\theta^{i,j}(\hat{\theta}_n) := \int_{\Omega^n} (\hat{\theta}_n^i(\vec{\omega}_n) - \theta^i)(\hat{\theta}_n^j(\vec{\omega}_n) - \theta^j) p_\theta^n(d\vec{\omega}_n)$ に注目することもある。対称行列の集合は全順序ではないため、完全に優越がつかないことも起りかねないが、確率分布族のパラメータ推定についての 1 次漸近論ではそのようなことは起きず、最適値な MSE 行列が一意に決まる。(量子状態族の場合には 1 次漸近論であっても、最適な MSE 行列は一意には決まらない。) パラメータが 1 つの場合と同様に、最尤推定量 $\hat{\theta}_{ML}$ は $\hat{\theta}_{ML}(\vec{\omega}_n) := \arg\max_\theta \prod_{i=1}^n p_\theta(\omega_i)$ で与えられ、以下の式が漸近的に成り立つ。

$$v_\theta^{i,j}(\hat{\theta}_n) \cong (J_\theta^{-1})^{i,j} \frac{1}{n}, \quad (J_\theta)_{i,j} := \int \left(\frac{\partial}{\partial \theta^i} \log p_\theta(\omega) \right) \left(\frac{\partial}{\partial \theta^j} \log p_\theta(\omega) \right) p_\theta(\omega) d\omega.$$

ここで行列 $(J_\theta)_{i,j}$ は Fisher 情報行列とよばれ、1 パラメータの場合の Fisher 情報量の代わりの役割を担う。さらに、推定量の列 $\vec{\hat{\theta}} = \{\hat{\theta}_n\}$ が収束に関するある種の一様性を満たすとする、極限で与えられる行列 $v^{i,j}(\vec{\hat{\theta}}) := \lim_{n \rightarrow \infty} n v_\theta^{i,j}(\hat{\theta}_n)$ は

$$v^{i,j}(\vec{\hat{\theta}}) \geq (J_\theta^{-1})^{i,j}$$

を満たす。

次に推定したいパラメータ θ の他に有限個の未知パラメータ $\phi \in \mathbb{R}^d$ がある場合について考える。以下では簡単のため推定したいパラメータの次元は 1 つとする。この場合、確率分布族 $\{p_{\theta,\phi}(\omega) | \theta \in \Theta \subset \mathbb{R}, \phi \in \mathbb{R}^d\}$ に対するパラメータ θ の推定を考えることになる。 θ についての Fisher 情報量 $J_{\theta,\phi;\theta}$ を偏微分を用いて

$$J_{\theta,\phi;\theta} := \int \left(\frac{\partial}{\partial \theta} \log p_{\theta,\phi}(\omega) \right)^2 p_{\theta,\phi}(\omega) d\omega$$

で定義すると、任意の推定量 $\hat{\theta}_n$ は ϕ が既知の場合の ϕ に依存したパラメータが 1 つの確率分布族 $\{p_{\theta,\phi}(\omega) | \theta \in \Theta \subset \mathbb{R}\}$ の推定量にもなるので、

$$\lim_{n \rightarrow \infty} n v_{\theta,\phi}(\hat{\theta}_n) \geq J_{\theta,\phi;\theta}^{-1}$$

が成り立つ。一方、確率分布族 $\{p_{\theta,\phi}(\omega) | \theta \in \Theta \subset \mathbb{R}, \phi \in \mathbb{R}^d\}$ に対する最尤推定量 $(\hat{\theta}_{ML,n}, \hat{\phi}_{ML,n})$ を考え、それから $\hat{\phi}_{ML,n}$ を忘れ、 $\hat{\theta}_{ML,n}$ のみを θ の推定量とした場合、その MSE は

$$v(\hat{\theta}_{ML,n}) \cong (J_{\theta,\phi}^{-1})^{0,0}$$

で与えられる。ただし $J_{\theta,\phi}$ は θ を 0 番目の座標とする $(d+1) \times (d+1)$ の対称行列で確率分布族 $\{p_{\theta,\phi}(\omega) | \theta \in \Theta \subset \mathbb{R}, \phi \in \mathbb{R}^d\}$ の Fisher 情報行列である。もちろん、 $J_{\theta,\phi;\theta} = (J_{\theta,\phi})_{0,0}$ であるが、一般には不等式

$$J_{\theta,\phi;\theta}^{-1} \leq (J_{\theta,\phi}^{-1})^{0,0} \quad (34)$$

が成り立ち、等号が成立するのはベクトル $(1, 0, \dots, 0)^T$ が行列 $J_{\theta, \phi}$ の固有ベクトルのときのみである。この事実は (29) で正定値なエルミート行列 A に対して示した事実から直ちに確認できる。

(34) で等号が成立しない場合、すなわち、 $(1, 0, \dots, 0)^T$ が行列 $J_{\theta, \phi}$ の固有ベクトルのときのみでない場合は、次に与える θ に対する推定量 $\hat{\theta}_{a,n}$ が限界 $J_{\theta, \phi}^{-1}$ を達成する。確率分布族 $\{p_{\theta, \phi}(\omega) | \theta \in \Theta \subset \mathbb{R}, \phi \in \mathbb{R}^d\}$ に対する最尤推定量 $(\hat{\theta}_{ML,n}, \hat{\phi}_{ML,n})$ を考え、その推定値 $\hat{\phi}_{ML,n}$ を用いて、確率分布族 $\{p_{\theta, \hat{\phi}_{ML,n}}(\omega) | \theta \in \Theta \subset \mathbb{R}\}$ に対する最尤推定量で $\hat{\theta}_{a,n}$ を定義する。すると、推定量 $\hat{\theta}_{a,n}$ についての極限 $\lim_{n \rightarrow \infty} n v(\hat{\theta}_{a,n})$ は $J_{\theta, \phi}^{-1}$ に一致する。

B.3 2 次の漸近論

次に MSE の 2 次のオーダーについて考える。簡単のため、1 次元のパラメータを持つ確率分布族 $\{p_{\theta}(\omega) | \theta \in \Theta \subset \mathbb{R}\}$ に対するパラメータ θ の推定のみをここでは扱う。2 次の漸近論では推定量の列 $\hat{\theta}_n$ で、その MSE $v_{\theta}(\hat{\theta}_n)$ の漸近的挙動が $v_{\theta}(\hat{\theta}_n) \cong J_{\theta}^{-1} \frac{1}{n} + C \frac{1}{n^2}$ となるものに限って議論し、係数 C に注目する。これを扱う一般論としては、バイアス $b_{\theta}(\hat{\theta}) := \int_{\Omega} (\hat{\theta}_n(\bar{\omega} - \theta)) p_{\theta}^n(d\bar{\omega}_n)$ の $\frac{1}{n}$ のオーダーの係数が一様に 0 になるものに限って、議論する場合が多い。一般にはバイアス $b_{\theta}(\hat{\theta}_n)$ の $\frac{1}{n}$ のオーダーの係数は必ずしも 0 にならないが、そのような推定量に対して若干の操作を施すと、 $\frac{1}{n}$ のオーダーのバイアスを 0 にすることができる (バイアス補正)。バイアス補正後の推定量については、常に C は正の値になり、最尤推定量にバイアス補正したものが最適となる。しかしながら、バイアス補正しない推定量まで含めた場合での最適性は十分には議論されていない。もちろん、指数型分布族の期待値パラメータに注目した場合は最尤推定量のバイアスは消え、MSE の 2 次のオーダーの係数も消える。しかし、指数型分布族であっても、全く異なるパラメータを選択した場合、最尤推定量であってもバイアスがあり、取り扱いがややこしい場合がある。

二項分布は $1/(1 + \exp(\theta))$, $\exp(\theta)/(1 + \exp(\theta))$ とパラメータを取ると、指数型分布族であることがわかる。ここで $1 - \eta$, η とパラメータ η を取ると、これは期待値パラメータになり、先に述べた一般論が適用可能となる。しかし、 $1/(1 + \theta)$, $\theta/(1 + \theta)$ とパラメータを取ると、これは期待値パラメータでもなく、自然パラメータでも無いので、その扱いは困難である。

n 回の独立試行で $1/(1 + \theta)$ に対応する事象が k 回得られた場合、最尤推定量 $\hat{\theta}_{ML,n}$ は $1/(1 + \hat{\theta}_{ML,n}) = k/n$ より、 $\hat{\theta}_{ML,n} := (n - k)/k$ で与えられる。一方、最尤推定量を変形して推定量 $\hat{\theta}'_n$ を $\hat{\theta}'_n := (n - k)/(k + 1)$ を考えることもできる。そのとき、それぞれのバイアス $b_{\theta}(\hat{\theta}_{ML,n})$, $b_{\theta}(\hat{\theta}'_n)$ は

$$b_{\theta}(\hat{\theta}_{ML,n}) \cong \frac{\theta(1 + \theta)}{n}, \quad b_{\theta}(\hat{\theta}'_n) = -\frac{\theta^{n+1}}{(1 + \theta)^n}$$

となり、 $\hat{\theta}'_n$ のバイアスは指数的に 0 に近づく。これらの MSE の 2 次のオーダーの漸近的な振る舞いは、それぞれ、

$$v_{\theta}(\hat{\theta}_{ML,n}) \cong \frac{\theta(1 + \theta)^2}{n} + \frac{\theta(7\theta + 2)(1 + \theta)^2}{n^2}, \quad v_{\theta}(\hat{\theta}'_n) \cong \frac{\theta(1 + \theta)^2}{n} + \frac{2\theta^2(1 + \theta)^2}{n^2}$$

となり、後者の方が θ によらず小さくなる。もちろん、この事実は最尤推定量にバイアス補正を施したものが、2 次のオーダーで最適であるという一般論と矛盾しない。なぜなら、推定量 $\hat{\theta}'_n$ は

最尤推定量 $\hat{\theta}_{ML,n}$ にバイアス補正を加えたものと見なせるからである。ここで注意しておかないといけないことは、バイアスや平均 2 乗誤差などの概念は、パラメータの取り方に依存するということである。この例では期待値パラメータの下でバイアスが無い最尤推定量であっても、パラメータを変えるとバイアスが生じるということである。また同時に、推定量の優劣もパラメータの取り方に依存してしまうことにも注意する必要がある。

C 群対称性と群不変なパラメータについての Fisher 情報量の最大化

ここで、量子相関を用いる場合の最適化について議論する前に、状態族が群対称性を満たし、群不変なパラメータのみを推定する場合の議論を要約するとともに、先の節での証明の細部を一般的設定の下で与える。

C.1 Fisher 情報量と最尤推定量

今、2つのパラメータ $r \in \mathbb{R}$ と $\theta \in \Theta$ を考え、コンパクトなパラメータ空間 Θ にはコンパクトな群 G が推移的に作用しているとする。すなわち、群 G の Θ への作用を f としたとき、任意の $\theta_1, \theta_2 \in \Theta$ に対して、 $\theta_2 = f(g)\theta_1$ となる $g \in G$ が存在する。さらに、群不変な距離 d 、すなわち、任意の $\theta_1, \theta_2 \in \Theta$ と $g \in G$ に対して、 $d(\theta_1, \theta_2) = d(f(g)\theta_1, f(g)\theta_2)$ となる距離が存在し、加えて、 $\theta_1, \theta_2, \theta_3 \in \Theta$ が $d(\theta_1, \theta_2) = d(\theta_1, \theta_3)$ となるとき、 $f(g)\theta_1 = \theta_1, f(g)\theta_2 = \theta_3$ となる $g \in G$ が取れるとする。この作用の下で不変な Θ 上の確率測度を $\mu(d\theta)$ で表し、群 G の不変測度を $\nu(dg)$ で表すことにする。それぞれコンパクトであるので右不変な確率測度と左不変な確率測度は一致する。

以下、群 G の Hilbert 空間 $\mathcal{H}$ への表現 V が与えられており、 $\mathcal{H}$ 上の状態族 $\{\rho_{r,\theta} | r \in \mathbb{R}, \theta \in \Theta\}$ が以下の条件を満たす場合を考える。

$$V_g \rho_{r,\theta} V_g^\dagger = \rho_{r,f(g)\theta}, \quad \forall g \in G.$$

Ω に測定値を持つ測定 M に対して、測定 M_g を

$$M_g(d\omega) := V_g^\dagger M(d\omega) V_g$$

で定義する。特に、空間 Ω に群 G が推移的に作用し、その作用が f_Ω で書ける場合、 Ω に測定値を持つ測定 M が以下の条件を満たすとき、共変的とよぶことにする。

$$M_g(d\omega) = M(f_\Omega(g) d\omega)$$

Ω は必ずしも Θ とは一致しない。

以下、群の作用と Fisher 情報量との関係について簡単にまとめる。 $J_{r,\theta;r}^M$ は r についての Fisher 情報量とする。付録 B.1 の議論から、測定 M を用いて推定する場合、収束に関する一様性を満たす推定量 $\hat{r}_n$ を用いると仮定すると、その MSE は

$$\lim_{n \rightarrow \infty} n v_{r,\theta}(M \times n, \hat{r}_n) \geq (J_{r,\theta;r}^M)^{-1} \quad (35)$$

となる.

補題 1 測定 M が共変的であるとき, (r, θ) の最尤推定量を $(\hat{r}_{ML,n}, \hat{\theta}_{ML,n})$ としたとき, r の推定量である $\hat{r}_{ML,n}$ は (35) で等号が成立する.

証明 以下, (r_0, θ_0) での Fisher 情報行列について議論するので, θ_0 の自然な座標を用いて議論することとする. すなわち, その座標の下では (r_0, θ_0) の近傍では

$$d(\theta_0, \theta + \eta) = d(\theta_0, \theta - \eta), \quad \eta \in \mathbb{R}^d$$

とする. (d は θ の次元とする.) ここで, $J_{r,\theta;r}^M$ は r についての Fisher 情報量なので θ の座標の取り方に依存しないことに注意せよ. このとき付録 B.2 の議論より, 上記の方法で構成した推定量 $\hat{r}_{ML,n}$ が (35) で等号が成立するための必要十分条件は, Hellinger 距離 $H(\cdot, \cdot)$ を用いると,

$$\lim_{\epsilon \rightarrow 0} \frac{H(P_{r_0+\epsilon, \theta_0+\epsilon\eta}^M, P_{r_0, \theta_0}^M)}{\epsilon^2} = \lim_{\epsilon \rightarrow 0} \frac{H(P_{r_0-\epsilon, \theta_0+\epsilon\eta}^M, P_{r_0, \theta_0}^M)}{\epsilon^2} \quad (36)$$

で表すことができる.

ところで, $\text{Tr } \rho_{r, f(g)\theta} M(d\omega) = \text{Tr } \rho_{r, \theta} V_g^\dagger M(d\omega) V_g = \text{Tr } \rho_{r, \theta} M_g(d\omega)$ となるので, $\omega_0 := f_\Omega(g)\omega$ とすると,

$$P_{r, \theta}^M(d\omega) = P_{r, f(g)\theta}^M(d\omega_0) \quad (37)$$

を得る.

M が共変的であるとき, $f(g)\theta_1 = \theta_1, f(g)\theta_2 = \theta_3, g \in G$ とすると,

$$H(P_{r_1, \theta_1}^M, P_{r_2, \theta_2}^M) = H(P_{r_1, f(g)\theta_1}^M, P_{r_2, f(g)\theta_2}^M) = H(P_{r_1, \theta_1}^M, P_{r_2, \theta_3}^M)$$

となる. したがって, (36) が得られ, 補題を得る. ■

ここで, 以下の議論に必要な性質をまとめておく, 測定 M_g を不変測度 $\nu(dg)$ で disjoint に重ね合わせた測定 $\overline{M}$ を考えると $\int_G J_{r, \theta; r}^{M_g} \nu(dg) = J_{r, \theta; r}^{\overline{M}}$ が成り立ち, 測定 $\overline{M}$ は共変的になる. なお測定 M の測定値の集合を Ω としたとき, 測定 $\overline{M}$ の測定値の集合は $G \times \Omega$ となる. また, (37) より,

$$J_{r, \theta; r}^M = J_{r, f(g)\theta; r}^M \quad (38)$$

が成り立つ.

C.2 ミニマックス法と Bayes 法

次に r, θ が未知で r のみを推定する問題を考える. この場合ミニマックス法では $\lim_{n \rightarrow \infty} n \max_{\theta \in \Theta} v_{r, \theta}(M \times n, \hat{r}_n)$ を, Bayes 法では $\lim_{n \rightarrow \infty} n \int_\Theta v_{r, \theta}(M \times n, \hat{r}_n) \mu(d\theta)$ をそれぞれ最小化することになる.

そして、収束に関する一様性を満たす推定量 $\hat{r}_n$ を用いると仮定すると、その MSE は

$$\lim_{n \rightarrow \infty} n v_{r,\theta}(M \times n, \hat{r}_n) \geq (J_{r,\theta;r}^M)^{-1} \quad (39)$$

となり、 $\hat{r}_n$ が最尤推定量のとき等号が成立する ($J_{r,\theta;r}^M$ は r についての Fisher 情報量). それゆえ、ミニマックス法では $\min_{\theta} J_{r,\theta;r}^M$ を最大にする M を見つけることが問題となる. このとき

$$\min_{\theta} J_{r,\theta;r}^M \leq \int_{\Theta} J_{r,\theta;r}^M \mu(d\theta) = \int_G J_{r,f(g)\theta;r}^M \nu(dg) = \int_{\Theta} J_{r,\theta;r}^M \mu(d\theta) = J_{r,\theta;r}^{\overline{M}}$$

となり、 $\overline{M}$ は共変的であるので、

$$\max_M \min_{\theta} J_{r,\theta;r}^M \leq \max_{M:\text{cov}} J_{r,\theta;r}^M \quad (40)$$

となる. (ここで cov は群共変的測定を表し、 $\max_{M:\text{cov}}$ は群共変的測定について最大化をとることを意味する.) 一方、群 G のパラメータ空間 Θ への作用の推移性と (38) より、

$$\min_{\theta} J_{r,\theta;r}^M = \int_{\Theta} J_{r,\theta_0;r}^M \mu(d\theta) = J_{r,\theta_0;r}^M$$

となり、(40) の逆向きの不等式が成立し、

$$\max_M \min_{\theta} J_{r,\theta;r}^M = \max_M \int_{\Theta} J_{r,\theta;r}^M \mu(d\theta) = \max_{M:\text{cov}} J_{r,\theta;r}^M \quad (41)$$

が成り立つ. すなわち、

$$\lim_{n \rightarrow \infty} n \max_{\theta \in \Theta} v_{r,\theta}(M \times n, \hat{r}_n) = \left(\max_{M:\text{cov}} J_{r,\theta;r}^M \right)^{-1}$$

が成り立つ.

一方 Bayes 法では $\int_{\Theta} (J_{r,\theta;r}^M)^{-1} \mu(d\theta)$ を最小化することになる. 関数 $x \mapsto 1/x$ が下に凸であることに注目し Jensen の不等式 (29) を用いると、

$$\int (J_{r,\theta}^M)^{-1} \mu(d\theta) \geq \left(\int J_{r,\theta}^M \mu(d\theta) \right)^{-1} = (J_{r,\theta;r}^{\overline{M}})^{-1}$$

が成り立つ. したがって、 $\overline{M}$ が共変的であることから、

$$\min_M \int (J_{r,\theta}^M)^{-1} \mu(d\theta) \geq \min_{M:\text{cov}} (J_{r,\theta;r}^M)^{-1} \quad (42)$$

が成り立つ. 逆に、 M が共変的であるとすると、 $J_{r,\theta}^M$ は θ に依存しないので、(42) の逆向きの不等式が示せ、

$$\lim_{n \rightarrow \infty} n \int_{\Theta} v_{r,\theta}(M \times n, \hat{r}_n) \mu(d\theta) = \left(\max_{M:\text{cov}} J_{r,\theta;r}^M \right)^{-1}$$

が成り立つ. これらの議論から、どちらの問題設定を採用しても、群共変的な測定についての Fisher 情報量の最大化問題 $\max_{M:\text{cov}} J_{r,\theta;r}^M$ に帰着される.

C.3 既約性と既約分解を用いた Fisher 情報量の最大化

次に、群 G の $\mathcal{H}$ への表現 V_g が既約である場合について考える。 $\mathcal{H}$ 上の状態 ρ を用いて測定値集合を G に持つ共変的な測定 M^ρ を以下で構成する。

$$M^\rho(dg) := (\dim \mathcal{H}) V_g \rho V_g^\dagger \nu(dg)$$

このとき、

$$\begin{aligned} V_{g_0}(\dim \mathcal{H}) \int_G V_g \rho V_g^\dagger \nu(dg) &= (\dim \mathcal{H}) \int_G V_{g_0 g} \rho V_g^\dagger \nu(dg) = (\dim \mathcal{H}) \int_G V_g \rho V_{g_0^{-1}g}^\dagger \nu(dg) \\ &= (\dim \mathcal{H}) \int_G V_g \rho V_g^\dagger \nu(dg) V_{g_0} \end{aligned}$$

となるので、表現の既約性から $(\dim \mathcal{H}) \int_G V_g \rho V_g^\dagger \nu(dg)$ は定数になる。さらにそのトレースは $\text{Tr}(\dim \mathcal{H}) \int_G V_g \rho V_g^\dagger \nu(dg) = (\dim \mathcal{H}) \int_G \nu(dg) = \dim \mathcal{H}$ となることから、 $V_{g_0}(\dim \mathcal{H}) \int_G V_g \rho V_g^\dagger \nu(dg)$ は恒等作用素になり、 M^ρ は POVM となる。 ρ が純粋状態でなく $\rho = \sum_i \lambda_i |\phi_i\rangle\langle\phi_i|$ とスペクトル分解できる場合は $M^\rho = \sum_i \lambda_i M^{|\phi_i\rangle\langle\phi_i|}$ となり、

$$J_{r,\theta}^{M^\rho} \leq \sum_i \lambda_i J_{r,\theta}^{M^{|\phi_i\rangle\langle\phi_i|}}$$

となる。なお ρ を動かさない $g \in G$ の集合は群になり（不変部分群 H ）その場合の M^ρ の測定値集合は等質空間 G/H になる。

逆に M を測定値を集合 Ω にとる共変的な測定とする。さらに Ω 上の確率測度 μ_Ω を

$$\mu_\Omega(d\omega) := \frac{1}{\dim \mathcal{H}} \text{Tr } M(d\omega)$$

で定義し、 $\rho(\omega)$ を

$$(\dim \mathcal{H}) \rho(\omega) \mu_\Omega(d\omega) = M(d\omega)$$

で定義する。さらに Ω を群 G の作用で割った空間を X とし、 μ_Ω から自然に定義される X 上の確率測度を μ_X とする。このとき、 $x = [\omega] \in X$ に対して、 POVM $M^x := M^{\rho(\omega)}$ を定義すると代表元 ω の取り方に依存せずに M^x が決まる。したがって、

$$M = \int_X M^x \mu_X(dx)$$

が成立し、

$$J_{r,\theta}^{M^\rho} \leq \int_X J_{r,\theta}^{M^x} \mu_X(dx)$$

が成り立つ。特に、複数の共変的 POVM M_i があり、その disjoint な和で表される POVM $M := \coprod_i \lambda_i M_i$ については

$$J_{r,\theta}^{M^\rho} = \sum_i \lambda_i J_{r,\theta}^{M_i}$$

が成り立つ。したがって、

$$\max_{M:\text{cov}} J_{r,\theta;r}^M = \max_{\rho:\text{pure}} J_{r,\theta;r}^{M^\rho} \quad (43)$$

が成り立つ。それゆえ、表現が $\mathcal{H}$ 上の純粋状態の集合に推移的に作用するときには、 ρ が純粋状態である限り、 $J_{r,\theta;r}^{M^\rho}$ は ρ の取り方に依存しない。もちろん r にも依存しない。例えば、 $\text{SU}(2)$ が $\mathbb{C}^2$ に作用する場合が挙げられ、この場合にこの議論を適用すると (17) が得られる。

次に、群 G の $\mathcal{H}$ への表現 V_g が互いに非同値な有限個の既約表現の直和 $\mathcal{H}_1 \oplus \cdots \oplus \mathcal{H}_l$ に分解できるとし、各直和成分 $\mathcal{H}_i$ への射影を P_i とする。さらに、各状態 $\rho_{r,\theta}$ は P_i と可換とする。このとき、 $\{P_i\}$ は POVM になる。ここで測定値を集合 Ω_i に取る $\mathcal{H}_i$ 上の測定 M_i を考え、これらの直和 $\oplus_i M_i$ を $\mathcal{H}$ 上の測定値を集合 $\coprod_i \Omega_i$ に取る測定として以下で定義する。各 Ω_i に対してその Borel 集合 B_i を選ぶと、 $\coprod B_i$ は $\coprod_i \Omega_i$ の Borel 集合になり、直和 $\oplus_i M_i$ を

$$\left(\bigoplus_i M_i \right) \left(\coprod_j B_j \right) := \bigoplus_i (M_i(B_i))$$

を満たすものとして定義する。このとき、以下に示すように

$$\max_{M:\text{cov}} J_{r,\theta;r}^M = \max_{M_i:\text{cov on } \mathcal{H}_i} J_{r,\theta;r}^{\oplus_i M_i} = \max_{\rho_i:\text{pure on } \mathcal{H}_i} J_{r,\theta;r}^{\oplus_i M^{\rho_i}} \quad (44)$$

が成り立つ。これにより、群共変的な測定で Fisher 情報量を最大化する問題は、個々の既約成分で独立に純粋状態 ρ_i を変数とする最大化問題を解くことで十分であることが分かった。このことは、問題の取り扱いを著しく容易にする。

M を $\mathcal{H}$ 上の共変的な POVM とすると、 P_i と $\rho_{r,\theta}$ は可換であり、 $\sum_i P_i = I$ となることから

$$\text{Tr } M(B) \rho_{r,\theta} = \sum_i \text{Tr } P_i M(B) P_i \rho_{r,\theta}$$

となる。 M_i を $\mathcal{H}_i$ 上の POVM として、 $M_i(B) := P_i M(B) P_i$ と定義すると、Fisher 情報量の単調性 (14) より、

$$J_{r,\theta;r}^M \leq J_{r,\theta;r}^{\oplus_i M_i}$$

となる。最後に (43) の導出と同様の議論を用いることにより (44) を得る。

D 光子系における重要なユニタリな時間発展

5.3 節では時間発展 (9) が重要なポイントを占めていた。この時間発展について考えるために、はじめに量子系 $\mathcal{H} = L^2(\mathbb{R})$ が 2 つ与えられた場合に、相互作用させた時間発展について考える。すなわち、1 番目の系 $\mathcal{H}_1 = L^2(\mathbb{R})$ の消滅演算子を a_1 とし、2 番目の系 $\mathcal{H}_2 = L^2(\mathbb{R})$ の消滅演算子を a_2 とし、合成系のハミルトニアン $H := i(a_2^\dagger a_1 - a_1^\dagger a_2)$ の下で時間発展を考える。すると、初期状態がコヒーレント状態のテンソル積 $|\alpha_1\rangle \otimes |\alpha_2\rangle$ であるとする、

$$\exp(itH)(|\alpha_1\rangle \otimes |\alpha_2\rangle) = |\alpha_1 \cos t + \alpha_2 \sin t\rangle \otimes |-\alpha_1 \sin t + \alpha_2 \cos t\rangle.$$

と時間発展する. 量子ガウス状態が初期状態である場合には

$$\exp(itH)(\rho_{\theta_1,N} \otimes \rho_{\theta_2,N}) \exp(-itH) = \rho_{\theta_1 \cos t + \theta_2 \sin t, N} \otimes \rho_{-\theta_1 \sin t + \theta_2 \cos t, N}.$$

と時間発展する. それゆえ, 逐次的な時間発展 $V_n := \prod_{j=2}^n \exp(it_j H_j)$, $H_j := i(a_j^\dagger a_1 - a_1^\dagger a_j)$ を考え, 各時間が $\cos t_j = \sqrt{\frac{j-1}{j}}$, $\sin t_j = \sqrt{\frac{1}{j}}$ を満たすとする, と

$$V_n \rho_{\theta,N}^{\otimes n} V_n^\dagger = \rho_{\sqrt{n}\theta, N} \otimes \rho_{0,N}^{\otimes(n-1)},$$

が得られる.

E 量子効率 t の個数測定

量子効率 t の個数測定とは 1 個の光子を検知できる確率が t となる (すなわち光子を検知できない確率が $1-t$ となる) 測定である. それゆえ, 入力状態が $|k\rangle\langle k|$ であるときに, 全く光子を検知できない確率は, $(1-t)^k$ となる. そして, 量子効率 t の個数測定の POVM での「出力が 0 個」に対応する作用素は $\sum_{k=0}^{\infty} |k\rangle\langle k|(1-t)^k$ となる.

さらに, 現実の個数測定では 1 個と 2 個以上を区別することが不可能である場合が多く, そのような事情を考慮すると, 測定値は, 光子を検出したか否かの 2 値となる. したがって, その POVM は

$$\left\{ \sum_{k=0}^{\infty} |k\rangle\langle k|(1-t)^k, I - \sum_{k=0}^{\infty} |k\rangle\langle k|(1-t)^k \right\} \quad (45)$$

となる [12].

F (23) の証明

$\log(\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger \text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger) = \log(\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger) + \log(\text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger)$ であるから,

$$\begin{aligned} & \int_{\text{SU}(2)} \left(\frac{d \log(\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger \text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger)}{dr} \right)^2 (\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger \text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger) \nu(dg) \\ &= \int_{\text{SU}(2)} \left(\frac{d \log(\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger)}{dr} \right)^2 (\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger \text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger) \nu(dg) \\ & \quad + \int_{\text{SU}(2)} \left(\frac{d \log(\text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger)}{dr} \right)^2 (\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger \text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger) \nu(dg) \\ & \quad + 2 \int_{\text{SU}(2)} \left(\frac{d \log(\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger)}{dr} \right) \left(\frac{d \log(\text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger)}{dr} \right) (\text{Tr} \rho_{r,\theta g} \rho_{1,k} g^\dagger \text{Tr} \rho_{r,\theta g} \rho_{2,k} g^\dagger) \nu(dg) \end{aligned}$$

となる. ρ_1 と ρ_2 はともに純粋状態であるので右辺の第 1 項と第 2 項は一致する. よって, 第 1 項と第 3 項のみ計算すれば十分である. $\rho_{1,k} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\rho_{2,k} = \begin{pmatrix} \beta_k & \sqrt{\beta_k(1-\beta_k)}e^{-ia} \\ \sqrt{\beta_k(1-\beta_k)}e^{ia} & 1-\beta_k \end{pmatrix}$

となる。また不変測度 ν は $\frac{1}{2\pi^2} \sin 2\theta d\phi d\theta d\psi$ で与えられる。したがって、

$$\begin{aligned}
& \int_{\text{SU}(2)} \left(\frac{d \log (\text{Tr } \rho_{r,\theta} g \rho_{1,k} g^\dagger)}{dr} \right)^2 \left(\text{Tr } \rho_{r,\theta} g \rho_{1,k} g^\dagger \text{Tr } \rho_{r,\theta} g \rho_{2,k} g^\dagger \right) \nu(dg) \\
&= \int_0^\pi \int_0^{\frac{\pi}{2}} \frac{1}{4\pi} \left(\frac{\cos 2\theta}{1+r \cos 2\theta} \right)^2 (1+r \cos 2\theta), \\
&\quad \left(1+r \left(\cos 2\theta(2\beta_k-1) + 2\sqrt{\beta_k(1-\beta_k)} \sin 2\theta \cos 2(\psi+a) \right) \right) \sin 2\theta d\theta d\psi \\
&= \frac{2\beta_k-1}{12} + \frac{\beta_k-1}{2r^2} + \frac{1-\beta_k}{4r^3} \log \left(\frac{1+r}{1-r} \right), \\
& \int_{\text{SU}(2)} \left(\frac{d \log (\text{Tr } \rho_{r,\theta} g \rho_{1,k} g^\dagger)}{dr} \right) \left(\frac{d \log (\text{Tr } \rho_{r,\theta} g \rho_{2,k} g^\dagger)}{dr} \right) \left(\text{Tr } \rho_{r,\theta} g \rho_{1,k} g^\dagger \text{Tr } \rho_{r,\theta} g \rho_{2,k} g^\dagger \right) \nu(dg) \\
&= \int_0^\pi \int_0^{\frac{\pi}{2}} \frac{1}{4\pi} (\cos 2\theta) \left(\cos 2\theta(2\beta_k-1) + 2\sqrt{\beta_k(1-\beta_k)} \sin 2\theta \cos 2(\psi+a) \right) \sin 2\theta d\theta d\psi \\
&= \frac{2\beta_k-1}{12}
\end{aligned}$$

となり (23) を得る。

参考文献

- [1] C. W. Helstrom, Minimum Mean-Square Error Estimation in Quantum Statistics, Physics Letters **25A**, 101-102 (1967).
- [2] H. P. Yuen, and M. Lax, Multiple-Parameter Quantum Estimation and Measurement of Nonselfadjoint Observables, IEEE Trans. **IT-19**, 740-750 (1973).
- [3] C. W. Helstrom, Quantum Detection and Estimation Theory, Academic Press, New York (1976).
- [4] A. S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory, North-Holland, Amsterdam (1982).
- [5] 林 正人, 松本 啓史, 量子系における統計的推測の最近の発展, 日本応用数理学会誌, Vol. 11, No. 3, Sep., 27-48 (2001).
- [6] R. Gill and S. Massar, "State estimation for large ensembles," Phys. Rev. A **61**, 042312 (2000); LANL E-print <http://xxx.lanl.gov/abs/quant-ph/9902063> (1999).
- [7] P. G. Kwiat, A. G. White, et al., Phys. Rev. A **60**, R773 (1999); Phys. Rev. Lett. **83**, 3103 (1999); Science **290**, 498 (2000); Nature **409**, 1014 (2001); LANL E-print <http://xxx.lanl.gov/abs/quant-ph/0108088> (2001).

- [8] K. Usami, Y. Nambu, Y. Tsuda, K. Matsumoto, and K. Nakamura, "A New Strategy of Quantum-State Estimation for Achieving the Cramer-Rao Bound," LANL E-print <http://xxx.lanl.gov/abs/quant-ph/0209074> (2003).
- [9] 林 正人, 量子推定と量子中心極限定理, 日本数学会誌「数学」 55 巻 4 号, 掲載予定.
- [10] Marlan O. Scully and M. Scharl Zubairy, Quantum optics, Cambridge University Press (1997).
- [11] M. Hayashi, Asymptotic quantum theory for the thermal states family, Quantum Communication, Computing, and Measurement 2 (edited by Kumar, P., D'ariano, G. M., and Hirota, O.), Plenum, New York, 99-104 (2000). LANL E-print <http://xxx.lanl.gov/abs/quant-ph/9809002> (1998).
- [12] S. K. Ozdemir, A. Miranowicz, M. Koashi, and N. Imoto, "Quantum-scissors device for optical state truncation: A proposal for practical realization," Phys. Rev. A **64**, 063818 (2001); LANL E-print <http://xxx.lanl.gov/abs/quant-ph/0107048> (2001).
- [13] Shigeki Takeuchi, Jungsang Kim, Yoshihisa Yamamoto, and Henry H. Hogue, "Development of a High-Quantum-Efficiency Single-Photon Counting System," Applied Physics Letters **74**, 1063 (1999).
- [14] 林 正人, 松本 啓史, "測定自由度をもつ統計モデルと量子力学," 数理解析研究所講究録 **1055**, 96-110 (1998).
- [15] 林 正人, "Spin 1/2 系 純粋状態ファミリーの漸近推定理論," 第 5 回「非平衡系の統計物理」シンポジウム報告集, 物性研究, Vol. 71, No. 5 (2 月号), 848-703 (1998).
- [16] Masahito Hayashi, "Asymptotic estimation theory for a finite dimensional pure state model," J. Phys. A: Math. and Gen. **31**, 4633-4655 (1998); LANL E-print <http://xxx.lanl.gov/abs/quant-ph/9704041> (1997).
- [17] K. Keyl and R. F. Werner, Phys. Rev. A **64**, 052311 (2001); LANL E-print <http://xxx.lanl.gov/abs/quant-ph/0102027>.
- [18] 松本 啓史, セミナー資料 (1999).
- [19] H. Nagaoka, "On the Parameter Estimation Problem for Quantum Statistical Models," 第 12 回情報理論とその応用学会シンポジウム予稿集, pp. 577-582 (1989).
- [20] 長岡 浩司, Kullback divergence と Fisher information の関係について —古典系から量子系へ—, 第 15 回情報理論とその応用シンポジウム予稿集, pp. 63-72 (1992).

- [21] M. Hayashi, “Two quantum analogues of Fisher information from a large deviation viewpoint of quantum estimation,” *J. Phys. A: Math. and Gen.* **35**, 7689-7727 (2002); LANL E-print <http://xxx.lanl.gov/abs/quant-ph/0202003> (2002).
- [22] M. Hendrych, M. Dusek, R. Filip, and J. Fiurasek, “Simple optical measurement of the overlap and fidelity of quantum states: An experiment,” LANL E-print <http://xxx.lanl.gov/abs/quant-ph/0208091> (2002).