

QUANTUM STATE ESTIMATION AND  
SYMMETRIC INFORMATIONALLY COMPLETE POMs

ZHU HUANGJUN

NATIONAL UNIVERSITY OF SINGAPORE

2012



**QUANTUM STATE ESTIMATION AND  
SYMMETRIC INFORMATIONALLY COMPLETE POMs**

**ZHU HUANGJUN**

(M.Sc., PEKING UNIVERSITY)

**A THESIS SUBMITTED FOR THE  
DEGREE OF DOCTOR OF PHILOSOPHY**

**NUS GRADUATE SCHOOL FOR INTEGRATIVE  
SCIENCES AND ENGINEERING  
CENTRE FOR QUANTUM TECHNOLOGIES  
NATIONAL UNIVERSITY OF SINGAPORE**


2012



## Declaration

I hereby declare that the thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

A handwritten signature in black ink, reading 'Zhu Huangjun', written in a cursive style. A horizontal line is drawn underneath the signature.

Zhu Huangjun

30 March 2012

# Acknowledgments

I am sincerely grateful to my supervisor Berthold-Georg Englert for his guidance, for giving me the opportunities and freedom to explore various interesting topics, and for encouraging me to attend conferences and make presentations. I would also like to thank Markus Grassl for numerous stimulating discussions, especially those on symmetric informationally complete probability operator measurements (SIC POMs), and for critical comments and suggestions on writing. Special thanks to Chen Lin, Masahito Hayashi, Teo Yong Siah, Wei Tzu-Chieh, Jaroslav Řeháček, Zdeněk Hradil, Valerio Scarani, Cyril Branciard, and Xu Aimin for fruitful discussions and collaborations. I am also grateful to Marcus Appleby, Christopher Fuchs, Ingemar Bengtsson, and Andreas Winter for discussions and encouragement. Special thanks also to Yao Penghui, Ng Hui Khoon, Amir Kalev, Philippe Raynal, Tan Si-Hui, Lü Xin, Wang Guangquan, Thi-ang Guo Chuan, Tomasz Paterek, Tomasz Karpiuk, Ma Jia Jun, Kwek Leong Chuan, Thomas Durt, Daniel Greenberger, and Andrew Scott for stimulating discussions. I would also like to thank Wang Jian-Sheng and Gong Jiangbin for serving on my thesis advisory committee. Special thanks to Dai Li and Lee Kean Loon for enthusiastic help related to the format of the thesis. Special thanks also to the examiners for reviewing this thesis and for providing generous comments and suggestions. I would like to acknowledge the financial support from NUS Graduate School for Integrative Sciences and Engineering (NGS), and I am grateful to many administrative staff for their dedication to the welfare of students. I am also grateful to Centre for Quantum Technologies and many administrative staff for providing a comfortable research environment and numerous timely help. Finally, many thanks to my parents and my sister for their continuous support and understanding, and to my friends for their encouragement.

During the PhD candidature, I have mainly worked on three related topics: quantum state estimation, SIC POMs, and multipartite entanglement. This thesis covers the main results concerning the first two topics, most of which have not been published. Chapters 3, 8, and 9 are based on the following three papers, respectively:

- 
- **H. Zhu** and B.-G. Englert, *Quantum state tomography with fully symmetric measurements and product measurements*, Phys. Rev. A **84**, 022327 (2011).
  - **H. Zhu**, *SIC POVMs and Clifford groups in prime dimensions*, J. Phys. A: Math. Theor. **43**, 305305 (2010).
  - **H. Zhu**, Y. S. Teo, and B.-G. Englert, *Two-qubit symmetric informationally complete positive-operator-valued measures*, Phys. Rev. A **82**, 042308 (2010).

Other papers not covered in this thesis:

- Y. S. Teo, **H. Zhu**, B.-G. Englert, J. Řeháček, and Z. Hradil, *Quantum-state reconstruction by maximizing likelihood and entropy*, Phys. Rev. Lett. **107**, 020404 (2011).
- L. Chen, **H. Zhu**, and T.-C. Wei, *Connections of geometric measure of entanglement of pure symmetric states to quantum state estimation*, Phys. Rev. A **83**, 012305 (2011).
- **H. Zhu**, L. Chen, and M. Hayashi, *Additivity and non-additivity of multipartite entanglement measures*, New J. Phys. **12**, 083002 (2010).
- L. Chen, A. Xu, and **H. Zhu**, *Computation of the geometric measure of entanglement for pure multiqubit states*, Phys. Rev. A **82**, 032301 (2010).
- C. Branciard, **H. Zhu**, L. Chen, and V. Scarani, *Evaluation of two different entanglement measures on a bound entangled state*, Phys. Rev. A **82**, 012327 (2010).
- **H. Zhu**, Y. S. Teo, and B.-G. Englert, *Minimal tomography with entanglement witnesses*, Phys. Rev. A **81**, 052339 (2010).
- Y. S. Teo, **H. Zhu**, and B.-G. Englert, *Product measurements and fully symmetric measurements in qubit-pair tomography: A numerical study*, Opt. Commun. **283**, 724 (2010).

# Contents

<b>Acknowledgments</b>	<b>ii</b>
<b>Summary</b>	<b>x</b>
<b>List of Tables</b>	<b>xi</b>
<b>List of Figures</b>	<b>xii</b>
<b>List of Abbreviations</b>	<b>xiv</b>
<b>List of Symbols</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Quantum state estimation . . . . .	1
1.2 Symmetric informationally complete POMs . . . . .	7
<b>2 Quantum state estimation</b>	<b>11</b>
2.1 Introduction . . . . .	11
2.2 Historical background . . . . .	12
2.3 Quantum states and measurements . . . . .	18
2.3.1 Simple systems . . . . .	18
2.3.2 Composite systems . . . . .	20
2.4 Quantum state reconstruction . . . . .	21
2.4.1 Linear state reconstruction . . . . .	22
2.4.2 Maximum-likelihood estimation . . . . .	23
2.4.3 Other reconstruction methods . . . . .	25



## Contents

---

2.5	Fisher information and Cramér–Rao bound . . . . .	27
<b>3</b>	<b>Fully symmetric measurements and product measurements</b>	<b>31</b>
3.1	Introduction . . . . .	31
3.2	Setting the stage . . . . .	34
3.2.1	Linear state tomography . . . . .	34
3.2.2	Tight IC measurements . . . . .	36
3.3	Applications of random-matrix theory to quantum state tomography . .	38
3.3.1	A simple idea . . . . .	38
3.3.2	Isotropic measurements . . . . .	40
3.3.3	Tight IC POMs and SIC POMs . . . . .	42
3.3.4	Qubit tomography . . . . .	45
3.4	Joint SIC POMs and Product SIC POMs . . . . .	50
3.4.1	Bipartite scenarios . . . . .	50
3.4.2	Multipartite scenarios . . . . .	53
3.5	Summary . . . . .	55
<b>4</b>	<b>The power of informationally overcomplete measurements</b>	<b>57</b>
4.1	Introduction . . . . .	57
4.2	Optimal state reconstruction . . . . .	59
4.2.1	Optimal reconstruction in the perspective of frame theory . . . .	60
4.2.2	Connection with the maximum-likelihood method . . . . .	63
4.3	Quantum state estimation with mutually unbiased measurements . . . .	64
4.4	Efficiency of covariant measurements . . . . .	68
4.5	Informationally overcomplete measurements on the two-level system . .	72
4.6	Summary . . . . .	75
<b>5</b>	<b>Optimal state estimation with adaptive measurements</b>	<b>77</b>
5.1	Introduction . . . . .	77
5.2	Quantum Fisher information and quantum CR bound . . . . .	80
5.2.1	One-parameter setting . . . . .	80

5.2.2	Multiparameter setting . . . . .	84
5.3	Gill–Massar trace and Gill–Massar bound . . . . .	85
5.3.1	Reexamination of the Gill–Massar inequality . . . . .	86
5.3.2	Gill–Massar bound for the scaled WMSE . . . . .	87
5.3.3	Gill–Massar bounds for the mean square Bures distance and the mean square HS distance . . . . .	89
5.4	Optimal quantum state estimation with adaptive measurements . . . . .	92
5.4.1	A general recipe . . . . .	93
5.4.2	Approximate saturation of the Gill–Massar bound for the MSB .	100
5.4.3	Degenerate two-level systems . . . . .	103
5.4.4	Comparison with nonadaptive schemes . . . . .	107
5.5	Summary and open problems . . . . .	110
<b>6</b>	<b>Quantum state estimation with collective measurements</b>	<b>113</b>
6.1	Introduction . . . . .	113
6.2	Efficiency of asymptotic state estimation . . . . .	115
6.2.1	Quantum Cramér–Rao bound based on the right logarithmic derivative . . . . .	115
6.2.2	Efficiency of the optimal state estimation in the asymptotic limit	118
6.3	Quantum state estimation with coherent measurements . . . . .	122
6.3.1	Schur–Weyl duality and its implications . . . . .	123
6.3.2	Highest-weight states and coherent states . . . . .	125
6.3.3	Coherent measurements . . . . .	127
6.3.4	Complementarity polynomials . . . . .	128
6.3.5	Estimation of highly mixed states with collective measurements .	133
6.4	Collective measurements in qubit state estimation . . . . .	136
6.4.1	A lower bound for the weighted mean square error . . . . .	137
6.4.2	Fisher information matrices for coherent measurements . . . . .	140
6.4.3	Complementarity polynomials . . . . .	144
6.4.4	Mean square error and mean square Bures distance . . . . .	146

## Contents

---

6.5	Summary and open problems . . . . .	149
<b>7</b>	<b>Symmetric informationally complete POMs</b>	<b>151</b>
7.1	Introduction . . . . .	151
7.2	Symmetry and group covariance . . . . .	156
7.2.1	Groups that can generate SIC POMs . . . . .	157
7.2.2	Orbits and equivalence of SIC POMs . . . . .	159
7.3	Heisenberg–Weyl group and Clifford group . . . . .	160
7.3.1	Heisenberg–Weyl group . . . . .	161
7.3.2	Special linear group . . . . .	163
7.3.3	Understanding the Clifford group from a homomorphism . . . . .	163
<b>8</b>	<b>SIC POMs in prime dimensions</b>	<b>167</b>
8.1	Introduction . . . . .	167
8.2	Group covariant SIC POMs are HW covariant . . . . .	168
8.3	Qubit SIC POMs . . . . .	168
8.4	SIC POMs in prime dimensions not equal to 3 . . . . .	169
8.5	SIC POMs in dimension 3 . . . . .	171
8.5.1	Symmetry of SIC POMs . . . . .	172
8.5.2	Infinitely many inequivalent SIC POMs . . . . .	176
8.6	Beyond prime dimensions . . . . .	180
8.7	Summary . . . . .	181
<b>9</b>	<b>Two-qubit SIC POMs</b>	<b>183</b>
9.1	Introduction . . . . .	183
9.2	Structure of SIC POMs in the four-dimensional Hilbert space . . . . .	184
9.2.1	Symmetry transformations within an HW covariant SIC POM . . . . .	185
9.2.2	Symmetry transformations among HW covariant SIC POMs . . . . .	186
9.2.3	SIC POM regrouping phenomena . . . . .	189
9.3	Two-qubit SIC POMs . . . . .	191
9.3.1	Two-qubit SIC POMs in the product basis . . . . .	192

9.3.2 Two-qubit SIC POMs in the Bell basis . . . . .	197
9.4 Summary . . . . .	198
<b>10 Symmetry and equivalence</b>	<b>199</b>
10.1 Introduction . . . . .	199
10.2 SIC POMs and graph automorphism problem . . . . .	200
10.2.1 Unitary symmetry and permutation symmetry . . . . .	201
10.2.2 A connection with the graph automorphism problem . . . . .	202
10.2.3 An algorithm . . . . .	206
10.3 HW covariant SIC POMs . . . . .	211
10.3.1 SIC POMs in dimension 3 revisited . . . . .	212
10.3.2 Symmetry and equivalence . . . . .	215
10.3.3 Nice error bases in the symmetry group . . . . .	216
10.4 Hoggar lines . . . . .	220
10.5 Quest for new SIC POMs . . . . .	224
10.6 Summary and open questions . . . . .	226
 <b>Appendix:</b>	
<b>A Several distance and distinguishability measures</b>	<b>229</b>
A.1 Hilbert–Schmidt distance and trace distance . . . . .	229
A.2 Fidelity and Bures distance . . . . .	230
<b>B Weighted <math>t</math>-designs</b>	<b>232</b>
<b>C Proof of Lemma 4.1</b>	<b>233</b>
<b>D Supplementary materials about adaptive measurements</b>	<b>234</b>
D.1 Derivation of Eqs. (5.23) and (5.24) . . . . .	234
D.2 Connection with pure-state estimation . . . . .	235
D.3 Discontinuity of the minimal scaled MSB . . . . .	237

## Contents

---

<b>E</b>	<b>Technical details about Chapter 6</b>	<b>240</b>
E.1	Proof of Eq. (6.49) for Slater-determinant states . . . . .	240
E.2	Proof of Conjecture 6.4 for symmetric and antisymmetric subspaces . . .	241
E.3	Proof of Theorem 6.7 . . . . .	243
E.4	Proof of Theorem 6.8 . . . . .	244
<b>F</b>	<b>Nice error bases</b>	<b>245</b>
<b>G</b>	<b><math>p</math>-groups and Sylow's theorem</b>	<b>246</b>
<b>H</b>	<b>Supplementary information about the Clifford group</b>	<b>247</b>
H.1	Trace of a Clifford unitary operator . . . . .	247
H.2	Normalizer of the Clifford group . . . . .	251
H.3	HW groups in the Clifford group in a prime dimension . . . . .	253
<b>I</b>	<b>Some basic concepts in graph theory</b>	<b>255</b>
	<b>Bibliography</b>	<b>257</b>
	<b>Index</b>	<b>275</b>

# Summary

This thesis studies two basic topics in quantum information science: quantum state estimation and symmetric informationally complete probability operator measurements (SIC POMs)<sup>1</sup>.

Part I of this thesis focuses on reliable and efficient estimation of mixed states of finite-dimensional quantum systems in the large-sample scenario. Four natural settings are investigated in the order of sophistication levels: independent and identical measurements with linear reconstruction, as well as optimal reconstruction, adaptive measurements, and collective measurements. We present an overview of the optimal estimation strategies and tomographic efficiencies under the four settings with respect to typical figures of merit, such as the mean square Hilbert–Schmidt distance, the mean square Bures distance, and the mean trace distance. The distinctive features of each setting and the efficiency differences among different settings are discussed in detail. Our study also highlights the connection between quantum state estimation and basic principles of quantum mechanics, especially the complementarity principle.

Part II of this thesis presents an overview on the symmetry properties of SIC POMs. We start by deriving several key attributes about group covariant SIC POMs. We then settle several persistent open problems concerning such SIC POMs in prime dimensions and clarify a few subtle points in the special case of dimension 3. Several peculiar features relevant to composite dimensions, such as regrouping phenomena and entanglement properties, are illustrated with two-qubit SIC POMs. Finally, we develop a powerful graph-theoretic approach, thereby determining the symmetry groups of all SIC POMs appearing in the literature and establishing complete equivalence relations among them. The connection between SIC POMs and nice error bases are also explicated. Our study indicates that, except for the set of Hoggar lines, all SIC POMs known so far are covariant with respect to the Heisenberg–Weyl groups.

---

<sup>1</sup>Also called symmetric informationally complete positive-operator-valued measures (SIC POVMs).

# List of Tables

3.1	Theoretical and numerical scaled mean trace distances in two-qubit state tomography with the joint SIC POM and the product SIC POM . . . . .	53
8.1	Geometric phases associated with triple products among fiducial states of SIC POMs in dimension 3 . . . . .	179
9.1	Arrangement of the 16 HW covariant SIC POMs in dimension 4 . . . . .	187
9.2	Generalized Bloch vector of a two-qubit state . . . . .	192
9.3	Generalized Bloch vectors of fiducial states of two-qubit SIC POMs in the product basis . . . . .	194
9.4	Invariants of two-qubit SIC POMs in the product basis . . . . .	195
9.5	Generalized Bloch vectors of fiducial states of two-qubit SIC POMs in the Bell basis . . . . .	197
10.1	Conjugacy classes of the stabilizer of a fiducial state of the Hoggar lines	222
10.2	SIC POMs generated by the nice error bases cataloged by Klappenecker and Rötteler . . . . .	226

# List of Figures

3.1	Scaled mean trace distances and scaled mean HS distances in state tomography with SIC POMs for dimensions from 2 to 45 . . . . .	43
3.2	Uncertainty ellipses and tomographic efficiencies in linear state tomography on a qubit with the MUB and the SIC POM . . . . .	48
3.3	Ratio of the MSE associated with the product SIC POM to that with the joint SIC POM . . . . .	52
3.4	Theoretical and numerical scaled mean trace distances for the joint SIC POMs and the product SIC POMs on multiqubit systems . . . . .	55
4.1	Tomographic efficiencies of covariant measurements with respect to the scaled MSE and the scaled MSB . . . . .	70
4.2	Tomographic efficiencies of the SIC, MUB, cube, and covariant measurements in qubit state estimation as well as uncertainty ellipses of the MUB measurement with the canonical and the optimal reconstructions . . . . .	74
5.1	Tomographic efficiencies of the optimal adaptive measurements with respect to the scaled MSH and the scaled MSB when the true states have the form $s 1\rangle\langle 1  + (1 - s)/d$ for $d = 2, 5, 10, 15, 20$ . . . . .	107
5.2	Tomographic efficiencies with respect to the scaled MSH of the standard state estimation, state estimation with covariant measurements, and state estimation with optimal adaptive measurements . . . . .	108
5.3	Tomographic efficiencies with respect to the scaled MSB of covariant measurements and optimal adaptive measurements . . . . .	109
6.1	Contour plots of the asymptotic maximal scaled MSE, MSB, and the minimal scaled GMT in the eigenvalue simplex for dimension 3. . . . .	121
6.2	Maximal scaled GMT and minimal scaled MSE at $\rho = 1/d$ over all collective measurements on $\rho^{\otimes N}$ for $d = 2, 3, \dots, 8$ and $N = 2, 3, \dots, 40$ . . . . .	135



## List of Figures

---

6.3	Maximal scaled GMT over all measurements on $N$ copies of a qubit state for $N = 1, 2, 3, 4, 5, 10, 20, 100, \infty$ . . . . .	145
6.4	Efficiencies of covariant coherent measurements and optimal coherent measurements in qubit state estimation with respect to the scaled MSH and the scaled MSB . . . . .	148
8.1	Geometric phases associated with triple products among fiducial states of SIC POMs in dimension 3 . . . . .	180
9.1	Symmetry transformations among HW covariant SIC POMs in dimension 4 . . . . .	188
10.1	Nice graph and “wicked” graph . . . . .	211
10.2	Graph representation of the angle matrices associated with HW covariant SIC POMs in dimension 3 . . . . .	213
D.1	Discontinuity of the minimal scaled MSB at the boundary of the state space . . . . .	239

# List of Abbreviations

BME	Bayesian mean estimation
CR	Cramér–Rao
GM	Gill–Massar
GMT	Gill–Massar trace
HMLE	Hedged maximum-likelihood estimation
HS	Hilbert–Schmidt
HW	Heisenberg–Weyl
IC	Informationally complete
LOCC	Local operations and classical communication
ME	Maximum entropy
ML	Maximum likelihood
MLE	Maximum-likelihood estimation
MSB	Mean square Bures distance
MSE	Mean square error
MSH	Mean square Hilbert–Schmidt distance
MUB	Mutually unbiased bases
POM	Probability operator measurement
POVM	Positive-operator-valued measure
QPT	Quantum process tomography
RLD	Right logarithmic derivative
ROP	Recursive ordered partition
SIC	Symmetric informationally complete
SLD	Symmetric logarithmic derivative
WMSE	Weighted mean square error

# List of Symbols

$\ \cdot\ _{\text{HS}}$	Hilbert–Schmidt norm	229
$\ \cdot\ _{\text{tr}}$	Trace norm	229
$\langle \mathbf{k}, \mathbf{q} \rangle$	$:= k_2 q_1 - k_1 q_2$ , symplectic form	162
$\text{Aut}(\cdot)$	Automorphism group	256
$\text{Aut}_{\mathbb{E}}(\cdot)$	Extended automorphism group	256
$C(d), \overline{C}(d)$	Clifford group in dimension $d$	163
$C(\theta)$	Scaled MSE matrix	28
$\mathcal{C}, \mathcal{C}(\rho)$	Scaled MSE matrix in superoperator form	35
$d$	Dimension of the Hilbert space	2
$\bar{d}$	$:= \begin{cases} d & \text{if } d \text{ is odd} \\ 2d & \text{if } d \text{ is even} \end{cases}$	163
$d_{\mu}$	Dimension of the representation $\mu$ of the symmetric group	123
$D, \overline{D}$	Heisenberg–Weyl group	161
$D_{\mathbf{k}}, D_{k_1, k_2}$	Displacement operator, element in the Heisenberg–Weyl group	161
$D_{\text{B}}(\rho, \sigma)$	Bures distance between $\rho$ and $\sigma$	231
$D_{\mu}$	Dimension of the representation $\mu$ of the unitary group	123
$\Delta \rho$	$:= \sqrt{N}(\hat{\rho} - \rho)$ , scaled deviation of the estimator	35
$\mathbb{E}(\cdot)$	Expectation value	28
$E_{jk}$	$:=  j\rangle\langle k $	29
$E_{jk}^+$	$:= \frac{1}{\sqrt{2}}( j\rangle\langle k  +  k\rangle\langle j )$	29
$E_{jk}^-$	$:= -\frac{i}{\sqrt{2}}( j\rangle\langle k  -  k\rangle\langle j )$	29
$\mathcal{E}(\rho)$	Scaled mean square error	35
$\mathcal{E}_{\text{HS}}(\rho)$	Scaled mean Hilbert–Schmidt distance	40
$\mathcal{E}_{\text{SB}}(\rho)$	Scaled mean square Bures distance	70
$\mathcal{E}_{\text{SH}}(\rho)$	Scaled mean square Hilbert–Schmidt distance	85
$\mathcal{E}_{\text{tr}}(\rho)$	Scaled mean trace distance	39
$\mathcal{E}_{\text{W}}(\rho)$	Scaled weighted mean square error	88

$\mathcal{E}_{\mathcal{W}}(\rho)$	Scaled weighted mean square error	88
$\text{EC}(d), \overline{\text{EC}}(d)$	Extended Clifford group in dimension $d$	163
$EG_{\text{sym}}, \overline{EG}_{\text{sym}}$	Extended symmetry group of a SIC POM	157
$\text{ESL}(2, \mathbb{Z}_d)$	Extended special linear group over $\mathbb{Z}_d$	163
$f_{\xi}$	Frequencies	22
$F_L$	$:= \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$	165
$F_Z$	$:= \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$ , Zauner matrix	165
$\mathcal{F}$	Frame superoperator	34
$\bar{\mathcal{F}}$	$:= \bar{\mathbf{I}}\mathcal{F}\bar{\mathbf{I}}$ , frame superoperator	34
$\mathcal{F}(\rho)$	Frame superoperator, Fisher information matrix	60
$\bar{\mathcal{F}}(\rho)$	$:= \bar{\mathbf{I}}\mathcal{F}(\rho)\bar{\mathbf{I}}$ , frame superoperator, Fisher information matrix	62
$(F, \chi)$	Element in the group $\text{ESL}(2, \mathbb{Z}_d) \times (\mathbb{Z}_d)^2$ or $\text{ESL}(2, \mathbb{Z}_{\bar{d}}) \times (\mathbb{Z}_d)^2$	163
$[F, \chi]$	Clifford operation, homomorphism image of $(F, \chi)$	165
$G_{\text{sym}}, \overline{G}_{\text{sym}}$	Symmetry group of a SIC POM	157
$ G $	Order of $G$	157
$\text{GL}(\mathcal{H})$	General linear group on $\mathcal{H}$	123
$\mathcal{H}$	Hilbert space	20
$h_{\mu}$		123
$H_{\mu}$	Projector onto $\mathcal{H}_{\mu}$	123
$\mathcal{H}_{\mu}$		123
$\text{ht}(\mu)$	Height of the partition $\mu$	123
$I, I(\theta)$	Scaled Fisher information matrix	27
$\mathbf{I}$	Identity superoperator	35
$\bar{\mathbf{I}}$	Projector onto the space of traceless Hermitian operators	34
$\Im$	Imaginary part of a complex number or matrix	*
$J$	SLD Fisher information matrix	81
	$:= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	163
$\tilde{J}$	RLD Fisher information matrix	116
$\hat{J}$	Complex-conjugation operator	165

## List of Symbols

---

$\mathcal{J}$	SLD Fisher information matrix in superoperator form . . . . .	83
$\bar{\mathcal{J}}$	$:= \bar{\mathbf{I}}\mathcal{J}\bar{\mathbf{I}}$ , SLD Fisher information matrix . . . . .	83
$\mathcal{K}_\mu$	An irreducible subspace in $\mathcal{H}^{\otimes N}$ of the symmetric group . . . . .	123
$L$	Symmetric logarithmic derivative (SLD) . . . . .	80
$L_j$	Symmetric logarithmic derivative (SLD) . . . . .	84
$\tilde{L}_j$	Right logarithmic derivative (RLD) . . . . .	115
$\mathcal{L}(\rho)$	Likelihood functional . . . . .	23
$\lambda_j, \lambda_k$	Eigenvalues of $\rho$ . . . . .	90
$\Lambda$	Angle matrix . . . . .	204
$\Lambda_{M,N}(\rho), \Lambda_N(\rho)$	. . . . .	129
$\mu$	Partition . . . . .	123
$\tilde{\mu}$	Dual partition of $\mu$ . . . . .	126
$ \mu $	$:= \sum_{j=1}^d \mu_j$ . . . . .	132
$N$	Number of copies of states measured for state estimation . . . . .	6
$\omega$	$:= e^{2\pi i/d}$ , a primitive $d$ th root of unity . . . . .	161
$p_\xi$	Probabilities . . . . .	18
$\Phi_t$	Frame potential of order $t$ . . . . .	232
$\Pi_\xi$	Measurement outcomes . . . . .	19
$\bar{\Pi}_\xi$	$:= \Pi_\xi - \text{tr}(\Pi_\xi)/d$ , measurement outcomes . . . . .	34
$\Re$	Real part of a complex number or matrix. . . . .	*
$\rho$	A generic quantum state . . . . .	6
$\hat{\rho}$	An estimator of $\rho$ . . . . .	22
$ \rho $	Determinant of $\rho$ . . . . .	140
$\rho_{,j}$	$:= \frac{\partial \rho}{\partial \theta_j}$ . . . . .	84
$S_N$	Symmetric group of $N$ letters . . . . .	123
$S_\mu$	Projector onto $\mathcal{S}_\mu$ . . . . .	124
$\mathcal{S}_\mu$	An irreducible subspace in $\mathcal{H}^{\otimes N}$ of the general linear group . . . . .	123
$s_\mu(\cdot)$	Schur symmetric polynomial . . . . .	124
$\text{SL}(2, \mathbb{Z}_d)$	Special linear group over $\mathbb{Z}_d$ . . . . .	163

$\sigma_x, \sigma_y, \sigma_z$	Pauli matrices.....*
$t(\cdot)$	Gill–Massar trace ..... 86
$t^N(\cdot)$	Complementarity polynomial ..... 131
$\bar{t}^N(\cdot)$	Homogeneous complementarity polynomial ..... 131
$t_\mu(\cdot)$	Complementarity polynomial for the subspace $\mathcal{S}_\mu$ ..... 131
$\bar{t}_\mu(\cdot)$	Homogeneous complementarity polynomial for the subspace $\mathcal{S}_\mu$ 131
$T_\mu(\rho)$	..... 129
$T_{jkl}$	Triple product ..... 203
tr	Trace of an ordinary operator ..... 35
Tr	Trace of a superoperator ..... 35
$\tau$	$:= -e^{\pi i/d}$ ..... 161
$\Theta_\xi$	Reconstruction operators ..... 22
$\vartheta_{jkl}$	Angle tensor ..... 203
$U_\rho$	Stabilizer of $\rho$ under the action of the unitary group ..... 93
$V$	Swap operator ..... 124
$V(j, k)$	Swap operator between party $j$ and party $k$ ..... 129
$V_F$	Clifford unitary transformation ..... 164
$V_L$	..... 165
$V_Z$	Zauner unitary transformation ..... 165
Var( $\cdot$ )	Variance ..... 27
$W, W(\theta)$	Weight matrix ..... 29
$\mathcal{W}$	Weight matrix in superoperator form ..... 84
$X$	Cyclic-shift operator ..... 161
$y_\mu$	..... 123
$Z$	Phase operator ..... 161
$\mathbb{Z}_d$	Ring of integers modulo $d$ ..... 161
$\mathbb{Z}_p$	Galois field of integers modulo prime $p$ ..... 169
$\mathbb{Z}_p^*$	Group of nonzero elements in $\mathbb{Z}_p$ ..... 254

# Introduction

---

## 1.1 Quantum state estimation

Quantum state estimation is a procedure for inferring the state of a quantum system from generalized measurements, known as probability operator measurements (POMs). It is a primitive of many quantum information processing tasks, such as quantum computation, quantum communication, and quantum cryptography, because all these tasks rely heavily on our ability to determine the state of a quantum system at various stages [133, 186, 208]. Owing to the complementarity principle [41] and the uncertainty relation [138], any measurement on a generic quantum system necessarily induces a disturbance, limiting further attempts to extract information from the system. Therefore, it is impossible to infer a generic unknown state from measurements on a single quantum system; that is, an ensemble of identically prepared systems is needed for reliable state determination. One of the main challenges in quantum state estimation is to infer quantum states as efficiently as possible and to determine the resources necessary to achieve a given accuracy, which can be quantified by various figures of merit, such as the mean trace distance, the mean square Hilbert–Schmidt distance (MSH), or the mean fidelity (see Appendix A).

A good state-estimation strategy entails judicial choices on both measurement schemes and data processing protocols for reconstructing the true state. Compared with measurement schemes, there is generally more freedom in choosing the reconstruction methods in practice, and a good choice is the first step towards a reliable and efficient estimator. On the other hand, given the measurement results, the optimization of data processing is basically a subject of classical statistical inference, although

attention is required to account for additional constraints, such as the positivity of the density matrices. When the sample is reasonably large, a suitable figure of merit is the weighted mean square error (WMSE) for a certain weight matrix, of which the MSH and the mean square Bures distance (MSB) are special examples. It is well known in classical statistical inference that the minimal error is determined by the Fisher information matrix [94] through the Cramér–Rao (CR) bound [68, 224].

The main departure of quantum state estimation from classical state estimation is the choice over measurements, which underlies the differences between quantum information processing and classical information processing. In practice, the set of permissible measurements is mainly determined by experimental settings. As technology advances, it is ultimately limited by the basic principles of quantum mechanics. For example, as a consequence of the complementarity principle, it is impossible to measure two noncommuting sharp observables simultaneously [204], which implies that no measurement can extract maximal information about both observables simultaneously. Put differently, any gain of information about one observable is necessarily accompanied with a loss of information about the other. To devise good measurement schemes, it is crucial to balance such information trade-off, which is a main challenge in quantum estimation theory.

Part I of this thesis (Chapters 2 to 6) studies reliable and efficient estimation of the mixed states of a  $d$ -level quantum system. The main concern is the large-sample scenario, in which the classical CR bound can be saturated approximately, and the main focus is to devise measurement schemes that yield the most information. Our analysis should be applicable to most scenarios in which precision estimation is desired. Four natural settings will be investigated in order of sophistication levels: independent and identical measurements with linear reconstruction, independent and identical measurements with optimal reconstruction, adaptive measurements, and collective measurements. Our main goal, yet not fully realized, is to determine the optimal estimation strategies and the optimal tomographic efficiencies under the four settings in terms of common figures of merit, such as the mean trace distance, the MSH, and the MSB. In



## 1.1. Quantum state estimation

---

this way, we hope to establish a fairly complete picture about the main characteristics in each setting as well as their differences, such as in the tomographic efficiency and in the complexity. Our study can help elucidate the efficiency gap between experimental quantum state estimation and the theoretic limit, as well as reduce resource consumption by increasing the tomographic efficiency. Meanwhile, it may stimulate reflections on foundational issues, such as the complementarity principle, the uncertainty relation, and the geometry of quantum states, from the information-theoretic perspective.

Chapter 2 presents an overview of quantum state estimation from the theoretical perspective. We start with a historical survey of the major achievements in the field during the past half a century and then introduce several basic ingredients in quantum state estimation, such as quantum states, measurements, state reconstruction, Fisher information, and CR bound.

Chapter 3 investigates state estimation with independent and identical measurements in conjunction with linear reconstruction, commonly known as linear state tomography. Our main concern is informationally complete (IC) measurements constructed out of weighted 2-designs [232, 244], called tight IC measurements according to Scott [244], who proved that such measurements are optimal in minimizing the MSE averaged over unitarily equivalent states. Prominent examples of tight IC measurements include symmetric informationally complete (SIC) measurements and mutually unbiased measurements, that is, measurements constructed from mutually unbiased bases (MUB). Our primary goal is to characterize the tomographic efficiency of tight IC measurements in terms of the mean trace distance and the mean HS distance, with special emphasis on the minimal tight IC measurements, SIC measurements. Another goal is to determine the efficiency gap between product measurements and joint measurements in the bipartite and multipartite settings.

First, we introduce random-matrix theory to study the tomographic efficiency of tight IC measurements. In particular, we derive analytical formulas for the mean trace distance and the mean HS distance, which demonstrate different scaling behaviors of the two error measures with the dimension of the Hilbert space. As a byproduct, we

discovered a special class of tight IC measurements that feature exceptionally symmetric outcome statistics and low fluctuation over repeated experiments. In the case of a qubit, we compare the similarities and the differences between the SIC POM and the MUB, as well as other measurements constructed out of platonic solids. We also discuss in detail the dependence of the reconstruction error on the Bloch vector of the unknown true state and make contact with experimental data.

Second, in the bipartite and multipartite scenarios, we show that product SIC POMs are optimal among all product measurements in the same sense as joint SIC POMs among joint measurements. For a bipartite system, there is only a marginal efficiency advantage of the joint SIC POM over the product SIC POM. Hence, it is not worth the trouble to perform joint measurements. For multipartite systems, however, the efficiency advantage of the joint SIC POM increases exponentially with the number of parts.

Chapter 4 considers optimal state estimation with informationally overcomplete measurements from the perspective of frame theory. To remedy the drawbacks in linear state tomography, we determine the set of optimal reconstruction operators in the pointwise sense, using the MSE matrix as a benchmark. It turns out that the resulting reconstruction scheme is equivalent to the maximum-likelihood (ML) method in the asymptotic limit. In contrast to the traditional approaches, our approach is parametrization independent and, as a consequence, is often much easier to work with. In addition, it is rooted in frame theory and has a close connection with linear state reconstruction. These merits enable us to better understand the difference between linear state reconstruction and optimal state reconstruction.

Based on the previous framework, we prove that, among all choices of  $d+1$  projective measurements, mutually unbiased measurements are optimal in minimizing the MSE averaged over unitarily equivalent true states. This conclusion generalizes the analogous result that SIC POMs are optimal among all minimal IC measurements [244]. Incidentally, our study leads to a conjecture that singles out SIC POMs and MUB as the only solutions to a state-estimation problem.

## 1.1. Quantum state estimation

---

Furthermore, we show that covariant measurements are optimal among all nonadaptive measurements in minimizing the WMSE based on any unitarily invariant distance, including the MSE and the MSB. Informationally overcomplete measurements can improve the tomographic efficiency significantly when the states of interest have high purity. Nevertheless, the average scaled MSB diverges at the boundary of the state space in the large-sample limit. And the same is true for the WMSE based on any monotone Riemannian metric as long as the measurement is nonadaptive. This observation breaks the intuitive belief that states with high purity are easier to estimate than those with low purity. On the other hand, it motivates us to study more sophisticated estimation strategies based on adaptive measurements and collective measurements, which are the focuses of the next two chapters.

Chapter 5 considers optimal state estimation with adaptive measurements. Thanks to the two-step adaptive strategy, it remains to construct measurements that are optimal locally. Although the problem in the one-parameter setting was solved by Helstrom [139, 141] many decades ago, the one in the multiparameter setting has largely remained open up to now since the optimal measurements corresponding to different parameters are generally incompatible. About a decade ago, Gill and Massar [107] investigated the trace of the product of the Fisher information matrix and the inverse quantum Fisher information matrix, which is now known as the Gill–Massar trace (GMT), and derived a simple inequality about this quantity that is applicable to any separable measurement. This inequality succinctly summarizes the information trade-off among different parameters and may be seen as a quantitative manifestation of the complementarity principle [41]. By means of this inequality, they derived a general lower bound, the GM bound, for the WMSE, which often turns out to be much tighter than bounds known previously. Except for the two-level system, however, little is known whether the GM bound is attainable or not. This open problem is the main motivation behind the present study.

We first derive the GM inequality in a much simpler way than the original one. Explicit formulas of the GM bounds for the MSH and the MSB are also calculated.

We then introduce a new optimization paradigm for minimizing the WMSE based on any unitarily invariant distance, which reduces the optimization domain from the set of POMs to the set of Fisher information matrices. In this way, the dimension of the parameter space decreases considerably and, moreover, the nonconvexity involved in traditional optimization procedures is avoided. Furthermore, we show that the GM bound for the MSB can be saturated approximately within a factor of two by constructing an explicit measurement scheme. Our numerical calculations indicate that the GM bounds for the MSB and the MSH are nearly tight, thereby effectively solving the long-standing open problem about the tomographic efficiency of adaptive measurements with respect to the two figures of merit. In addition, adaptive measurements can improve the tomographic efficiency significantly over all nonadaptive ones.

Chapter 6 investigates the tomographic efficiencies and distinctive features of collective measurements in contrast with individual measurements. Owing to technical reasons, most previous studies on this topic presume the capability of performing arbitrary collective measurements, which is hardly accessible in practice. Our study is tailored to deal with realistic scenarios in which the experimentalist is able to perform collective measurements but only on a limited number of systems each time.

To circumvent the difficulty associated with traditional approaches, we introduce the concept of coherent measurements, which are composed of (generalized) coherent states as outcomes. Coherent measurements are a very special class of collective measurements that, in a sense, are closest to separable measurements. Surprisingly, it turns out that they are optimal or nearly optimal for many state estimation tasks. Meanwhile, they exhibit many nice features which make them an ideal starting point for studying collective measurements. We prove that the GMT of any coherent measurement on the joint state  $\rho^{\otimes N}$  of  $N$  identically prepared quantum systems is a symmetric polynomial of the eigenvalues of  $\rho$ . In addition, this polynomial is the maximum of the GMT over all possible measurements on  $\rho^{\otimes N}$  when either  $N = 2$  or  $d = 2$ . We believe that this conclusion holds in general. This polynomial succinctly summarizes the information trade-off among different parameters in the case of collective measurements on

## 1.2. Symmetric informationally complete POMs

---

$N$  identically prepared quantum systems. It has profound implications for understanding the tomographic efficiencies and distinctive features of collective measurements. It is useful not only for determining the efficiency gap between separable measurements and collective measurements but also for explicating the emergence of universality in optimal state estimation as  $N$  increases and the importance of adaption decreases.

In the case of a two-level system, we first provide a new lower bound for the WMSE that is generally much tighter than any bound known previously. We then derive the set of Fisher information matrices of all coherent measurements on  $\rho^{\otimes N}$  and the maximal GMT over all measurements on  $\rho^{\otimes N}$ . Our study settles a conjecture posed by Slater [252] more than ten years ago. Furthermore, we determine the tomographic efficiencies of coherent measurements in terms of the MSH and the MSB. It turns out that all coherent measurements are nearly optimal globally whenever  $N \geq 2$ , in sharp contrast with state estimation based on individual measurements, in which the optimal measurement heavily depends on the true state and the figure of merit.

## 1.2 Symmetric informationally complete POMs

In a  $d$ -dimensional Hilbert space, a SIC POM is composed of  $d^2$  subnormalized projectors onto pure states  $\Pi_j = |\psi_j\rangle\langle\psi_j|/d$  with equal pairwise fidelity [232, 275],

$$|\langle\psi_j|\psi_k\rangle|^2 = \frac{d\delta_{jk} + 1}{d + 1}, \quad j, k = 0, 1, \dots, d^2 - 1. \quad (1.1)$$

It is an appealing candidate for a fiducial POM owing to its high symmetry and high tomographic efficiency. Besides, SIC POMs have attracted much attention because of their connections with MUB, equiangular lines, Lie algebras, and foundational studies. All SIC POMs known so far are group covariant in the sense that each of them can be generated from a single state—the *fiducial state*—under the action of a group composed of unitary operators. Moreover, most group covariant SIC POMs are covariant with respect to the Heisenberg–Weyl (HW) group. Up to now, analytical solutions of HW covariant SIC POMs have been constructed in dimensions 2–16, 19, 24, 31, 35, 37, 43,

48; numerical solutions with high precision have been found up to dimension 67. All these results support the belief that HW covariant SIC POMs exist in any Hilbert space of finite dimension. In sharp contrast with this wealth of evidence, there is neither a general existence proof nor an efficient way for constructing SIC POMs. What is worse, many basic properties of SIC POMs have remained elusive. The implication of the equiangular condition is largely a mystery, although it looks so simple. Actually, SIC POMs in dimension 3 already exhibit a plethora of surprises.

Part II of this thesis (Chapters 7 to 10) explores the structure of SIC POMs with a special emphasis on the symmetry problem: What symmetry can a SIC POM possess? and the equivalence problem: How can we determine whether two SIC POMs are equivalent or not. In this way, we hope to establish a clear picture about known SIC POMs and shed some light on those SIC POMs yet to be discovered.

Chapter 7 introduces some preliminary concepts followed by several new results. We first derive a necessary condition on the groups that can generate SIC POMs based on the works of Zauner [275] and Grassl [119], which signifies the crucial role of nice error bases in the study of SIC POMs. We then establish a simple criterion for determining equivalence relations among SIC POMs that are covariant with respect to the same group. Finally, we review the basic properties of the HW group and the Clifford group. For the convenience of later discussions, some supplementary materials concerning the Clifford group are presented in Appendix H.

Chapter 8 settles several persistent open problems about group covariant SIC POMs in prime dimensions. We prove that, in any prime dimension not equal to 3, each group covariant SIC POM is covariant with respect to a unique HW group; its symmetry group is a subgroup of the Clifford group. Hence, SIC POMs on different orbits are not equivalent. In dimension 3, each group covariant SIC POM may be covariant with respect to three or nine HW groups; its symmetry group is a subgroup of at least one of the Clifford groups associated with these HW groups, respectively. There may exist two or three orbits of equivalent SIC POMs depending on the order of the symmetry group. In addition, we establish complete equivalence relations among group covariant

## 1.2. Symmetric informationally complete POMs

---

SIC POMs in dimension 3 and classify inequivalent ones according to the geometric phases associated with fiducial states.

Finally, we briefly discuss the situation beyond prime dimensions. In particular, we prove that two HW covariant SIC POMs in any prime-power dimension not equal to 3 are unitarily or antiunitarily equivalent if and only if they are on the same orbit of the extended Clifford group. In addition, the set of Hoggar lines is not covariant with respect to the usual HW group, in agreement with a long-standing speculation.

Chapter 9 focuses on HW covariant SIC POMs in the four-dimensional Hilbert space, which exhibit remarkable additional symmetry beyond what is reflected in the name<sup>1</sup>. It is known that there exists a single orbit of 256 fiducial states, constituting 16 SIC POMs [10, 232, 245]. We characterize these fiducial states and SIC POMs by examining the symmetry transformations within a given SIC POM and among different SIC POMs. The symmetry group of each SIC POM is shown to be a subgroup of the Clifford group, thereby extending previous results on prime dimensions. Furthermore, we find 16 additional SIC POMs by a suitable regrouping of the 256 fiducial states, and show that they are unitarily equivalent to the 16 original SIC POMs. We also determine all similar regrouping phenomena on the orbits of SIC POMs cataloged by Scott and Grassl [245] and provide a unified explanation of these phenomena based on a peculiar structure of the Clifford group and its normalizer explicated in Appendix H.2.

We then reveal additional structure of these SIC POMs when the four-dimensional Hilbert space is perceived as the tensor product of two qubit Hilbert spaces. A concise representation of the fiducial states is introduced in terms of generalized Bloch vectors, which allows us to explore the intriguing symmetry of the two-qubit SIC POMs. In particular, when either the standard product basis or the Bell basis is chosen as the defining basis of the HW group, in eight of the 16 HW covariant SIC POMs, all the fiducial states have the same concurrence of  $\sqrt{2/5}$ . These SIC POMs are particularly appealing for an experimental implementation, because all fiducial states can be turned into each other with just local unitary transformations.

---

<sup>1</sup>This work represents a collaboration with Teo Yong Siah and Berthold-Georg Englert [283].

Chapter 10 starts a graph-theoretic approach to the symmetry and the equivalence problems of SIC POMs. We establish a simple connection between the symmetry problem of a SIC POM and the automorphism problem of a graph constructed out of the triple products among the states in the SIC POM. Based on this connection, we develop an efficient algorithm for determining the symmetry group of the SIC POM, which is much faster than any algorithm known before. A variant of the algorithm allows solving the SIC POM equivalence problem, which can be reduced to the graph isomorphism problem. In addition to its applications to practical calculations, the graph-theoretic approach also provides a fresh perspective for understanding SIC POMs, which complements the group-theoretic approach explored previously.

As an application of the graph-theoretic approach, we determine the symmetry groups of all SIC POMs known in the literature and establish complete equivalence relations among them. We also figure out all nice error bases contained in the symmetry groups of these SIC POMs. It turns out that, except in dimension 3, the (extended) symmetry group of any known HW covariant SIC POM is a subgroup of the (extended) Clifford group and contains only one HW group, in agreement with a long-standing conjecture. As a consequence, two such SIC POMs are unitarily or antiunitarily equivalent if and only if they are on the same orbit of the extended Clifford group. Furthermore, our study indicates that all SIC POMs known so far are covariant with respect to the HW groups, except for the set of Hoggar lines, which is covariant with respect to the three-qubit Pauli group.

As a caveat, we emphasize that Part II of the thesis may reuse some symbols used in Part I that have completely different meanings. In addition, to simplify the notation, the indices of basis elements of the Hilbert space are chosen to run from 1 to  $d$  in Part I of the thesis, but from 0 to  $d - 1$  in Part II.



# Quantum state estimation

---

## 2.1 Introduction

The development of quantum estimation theory has followed two different lines of thinking. The first line is mainly concerned with reliable and efficient state estimation in practice; see Refs. [176, 186, 208] for an overview. It was initiated in the late 1950s by Fano [92], inspired by the question: How can we determine the state of a quantum system from observable quantities? The benchmark was the introduction of the concept of a quorum, a complete set of observables that uniquely determines the state of a quantum system, which may be seen as the precursor of the concept of informational completeness [52, 223]. The second line is mainly concerned with the optimal strategies and optimal efficiency allowed by quantum mechanics; see Refs. [133, 141, 147] for an overview. It was initiated in the late 1960s by Helstrom [139, 141], inspired by the question: What is the minimal MSE in estimating certain parameter that characterizes the quantum state? The benchmark was the introduction of quantum analogs of the Fisher information and the CR bound based on the symmetric logarithmic derivative (SLD), which enabled solving the optimization problem in the one-parameter setting. Both lines of thinking have proved to be very useful in the development of quantum estimation theory. Unfortunately, they have run almost independently for many decades, and the lack of communication between the two communities has remained a source of many confusions. Recently, there appeared a trend of convergence of the two approaches, especially in the study of quantum metrology [108, 109, 110]. As the requirement for precision measurements increases, the integration of the two approaches is due to play an increasingly important role.

In this chapter, we first present a historical survey of the development of quantum estimation theory and quantum state estimation in particular. We then introduce several basic elements in the field of quantum state estimation, such as quantum states, measurements, state reconstruction, Fisher information, and CR bound.

## 2.2 Historical background

The idea of determining the state of a quantum system from measurements can be traced back to Pauli when he asked whether the position distribution and momentum distribution suffice to determine the wave function of a quantum system [211]. However, a systematic study was not initiated until the 1950s when Fano introduced the concept of a quorum [92]. Following Fano's work, state determination for spin systems was studied by Gale, Guth, and Trammell [106], as well as Newton and Young [205]; more general settings were investigated by Band and Park [23, 24, 25, 209], who considered one-dimensional spinless particle in addition to spin systems. Later, Ivanović [155] explored the state estimation problem from a geometric perspective, with a special emphasis on mutually unbiased measurements, an idea first conceived by Schwinger [243]. He also constructed a complete set of mutually unbiased measurements when the dimension is a prime, followed by a generalization to prime-power dimensions by Wootters and Fields [272]. Based on the concept of mutually unbiased measurements, Wootters [269, 270] introduced a formulation of quantum mechanics in terms of probabilities instead of probability amplitudes and generalized the Wigner functions to systems with discrete degrees of freedom. Meanwhile, tomographic approaches to the traditional Wigner functions were initiated by Bertrand and Bertrand [36], as well as Vogel and Risken [260] (see also the works of Royer [236, 237]), who showed that Wigner functions can be reconstructed from probability distributions for the rotated quadrature operators by means of the inverse Radon transform. Density operators can then be determined based on their correspondence with Wigner functions. A more efficient reconstruction method that is based on pattern functions was later developed by D'Ariano et al. [72] and Leonhardt et al. [176, 177, 178].

## 2.2. Historical background

---

Inspired by the observation of Vogel and Risken [260], Smithey et al. [253] performed the first measurements of the quadrature probability distributions of an optical mode based on optical homodyne detection [273], and reconstructed the Wigner function and the density operator, which marked the birth of optical homodyne tomography [186, 208]. Following their experiment, states of many other quantum systems were also characterized, such as the vibrational state of a diatomic molecule [82], the motional state of a trapped ion [174], the state of an ensemble of helium atoms [171], and entangled states of polarized photon pairs [156, 267]. See Refs. [186, 208] for an overview about experimental progress in quantum state estimation.

The advance of experimental techniques and the emergence of quantum information science further stimulated the development of quantum estimation theory. Traditional tomographic schemes, such as linear inversion, which are suitable for the proof of principle, often could not meet practical requirements. Thus, great efforts were directed to search for reliable and efficient alternatives. The problem of reconstructing quantum states from informationally incomplete measurements was addressed in the middle 1990s by Bužek et al. [53, 54, 55], who proposed a method for selecting the most objective estimator by means of Jaynes principle of maximum entropy (ME) [157, 158]. Meanwhile, ML estimation (MLE) was advocated by Hradil [152], who developed an efficient algorithm for computing the ML estimator, which avoids the problems of non-positivity and choice ambiguity associated with linear estimators. Recently, as an alternative to MLE, hedged maximum-likelihood estimation (HMLE) was proposed by Blume-Kohout [38] to eliminate the zero-eigenvalue problem, which is not desirable for predicative tasks. Based on the ML and ME principles, Teo et al. [256] developed a general procedure for selecting the most-likely state with the largest entropy, which enables us to obtain a unique and objective estimator even from noisy data of informationally incomplete measurements. Out of a different vein, Gross et al. [123] proposed a tomographic method based on compressed sensing [57, 58, 59, 60, 79], which can improve the efficiency significantly, provided that the states of interest have high purities.

In contrast to full tomography, direct estimation of certain quantities of interest is generally more efficient and has thus received increasing attention in the past decade. Prominent examples include direct estimation of linear or nonlinear functional, such as the purity of density operators [87]; direct detection and characterization of quantum entanglement [44, 149]; entanglement verification based on the likelihood ratio test [40]; and direct fidelity estimation from Pauli measurements [96].

As an extension to quantum state tomography, quantum process tomography (QPT) focuses on characterizing unknown quantum processes or dynamics instead of quantum states, which is crucial to ensuring the performance of many quantum information processing protocols. Its development has drawn much inspiration from quantum state tomography. Standard QPT (SQPT) was introduced by Poyatos, Cirac, and Zoller [221], as well as by Chuang and Nielsen [66] in the late 1990s. To characterize a quantum process, a set of reference states is prepared and then reconstructed by quantum state tomography after subjecting them to a given quantum process, which can then be determined if the set of reference states spans the operator space. SQPT has been applied to characterize the control-not gate [65, 207] and Bell-state filters [197]. As an alternative to SQPT, ancilla assisted QPT (AAQPT) was proposed by Leung [179, 180], as well as by D'Ariano and Presti [70], followed by experimental realizations [4, 189]. By introducing an ancilla system, it requires only one preparation and tomography of the reference state. Later, an algorithm for direct characterization of quantum dynamics (DCQD) was developed by Mohseni and Lidar [198, 199] and applied to determine the dynamics of a photon qubit [262] and that of nuclear spins in the solid state [89]. In contrast with the previous two methods, DCQD does not need quantum state tomography, but relies on error-detection techniques. It is especially suitable when one is interested in a few parameters rather than full information about a quantum process, in which case it can reduce the number of necessary experimental configurations significantly. A survey on the three alternative strategies was presented in Ref. [200].

A central problem in quantum estimation theory is to determine the optimal strat-

## 2.2. Historical background

---

egy for estimating the parameters that characterize a quantum system. This problem was first addressed in the 1960s by Helstrom [139, 140, 141, 142], who derived the quantum CR bound based on the SLD Fisher information matrix and solved the optimization problem in the one-parameter setting, in which case the bound is tight. Incidentally, the optimal strategy can be realized with only individual measurements. The situation in the multiparameter setting turned out to be much more involved; nevertheless, breakthroughs were made in a few special yet important cases. The problem of estimating the complex amplitude of coherent signal in Gaussian noise was solved by Yuen and Lax [274] by means of another quantum analog of the CR bound based on the right logarithmic derivative (RLD), which is often tighter than the SLD bound in the multiparameter setting. Based on a similar approach, Holevo [147] solved the estimation problem about the mean value of Gaussian states. He also introduced a new quantum CR bound, known as Holevo bound, which is tighter than both the SLD bound and the RLD bound. However, this bound is generally not easy to calculate since the definition itself involves a tough optimization procedure. The main achievements of the pioneering works in the 1960s and 1970s are summarized in the books of Helstrom [141] and of Holevo [147].

In the late 1980s, the development of quantum estimation theory was revitalized after a short period of slowdown, as witnessed by the introduction of several quantum CR bounds that are applicable to separable measurements and are usually much tighter than those bounds known previously; see Ref. [133] for more details. Nagaoka introduced the concept of the most informative or attainable CR bound and studied its general properties [201]. He also introduced a new CR bound in the two-parameter setting based on an inequality concerning simultaneous approximate measurements of noncommuting observables [203], and showed that it is tight for the two-level system. Using the duality theorem in linear programming, Hayashi [129] generalized the result of Nagaoka and derived the attainable CR bound for any family of states describing the two-level system. Later, Gill and Massar [107] introduced a novel approach that naturally incorporates the information trade-off among different parameters. Based on

this approach, they derived a general lower bound for the WMSE that is applicable to any separable measurement on a  $d$ -level system. This bound is tight for the two-level system [107], in agreement with the analysis of Hayashi [129]. In general, however, little is known whether it is attainable or not.

Since the late 1990s, significant progress has been achieved in quantum state estimation with collective measurements in the asymptotic setting. Hayashi studied the estimation problem of the displaced thermal states and showed that the RLD bound for the MSE can be saturated with collective measurements [131]. He also applied quantum central-limit theorem [111, 215] to studying quantum state estimation and demonstrated that the Holevo bound [147] can be saturated asymptotically [135]. Based on this idea, optimal state determination for the two-level system was later analyzed in detail by Hayashi and Matsumoto [137]. Recently, another breakthrough was made by Kahn and Guță et al. [125, 126, 160], who demonstrated local asymptotic normality for finite-dimensional quantum systems, which states that a quantum statistical model consisting of an ensemble of identically prepared systems can be approximated by a statistical model consisting of classical and quantum Gaussian variables in the asymptotic limit. This observation is crucial to devising optimal estimation strategies in the asymptotic setting. Incidentally, above studies presume the capability of collective measurements on arbitrary number of identically prepared quantum systems, which is hardly accessible in practice. A major open problem is to determine the optimal estimation strategies and the corresponding tomographic efficiencies in case of limited access to collective measurements.

Quantum statistical models consisting of pure states exhibit many distinctive features. Since the density operators are not invertible, the SLD and RLD bounds are not well defined, and many traditional methods do not apply. Surprisingly, it turned out that the problem was actually more amenable compared with the problem in mixed-state setting thanks to the simplification brought by the new features [133]. Systematic studies of pure-state models were initiated in the middle 1990s by Fujiwara and Nagaoka [103, 104, 105], who derived the obtainable CR bounds for a one-dimensional

## 2.2. Historical background

---

model and a two-dimensional coherent model. Later, Matsumoto [193] introduced a powerful approach and derived the obtainable CR bounds for a wide range of pure-state models. According to his study, the use of quantum correlations cannot improve these bounds for pure-state models, in sharp contrast with mixed-state models. It should be noted that this conclusion is applicable only to asymptotic state estimation: In the finite-sample scenario, quantum correlations are useful even for pure-state models, as we shall see in the next paragraph.

As an alternative to the CR approach, the Bayesian approach got momentum in the 1990s, thereby yielding fruitful results. With this approach, it is generally easier to determine the optimal measurements in the case of finite samples. Coincidentally, the study of optimal state estimation was interlaced with that of optimal quantum cloning [241]. In both fields, most of the priors considered were unitarily invariant, and the mean fidelity was the most popular figure of merit. The optimal measurements in estimating qubit pure states were first derived by Massar and Popescu [191], who also proved that the optimal strategy cannot be realized by individual measurements. Their study showed that collective measurements on an ensemble as a whole can provide more information than individual measurements, thereby confirming a conjecture posed by Peres and Wootters [214]. Later, a universal algorithm for constructing the optimal measurements for estimating pure states in more general settings was developed by Derka, Bužek, and Ekert [78]. Meanwhile, Bruß, Ekert, and Macchiavello [48] demonstrated the equivalence between optimal state estimation and asymptotic cloning. Based on this connection and a result on optimal cloning derived by Werner [265], Bruß and Macchiavello determined the optimal measurements for estimating pure states of a  $d$ -level system. A direct derivation of their result was later proposed by Hayashi, Hashimoto, and Horibe [128]. The optimal strategy for estimating qubit mixed states was first derived by Vidal et al. [259] (see also Ref. [254]) based on a special formula for the fidelity, which has no analog in higher dimensions. Detailed comparison between separable measurements and collective measurements in qubit state estimation was later presented by Bagan et al. [18, 19, 20]. The problem of estimating mixed

states in higher dimensions is largely open.

## 2.3 Quantum states and measurements

### 2.3.1 Simple systems

The state of a quantum system encodes all information about the quantum system and determines the statistics of all potential measurements on it. Mathematically, a pure state is generally represented by a normalized ket often labeled by  $|\psi\rangle$ . According to the basic postulates of quantum mechanics, any superposition of kets also represents a legitimate state; all unnormalized kets form a vector space, known as the Hilbert space. Since kets that are proportional to each other represent the same state, there is a one-to-one correspondence between the rays in the Hilbert space and the pure states. In general, the state of a quantum system can be represented by a positive semidefinite matrix of unit trace, known as the density matrix or density operator and often denoted by  $\rho$ . Density operators of rank one represent pure states, whereas those of higher ranks represent mixed states. In practice, the state is usually determined by the preparation procedure, which may be characterized by one or more parameters. For example, the alignment of the polarizer determines the polarization state of the photon after passing through it.

A *generalized measurement* [206] is described by a set of measurement operators  $M_\xi$  corresponding to a set of measurement outcomes that satisfy the completeness condition

$$\sum_{\xi} M_{\xi}^{\dagger} M_{\xi} = 1. \quad (2.1)$$

Given a quantum system on the state  $\rho$  before the measurement, the probability  $p_{\xi}$  of obtaining outcome  $\xi$  is given by the Born rule

$$p_{\xi} = \text{tr}(M_{\xi} \rho M_{\xi}^{\dagger}). \quad (2.2)$$

As a consequence of the completeness condition, these probabilities are normalized; that



### 2.3. Quantum states and measurements

---

is,  $\sum_{\xi} p_{\xi} = 1$ . If outcome  $\xi$  occurs, then the quantum system after the measurement is described by the state operator

$$\frac{M_{\xi} \rho M_{\xi}^{\dagger}}{\text{tr}(M_{\xi} \rho M_{\xi}^{\dagger})}. \quad (2.3)$$

A measurement is a *projective* or von Neumann measurement if the measurement operators  $M_{\xi}$  are orthogonal projectors. In that case, there exists an observable with  $M_{\xi}$  as eigenprojectors. For example, in the case of the qubit, the projective measurement composed of the two measurement operators  $|1\rangle\langle 1|$  and  $|2\rangle\langle 2|$  is equivalent to the measurement of the spin of a particle along the  $z$  direction, as realized in the Stern-Gerlach experiment. A projective measurement is repetitive in the sense that repeated measurements yield the same outcome as the first one and thus provide no additional information about the original quantum system.

If we are interested only in the outcome statistics but not the state after the measurement, then the measurement can be effectively described by the set of positive operators  $\Pi_{\xi} = M_{\xi}^{\dagger} M_{\xi}$ , which sum up to the identity. In that case, the measurement may be referred to as a *probability operator measurement* (POM)<sup>1</sup>, and the set of operators  $\Pi_{\xi}$  may be identified with the outcomes of the measurement. According to Neumark's dilation theorem [213], any POM can be realized as a projective measurement on a larger system. The merit of the POM formalism lies in allowing us to focus on the system under study, without worrying about the detailed realization of the measurement. Besides, POMs are generally easier to handle than projective measurements thanks to their nicer mathematical structure. For example, any convex combination of POMs is still a POM. This observation is crucial to constructing sophisticated POMs from simple ones.

A measurement is *informationally complete* (IC) if every state is completely determined by the outcome statistics [52, 223] or, equivalently, if the outcomes of the measurement span the space of Hermitian operators. In a  $d$ -dimensional Hilbert space, an

---

<sup>1</sup>Also known as positive-operator-valued measure (POVM) in the mathematical community.

IC measurement consists of at least  $d^2$  outcomes, whereas a *minimal* IC measurement consists of no more than  $d^2$  outcomes. An informationally overcomplete measurement is an IC measurement with more than  $d^2$  outcomes. Is there any advantage in choosing informationally overcomplete measurements? This question will be investigated in Chapter 4.

### 2.3.2 Composite systems

Compared with simple systems, a distinctive feature of composite systems is the appearance of quantum correlations known as quantum entanglement (see Ref. [150] for a review), as emphasized by Einstein [85] and Schrödinger [242]. Quantum entanglement is not only a characteristic feature of quantum physics, but also a crucial resource for many information processing tasks [150], such as quantum teleportation [31], superdense coding [33], quantum key distribution [86], and quantum computation [159, 225]. Its connection with quantum state estimation can be elaborated in two aspects. On the one hand, tomographic techniques provide basic means of detecting, quantifying, and characterizing entanglement [40, 43, 64, 148, 150, 282]. On the other hand, entanglement is a basic ingredient for many collective measurements [133, 191], the most general measurements allowed by quantum mechanics. This latter aspect is the main focus of the present discussion.

Consider a bipartite composite system as an example. Suppose the Hilbert spaces of Alice and Bob are  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively, then the Hilbert space  $\mathcal{H}$  of the whole system is the tensor product  $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ . A pure state  $|\Psi\rangle \in \mathcal{H}$  is *separable* if it is a tensor product of two states in their respective Hilbert spaces; otherwise, it is *entangled*. Alternatively, a pure state is separable if and only if each reduced state is pure. A mixed state  $\rho$  is separable if it can be written as a convex combination of separable pure states and is entangled otherwise [264]. Similar concepts also apply to a system composed of more than two parties [150].

A measurement on a composite system is *collective* if it cannot be decomposed into individual measurements on the constituent subsystems. Sometimes collective

## 2.4. Quantum state reconstruction

---

measurements may also refer to all possible measurements on a composite system.

Analogous to a quantum state, a measurement is *separable* if each outcome is a convex combination of tensor products of positive operators<sup>2</sup> or, equivalently, if each outcome corresponds to a separable state, which is not necessarily normalized. A simple example of separable measurements are *product measurements*, which can be decomposed into independent measurements on the constituent subsystems. In the bipartite scenario, suppose  $\Pi_{\xi_1}$  and  $\Pi_{\xi_2}$  are the outcomes of the measurements on the two subsystems, respectively. Then the outcomes of the product measurement are given by  $\Pi_{\xi_1\xi_2} = \Pi_{\xi_1} \otimes \Pi_{\xi_2}$ .

In a more sophisticated scenario, Alice and Bob may perform local measurements and tell each other the outcomes of their measurements through classical communication. Conditioned on these outcomes, they may perform further local measurements, and so forth. Obviously, such measurements are separable since they can be realized by local operations and classical communication (LOCC); the converse, however, is not true in general [32].

A measurement is *entangled* if it is not separable. A simple example of entangled measurements in the two-qubit setting is the Bell measurement. In practice, it is generally much harder to realize entangled (collective) measurements than separable (individual) measurements. A major open question in quantum estimation theory is by how much can the efficiency be increased with entangled (collective) measurements compared with separable (individual) measurements. Besides practical interest, this question is also of paramount importance in understanding the difference between quantum information processing and classical information processing.

## 2.4 Quantum state reconstruction

Quantum state reconstruction is a procedure for inferring the state of a quantum system from measurement results. It has a close analog in classical statistical inference.

---

<sup>2</sup>Some authors define separable measurements as those measurements that can be realized by LOCC [107]; as for our definition, not all separable measurements can be realized by LOCC.

Accordingly, most reconstruction methods have classical counterparts. However, to devise a good quantum estimation strategy, it is indispensable to take into account additional requirements pertinent to quantum systems, such as the positivity constraint. In addition, the choice may also depend on the system under consideration and the applications in mind. In this section, we review several mainstream reconstruction methods investigated in the literature, with brief comments on the pros and the cons of each method; see also [186, 208].

### 2.4.1 Linear state reconstruction

Linear state reconstruction is one of the simplest reconstruction methods; it was first conceived by Fano [92], followed by many other researchers [23, 24, 25, 106, 205, 209, 244]. Suppose we are given  $N$  identically prepared quantum systems, each in the state  $\rho$ , and perform  $N$  identical and independent measurements described by the POM  $\sum_{\xi} \Pi_{\xi} = 1$ . If the outcome  $\xi$  occurs  $n_{\xi}$  times after the measurements, then the frequency of the outcome  $\xi$  is  $f_{\xi} = n_{\xi}/N$ . In linear state reconstruction, we search for an estimator  $\hat{\rho}$  that matches the observed frequencies, that is,

$$\text{tr}(\hat{\rho}\Pi_{\xi}) = f_{\xi} \quad \text{for all } \xi. \quad (2.4)$$

Incidentally, linear state reconstruction is sometimes called linear inversion. If the measurement is IC, then there is at most one solution. If in addition the measurement is minimal, then there exists a unique solution. In that case, the outcomes  $\Pi_{\xi}$  form a basis in the operator space, and there exists a unique dual basis composed of Hermitian operators  $\Theta_{\xi}$  such that  $\text{tr}(\Pi_{\xi}\Theta_{\zeta}) = \delta_{\xi\zeta}$ . Once the dual basis is known, the estimator can be computed immediately using the formula  $\hat{\rho} = \sum_{\xi} f_{\xi}\Theta_{\xi}$ . Therefore, the  $\Theta_{\xi}$ s are also known as *reconstruction operators*. For a generic IC measurement, the system of equations in Eq. (2.4) can become incompatible because of the statistical noise associated with the frequencies, and there is generally no estimator that is compatible with the frequencies. It is still possible to find a set of reconstruction operators as before, but the choice is no longer unique.

## 2.4. Quantum state reconstruction

---

The main merit of linear state reconstruction is its simplicity. It is a good starting point in theoretical analysis, but not a good choice in practice owing to several defects: First, the estimator is sometimes not positive semidefinite and thus does not represent a legitimate quantum state. This happens quite often if the true state has a high purity and the sample size is small. Second, there is some arbitrariness in the choice of the reconstruction operators when the measurement is informationally overcomplete, and the information encoded in the measurement statistics cannot be extracted efficiently if the reconstruction operators are chosen a priori. An ad hoc method for solving the first problem is to mix the estimator with some noise (the completely mixed state for instance) until it is positive semidefinite. To solve the second problem, we need to choose the reconstruction operators adaptively according to the measurement results. An alternative recipe that can circumvent the two problems simultaneously is MLE, which is the subject matter of the next section.

### 2.4.2 Maximum-likelihood estimation

In MLE, instead of searching for the state that matches the observed frequencies, we seek the state that maximizes the likelihood function. The principle of ML was proposed by R. A. Fisher [93] in the 1920s and has become a basic ingredient in statistical inference. During the past decade, it has found extensive applications in quantum state estimation [152, 186, 208, 228, 229]. In addition, it is useful for entanglement detection [40] and characterization [64].

Following the previous notation, the *likelihood functional* is defined as [152, 208]

$$\mathcal{L}(\rho) = \prod_{\xi} p_{\xi}^{n_{\xi}}, \quad (2.5)$$

where  $p_{\xi} = \text{tr}(\rho \Pi_{\xi})$  is the probability of obtaining the outcome  $\xi$  given the true state  $\rho$ . In practice, it is often more convenient to work with the *log-likelihood functional*

$$\ln \mathcal{L}(\rho) = \sum_{\xi} n_{\xi} \ln p_{\xi} = N \sum_{\xi} f_{\xi} \ln p_{\xi}. \quad (2.6)$$

MLE consists in choosing a state  $\hat{\rho}_{\text{ML}}$  that maximizes the likelihood functional or, equivalently, the log-likelihood functional, as an estimator of the true state [152, 186, 208, 228, 229]. If there exists a state that matches the observed frequencies in the sense of satisfying Eq. (2.4), then the state is also an ML estimator. This conclusion is an immediate consequence of the inequality

$$\sum_{\xi} f_{\xi} \ln p_{\xi} \leq \sum_{\xi} f_{\xi} \ln f_{\xi}. \quad (2.7)$$

In general, it is not easy to find a closed formula for the ML estimator. Fortunately, the estimator can be computed efficiently with an algorithm proposed by Hradil [152]. Since the log-likelihood functional is concave and the state space is convex, the search for the ML estimator can be turned into a convex optimization problem, which can be solved based on the idea of steepest ascent. Starting from an initial guess, say,  $\rho^{(k)} = 1/d$  with  $k = 0$ , we can obtain the ML estimator by implementing the following successive iterations [152, 208, 255]:

1. Compute the operator

$$R_k = \sum_{\xi} \frac{f_{\xi} \Pi_{\xi}}{\text{tr}(\hat{\rho}^{(k)} \Pi_{\xi})}. \quad (2.8)$$

2. Choose a small parameter  $\epsilon_k$  and update the estimator,

$$\hat{\rho}^{(k+1)} = \frac{(1 + \epsilon_k R_k) \hat{\rho}^{(k)} (1 + \epsilon_k R_k)}{\text{tr}\{(1 + \epsilon_k R_k) \hat{\rho}^{(k)} (1 + \epsilon_k R_k)\}}. \quad (2.9)$$

3. Stop the iteration if the trace distance between  $\hat{\rho}^{(k+1)}$  and  $\hat{\rho}^{(k)}$  is smaller than a given threshold; otherwise, replace  $k$  with  $k + 1$  and repeat the above steps.

The parameter  $\epsilon_k$  can be chosen a priori; for example, the choice  $\epsilon_k = 0.5$  works quite well when  $d$  is small. In general, a suitable line-optimization procedure can help speed up the algorithm.

The ML estimator is unique if the measurement is IC; otherwise, the estimator is generally not unique, and there exists a plateau in the contour of the likelihood functional. Recently, a nice solution to this problem was proposed by Teo et al. [256] based

## 2.4. Quantum state reconstruction

---

on the ML principle [93] and the ME principle [157, 158]. They developed an efficient algorithm for computing the most objective estimator—the state with the highest von Neumann entropy among all the states that maximize the likelihood functional.

As one of the most popular estimators used in practice, the ML estimator has many nice features: It is always positive semi-definite and thus represents a legitimate quantum state; it is asymptotically unbiased; it is asymptotically efficient in the sense of saturating the CR bound in the large-sample limit [173]; it can be computed efficiently with a simple algorithm [152]. The main drawback of the ML estimator is the zero-eigenvalue problem: The estimator is often rank deficient when the true state has a high purity. These zero eigenvalues represent unrealistic confidence over the outcomes of certain potential measurements, which is undesirable for applications such as data compression, betting, and cryptography [39].

### 2.4.3 Other reconstruction methods

Over the past few years, several alternatives to MLE have been proposed: Prominent examples are Bayesian mean estimation [39] and hedged maximum-likelihood estimation [38]. Meanwhile, several methods have been developed to deal with large quantum systems, such as compressed sensing [123, 249] and direct fidelity estimation [96]. Here we shall briefly discuss the first two methods.

#### 2.4.3.1 Bayesian mean estimation

In Bayesian mean estimation (BME) [39], we choose a prior  $p_0(\rho)$  over the state space and derive the posterior distribution  $p_f(\rho)$  by normalizing the product of the prior and the likelihood functional, that is,  $p_f(\rho) \propto p_0(\rho)\mathcal{L}(\rho)$ . The Bayesian mean estimator is the average over the posterior,

$$\hat{\rho}_{\text{BM}} = \int d\mu(\rho)p_f(\rho)\rho. \quad (2.10)$$

Common choices for the prior include the uniform distribution with respect to the HS measure and the one with respect to the Bures measure [30, 50]. With a suitable

choice of the prior, BME can avoid the zero-eigenvalue problem and is thus more appealing than MLE if the estimator is to be used for predictive tasks such as betting or data compression. In addition, BME often outperforms MLE when the sample size is small [39]. There are two major problems with BME. One problem is the ambiguity in the choice of the prior: There is no universal criterion for selecting the prior. While some natural restrictions can be imposed on the prior based on symmetry consideration, unitary invariance for instance, such restrictions generally cannot fix a unique prior. In the case of the qubit, for example, there is no consensus on the radial distribution of the prior over the Bloch ball. Another serious problem is the difficulty in computing the estimator even numerically since the computation involves a high-dimensional integral. There is still no reliable and efficient algorithm for this purpose; Monte Carlo methods have been proposed to attack this problem [39].

#### 2.4.3.2 Hedged maximum-likelihood estimation

Hedged maximum-likelihood estimation (HMLE) was proposed by Blume-Kohout [39] as an alternative to MLE and was tailored to solve the zero-eigenvalue problem. It generalizes an idea in classical statistical inference known as the “add  $\beta$ ” rule, which was proposed by Lidstone [182] in the 1920s. In HMLE, the likelihood functional  $\mathcal{L}(\rho)$  is multiplied by a *hedging functional* [38]

$$h(\rho) = \det(\rho)^\beta, \tag{2.11}$$

where the hedging parameter  $\beta$  usually assumes a value between 0 and 1. The maximum of the functional  $\mathcal{L}(\rho)h(\rho)$  defines the estimator, which is guaranteed to have full rank. Since  $\ln \det(\rho)$  is concave in  $\rho$ , the functional  $\ln[\mathcal{L}(\rho)h(\rho)]$  is also concave. Therefore, the estimator can be computed efficiently with a similar algorithm as in MLE. These two attractive features make HMLE an appealing alternative to MLE and BME. A major problem with HMLE is that there is no universal criterion for choosing the hedging functional, which may depend on both the prior knowledge available and the figure of merit adopted. Quite often the choice is made on an ad hoc basis.



## 2.5 Fisher information and Cramér–Rao bound

The Fisher information [94] and the CR bound [68, 224] are two basic ingredients in statistical inference: The former quantifies the amount of information yielded by an observation or a measurement concerning certain parameters of interest, while the latter quantifies the minimal error that goes with the inference of these parameters.

Consider a family of probability distributions  $p(\xi|\theta)$  parameterized by  $\theta$ . Our task is to estimate the value of  $\theta$  as accurately as possible based on the measurement outcomes. Given an outcome  $\xi$ , the function  $p(\xi|\theta)$  of  $\theta$  is called the likelihood function. The *score* is defined as the partial derivative of the log-likelihood function with respect to  $\theta$  and reflects the sensitivity of the log-likelihood function with respect to the variation of  $\theta$ . Its first moment vanishes; its second moment is known as the *Fisher information* [94, 173] and is given by

$$I(\theta) = \text{Var}\left(\frac{\partial \ln p(\xi|\theta)}{\partial \theta}\right) = \sum_{\xi} p(\xi|\theta) \left(\frac{\partial \ln p(\xi|\theta)}{\partial \theta}\right)^2 = \sum_{\xi} \frac{1}{p(\xi|\theta)} \left(\frac{\partial p(\xi|\theta)}{\partial \theta}\right)^2. \quad (2.12)$$

The Fisher information represents the average sensitivity of the log-likelihood function with respect to the variation of  $\theta$ . Intuitively, the larger the Fisher information, the better we can estimate the value of the parameter  $\theta$ .

An estimator  $\hat{\theta}(\xi)$  of the parameter  $\theta$  is *unbiased* if its expectation value is equal to the true parameter; that is,

$$\sum_{\xi} p(\xi|\theta) [\hat{\theta}(\xi) - \theta] = 0. \quad (2.13)$$

Taking the derivative with respect to  $\theta$  and applying the Cauchy–Schwarz inequality (using the fact that  $\sum_{\xi} p(\xi|\theta) = 1$ ), we obtain the well-known *CR bound*

$$\text{Var}(\hat{\theta}) \geq \frac{1}{I(\theta)}, \quad (2.14)$$

which states that the MSE of any unbiased estimator is bounded from below by the inverse of the Fisher information [68, 224].

In the multiparameter setting, the Fisher information takes on a matrix form,

$$I_{jk}(\theta) = \mathbb{E} \left[ \left( \frac{\partial \ln p(\xi|\theta)}{\partial \theta_j} \right) \left( \frac{\partial \ln p(\xi|\theta)}{\partial \theta_k} \right) \right]. \quad (2.15)$$

and the CR bound for any unbiased estimator turns out to be a matrix inequality,

$$C(\theta) \geq I^{-1}(\theta), \quad (2.16)$$

where  $C(\theta)$  is the MSE matrix (also known as the covariance matrix),

$$C_{jk}(\theta) = \mathbb{E}[(\hat{\theta}_j - \theta_j)(\hat{\theta}_k - \theta_k)]. \quad (2.17)$$

Since the likelihood function is multiplicative, the Fisher information matrix is additive; that is, the total Fisher information matrix of several independent measurements is equal to the sum of the respective Fisher information matrices of individual measurements. In particular, the Fisher information matrix of  $N$  identical and independent measurements is  $N$  times that of one measurement. Accordingly, the MSE matrix of any unbiased estimator based on  $N$  measurements satisfies the inequality  $C^N(\theta) \geq 1/N I(\theta)$ . Thanks to Fisher's theorem [93, 94], the lower bound can be saturated asymptotically with the ML estimator. In the large-sample scenario, the *scaled MSE matrix*  $N C^N(\theta)$  is generally independent of the sample size. It is also denoted by  $C(\theta)$  when there is no confusion.

In quantum state estimation, we are interested in the parameters that characterize the state  $\rho(\theta)$  of a quantum system. To estimate the values of these parameters, we may perform generalized measurements. Given a measurement  $\Pi$  with outcomes  $\Pi_\xi$ , the probability of obtaining the outcome  $\xi$  is  $p(\xi|\theta) = \text{tr}\{\rho(\theta)\Pi_\xi\}$ . The corresponding Fisher information matrix  $I_{jk}(\Pi, \theta)$  is given by

$$I_{jk}(\Pi, \theta) = \sum_{\xi} \frac{1}{p(\xi|\theta)} \text{tr} \left\{ \frac{\partial \rho(\theta)}{\partial \theta_j} \Pi_\xi \right\} \text{tr} \left\{ \frac{\partial \rho(\theta)}{\partial \theta_k} \Pi_\xi \right\}. \quad (2.18)$$

Once a measurement is chosen, the inverse Fisher information matrix sets a lower bound

## 2.5. Fisher information and Cramér–Rao bound

---

for the MSE matrix of any unbiased estimator, which can be saturated asymptotically by the ML estimator, as in the case of classical parameter estimation. It should be noted that the bound depends on the specific measurement. Measurement-independent bounds will be introduced in Chapters 5 and 6.

In practice, it is often more convenient to use a single number rather than a matrix to quantify the error. A common choice is the scaled MSE  $\text{tr}\{C(\theta)\}$ ; a more general alternative is the scaled WMSE  $\text{tr}\{W(\theta)C(\theta)\}$ , where  $W(\theta)$  is a positive semidefinite weight matrix, which may depend on  $\theta$ . The CR bound implies that  $\text{tr}\{W(\theta)C(\theta)\} \geq \text{tr}\{W(\theta)I^{-1}(\theta)\}$ ; again, this bound can be saturated asymptotically with the ML estimator. A problem with the MSE is that it depends on the parametrization, which is somehow arbitrary. With a suitable choice of the weight matrix, the WMSE is free from this problem. For example, as special cases of the WMSE, the MSH and the MSB are parametrization independent.

In Part I of this thesis, we shall often use the affine parametrization of quantum states

$$\rho(\theta) = \frac{1}{d} + \sum_{j=1}^{d^2-1} \theta_j E_j, \quad (2.19)$$

where the  $E_j$ s form an orthonormal basis in the space of traceless Hermitian operators. In that case, the MSE is identical with the MSH. In addition, a convenient choice for the operator basis is given by

$$E_{jk} := |j\rangle\langle k|, \quad j, k = 1, 2, \dots, d. \quad (2.20)$$

Note that the indices of basis elements of the Hilbert space run from 1 to  $d$  in Part I of the thesis. An alternative candidate is

$$E_{jk}^+ := \frac{1}{\sqrt{2}}(|j\rangle\langle k| + |k\rangle\langle j|), \quad E_{jk}^- := -\frac{i}{\sqrt{2}}(|j\rangle\langle k| - |k\rangle\langle j|), \quad j \leq k. \quad (2.21)$$

By convention,  $E_j$  refers to a generic element in an orthonormal operator basis, whereas  $E_{jk}$ ,  $E_{jk}^+$ , and  $E_{jk}^-$  refer to specific basis elements defined above.



# Quantum state estimation with fully symmetric measurements and product measurements<sup>1</sup>

---

## 3.1 Introduction

Quantum state estimation is a procedure for inferring the state of a quantum system from generalized measurements. Given an ensemble of identically prepared quantum systems, the simplest measurement scheme consists of identical and independent measurements on individual copies. A measurement is IC if any state is determined completely by the measurement statistics [52, 74, 223]. A particularly appealing choice of IC measurements are those constructed out of *weighted 2-designs* [232, 244] (see Appendix B), called *tight* IC measurements according to Scott [244]. In linear quantum state tomography, they not only feature a simple state reconstruction formula but also minimize the MSE, the mean square HS distance between the estimator and the true state. The construction of tight IC measurements was discussed in detail in Ref. [234].

A prominent example of tight IC measurements are SIC POMs [6, 232, 245, 275], which turn out to be the only minimal tight IC measurements [244]. They may be considered as fiducial measurements for state tomography owing to their high symmetry and high tomographic efficiency [6, 226, 232, 244, 245]. In addition to applications in

---

<sup>1</sup>This chapter is based on the following paper: H. Zhu and B.-G. Englert, *Quantum state tomography with fully symmetric measurements and product measurements*, Phys. Rev. A **84**, 022327 (APS, 2011).

### Chapter 3. Fully symmetric measurements and product measurements

---

quantum state tomography, SIC POMs have attracted much attention because of their connections with mutually unbiased bases (MUB) [8, 83, 155, 271, 272], equiangular lines [175], Lie algebras [14], and foundational studies [99] (see Chapter 7).

The trace distance is one of the most important distance and distinguishability measures in quantum mechanics, and is widely used in quantum state tomography, quantum cryptography, and entanglement theory [30, 102, 150, 206, 208], as well as other contexts. It is also closely related to other prevalent figures of merit, such as the fidelity and the Shannon distinguishability [102, 206]. However, little is known about the tomographic resources required to achieve a given accuracy as quantified by the trace distance since its definition, which involves taking the square root of a positive operator, makes analytical studies difficult. Even for the qubit SIC POM, no analytical formula is known for computing the mean trace distance between the estimator and the true state. One motivation behind the present study is to solve this open problem.

In the case of a bipartite or multipartite system, it is technologically much more challenging to perform joint measurements, such as a SIC POM, on the whole system. Moreover, in some important realistic scenarios, such as tomographic quantum key distribution [47, 84, 91, 181], all parties are spatially separated from each other, so it is impractical to perform full joint measurements. Nevertheless, each party can perform a local SIC POM and reconstruct the global state after gathering all the data obtained. Such a POM is henceforth referred to as a product SIC POM; by contrast, the SIC POM for the whole system is referred to as the joint SIC POM. The product SIC POM is particularly appealing in tomographic quantum key distribution since it minimizes the redundant information and classical communication required to exchange measurement data among different parties [84]. However, even less is known concerning its tomographic efficiency except for numerical studies in the two-qubit setting [51, 255].

In this chapter, we characterize the tomographic efficiency of tight IC measurements in terms of the mean trace distance and the mean HS distance, with special emphasis placed on the minimal tight IC measurements—the SIC POMs. We also determine the efficiency gap between product measurements and joint measurements in the bipartite

### 3.1. Introduction

---

and multipartite settings. Incidentally, all SIC POMs used in our numerical simulations are generated by the Heisenberg–Weyl group (see Sec. 7.3.1) from the fiducial states of Ref. [232]. However, all theoretical analysis is independent of this specific choice.

First, we introduce random-matrix theory [196] to study the tomographic efficiency of tight IC measurements, thereby deriving analytical formulas for the mean trace distance and the mean HS distance. We illustrate the general result with SIC POMs and show different scaling behaviors of the two error measures with the dimension of the Hilbert space. As a byproduct, our study uncovers a special class of tight IC measurements that feature exceptionally symmetric outcome statistics and low fluctuation over repeated experiments. In the case of a qubit, we compare the similarities and differences between the SIC POM and the MUB, as well as other measurements constructed out of platonic solids. We also explicate the dependence of the reconstruction error on the Bloch vector of the unknown true state and make contact with experimental data.

Next, in bipartite and multipartite settings, we show that product SIC POMs are optimal among all product measurements in the same sense as joint SIC POMs among joint measurements. For bipartite systems, there is only a marginal efficiency advantage of joint SIC POMs over product SIC POMs, and it is not worth the trouble to perform joint measurements. However, for multipartite systems, the efficiency advantage increases exponentially with the number of parties.

To provide a simple picture of the tomographic efficiencies of SIC POMs and product SIC POMs, we restrict our attention to the scenario in which the number of copies of true states available is large enough to yield a reasonably good estimator and focus on linear state-reconstruction [208, 244]. The analysis of other reconstruction schemes, such as MLE [152, 208, 255], is much more involved and will be postponed to Chapter 4. Hopefully, our analysis may serve as a starting point and may be generalized to deal with those more complicated situations. Moreover, for minimal tomography on a large sample, the estimator given by linear-reconstruction is identical to that determined by MLE with quite a high probability, except when the true state is very close to the boundary of the state space. This is because the former maximizes the likelihood

functional whenever it is positive semidefinite (see Sec. 2.4.2). Therefore, the efficiencies of the two alternative schemes are close to each other in this scenario.

## 3.2 Setting the stage

### 3.2.1 Linear state tomography

A generalized measurement is composed of a set of outcomes represented mathematically by positive operators  $\Pi_j$  that sum up to the identity 1. Given an unknown true state  $\rho$ , the probability of obtaining the outcome  $\Pi_j$  is given by the Born rule:  $p_j = \text{tr}(\Pi_j \rho)$ . A measurement is IC if we can reconstruct any state according to the statistics of measurement results, namely, the set of probabilities  $p_j$ . When both the state  $\rho$  and the outcome  $\Pi_j$  are represented by vectors in the space of Hermitian operators, the probability can be expressed as an inner product  $\langle\langle \Pi_j | \rho \rangle\rangle := \text{tr}(\Pi_j \rho)$ , where the double ket (bra) notation is borrowed from Refs. [71, 73]. Furthermore, superoperators, such as the out product  $|\Pi_j\rangle\rangle\langle\langle \Pi_j|$ , act on this space just as operators on the ordinary Hilbert space (the arithmetics of superoperators can be found in Refs. [238, 239]). With this background, one can show that a measurement is IC if and only if the *frame superoperator*

$$\mathcal{F} = d \sum_j \frac{|\Pi_j\rangle\rangle\langle\langle \Pi_j|}{\text{tr}(\Pi_j)} \quad (3.1)$$

is invertible [61, 73, 81, 244], where the factor  $d$  is introduced for the convenience of later discussions. The frame superoperator  $\mathcal{F}$  can be written as [244]

$$\mathcal{F} = |1\rangle\rangle\langle\langle 1| + \bar{\mathcal{F}}, \quad (3.2)$$

where  $\bar{\mathcal{F}}$  is the projection of  $\mathcal{F}$  onto the space of traceless Hermitian operators. Let  $\bar{\mathbf{I}}$  denote the identity superoperator on this space and  $\bar{\Pi}_j = \Pi_j - \text{tr}(\Pi_j)/d$ ; then we have

$$\bar{\mathcal{F}} = \bar{\mathbf{I}}\mathcal{F}\bar{\mathbf{I}} = d \sum_j \frac{|\bar{\Pi}_j\rangle\rangle\langle\langle \bar{\Pi}_j|}{\text{tr}(\Pi_j)}. \quad (3.3)$$



### 3.2. Setting the stage

---

Obviously,  $\mathcal{F}$  is invertible if and only if  $\bar{\mathcal{F}}$  is invertible in this space. In the rest of this thesis,  $\bar{\mathcal{F}}$  is also referred to as the frame superoperator unless otherwise stated. In addition,  $\bar{\mathcal{F}}^{-1}$  denotes the inverse of  $\bar{\mathcal{F}}$  in the space of traceless Hermitian operators, and the same applies to other superoperators supported on this space.

When  $\mathcal{F}$  is invertible, there exists a set of reconstruction operators  $\Theta_j$  satisfying  $\sum_j |\Theta_j\rangle\rangle\langle\langle\Pi_j| = \mathbf{I}$ , where  $\mathbf{I}$  is the identity superoperator. Given a set of reconstruction operators, any state can be recovered from the set of probabilities  $p_j$ :  $\rho = \sum_j p_j \Theta_j$ . In a realistic scenario, given  $N$  copies of the unknown true state, what we really get in an experiment are frequencies  $f_j$  rather than probabilities  $p_j$ . The estimator based on these frequencies  $\hat{\rho} = \sum_j f_j \Theta_j$  is thus different from the true state. Nevertheless, the deviation  $\hat{\rho} - \rho$  vanishes in the large- $N$  limit as long as the measurement is IC. In general, these frequencies obey a multinomial distribution with the scaled MSE matrix  $\Sigma_{jk} = p_j \delta_{jk} - p_j p_k$ . The scaled MSE matrix of the estimator  $\hat{\rho}$  can be derived by virtue of the principle of error propagation,

$$\mathcal{C}(\rho) = \sum_{j,k} |\Theta_j\rangle\rangle\langle\langle\Sigma_{jk}|\Theta_k| = \sum_j |\Theta_j\rangle\rangle\langle\langle\Pi_j|\rho\rangle\rangle\langle\langle\Theta_j| - |\rho\rangle\rangle\langle\langle\rho|. \quad (3.4)$$

Denote by  $\Delta\rho = \sqrt{N}(\hat{\rho} - \rho)$  the scaled deviation of the estimator from the true state. Then the scaled MSE reads

$$\mathcal{E}(\rho) := \mathbb{E}(\|\Delta\rho\|_{\text{HS}}^2) = \text{Tr}\{\mathcal{C}(\rho)\} = \sum_j p_j \text{tr}(\Theta_j^2) - \text{tr}(\rho^2). \quad (3.5)$$

Here “Tr” denotes the trace of a superoperator, and “tr” of an ordinary operator.

The set of reconstruction operators is unique for a minimal IC measurement, such as a SIC POM or a product SIC POM, but not for a generic IC measurement. Among all the candidates, the set of *canonical reconstruction operators*

$$|\Theta_j\rangle\rangle = \frac{d\mathcal{F}^{-1}|\Pi_j\rangle\rangle}{\text{tr}(\Pi_j)} \quad (3.6)$$

is the best choice for linear state reconstruction in the sense of minimizing the MSE

averaged over unitarily equivalent true states and is thus widely used in practice [244] (see also Sec. 4.2). In the rest of this chapter, we consider only canonical reconstruction operators. It is then straightforward to verify that  $|1\rangle\rangle$  is an eigenvector of  $\mathcal{C}(\rho)$  with eigenvalue 0; in other words,  $\mathcal{C}(\rho)$  is supported on the space of traceless Hermitian operators as is  $\bar{\mathcal{F}}$ . The other eigenvalues of  $\mathcal{C}(\rho)$  determine the variances along the principle axes and thus the shape of the uncertainty ellipsoid.

When  $N$  is sufficiently large, the multinomial distribution approximates a Gaussian distribution, which is completely determined by its mean and MSE matrix. The Gaussian approximation is already quite good for moderate values of  $N$  if we are mainly concerned with quantities like the mean trace distance and the mean HS distance, which are the most popular figures of merit in quantum state tomography. We thus assume the validity of this approximation in the following discussion. Now, the variance of the scaled square error  $\|\Delta\rho\|_{\text{HS}}^2$  is given by the simple formula

$$\mathcal{V}(\rho) := \text{Var}(\|\Delta\rho\|_{\text{HS}}^2) = 2 \text{Tr}\{\mathcal{C}(\rho)^2\}. \quad (3.7)$$

In practice,  $\sqrt{\mathcal{V}(\rho)}$  quantifies the amount of fluctuation in  $\|\Delta\rho\|_{\text{HS}}^2$  over repeated experiments, that is, the typical error in estimating  $\mathcal{E}(\rho)$  with just one experiment, assuming that the true state is known. This error can be reduced by a factor of  $\sqrt{N_e}$  by repeating the experiment  $N_e$  times and taking the average of  $\|\Delta\rho\|_{\text{HS}}^2$ . Once  $\mathcal{E}(\rho)$  is fixed,  $\mathcal{V}(\rho)$  also quantifies the dispersion of the eigenvalues of  $\mathcal{C}(\rho)$  or the degree of anisotropy in the distribution of the estimators.

### 3.2.2 Tight IC measurements

An IC measurement is *tight* if the frame superoperator  $\bar{\mathcal{F}}$  is proportional to  $\bar{\mathbf{I}}$ ; that is,  $\bar{\mathcal{F}} = a\bar{\mathbf{I}}$  for  $a > 0$ . According to Scott [244], the coefficient  $a$  is upper bounded by  $d/(d+1)$  for any tight IC measurement, and the bound is saturated if and only if the measurement is rank one. Rank-one tight IC measurements are optimal for linear state tomography in the sense of minimizing the average MSE over unitarily equivalent states. Here we recapitulate his main idea in a way that suits our subsequent discussion.

### 3.2. Setting the stage

---

According to Eqs. (3.4) and (3.5), it is enough to show the optimality of rank-one tight IC measurements when the true state is the completely mixed state, which is the average of any set of states that is unitarily invariant. In that case, the scaled MSE matrix and MSE reduce to

$$\mathcal{C}\left(\frac{1}{d}\right) = \mathcal{F}^{-1} - \frac{|1\rangle\rangle\langle\langle 1|}{d^2} = \bar{\mathcal{F}}^{-1}, \quad \mathcal{E}\left(\frac{1}{d}\right) = \text{Tr}(\bar{\mathcal{F}}^{-1}). \quad (3.8)$$

The first equation endows the frame superoperator  $\bar{\mathcal{F}}$  with a concrete operational meaning as the inverse of the scaled MSE matrix evaluated at the point  $\rho = 1/d$ . According to the definitions of the frame superoperators  $\mathcal{F}$  and  $\bar{\mathcal{F}}$  (see Sec. 3.2.1),

$$\text{Tr}(\bar{\mathcal{F}}) = \text{Tr}(\mathcal{F}) - d \leq d \sum_j \text{tr}(\Pi_j) - d = d(d-1), \quad (3.9)$$

and the inequality is saturated if and only if the measurement is rank one. Therefore,

$$\mathcal{E}\left(\frac{1}{d}\right) = \text{Tr}(\bar{\mathcal{F}}^{-1}) \geq \frac{1}{d}(d+1)(d^2-1), \quad (3.10)$$

recalling that  $\bar{\mathcal{F}}$  is supported on the space of traceless Hermitian operators, which has dimension  $d^2-1$ . The above inequalities are saturated if  $\bar{\mathcal{F}} = d\bar{\mathbf{I}}/(d+1)$  and only then. So rank-one tight IC measurements are indeed optimal in minimizing the MSE [244]. In that case, the MSE matrix is proportional to  $\bar{\mathbf{I}}$  when  $\rho = 1/d$ , so that the uncertainty ellipsoid is isotropic in the space of traceless Hermitian operators. This feature is quite useful to our later discussions.

A rank-one tight IC measurement with outcomes  $\Pi_j = |\psi_j\rangle w_j \langle\psi_j|$  features particularly simple canonical reconstruction operators

$$\Theta_j = |\psi_j\rangle (d+1) \langle\psi_j| - 1 \quad (3.11)$$

and, thus, easy state reconstruction. The scaled MSE follows from Eq. (3.5),

$$\mathcal{E}(\rho) = d^2 + d - 1 - \text{tr}(\rho^2), \quad (3.12)$$

which turns out to be unitarily invariant. Given any measurement in conjunction with linear state tomography, Eq. (3.12) sets a lower bound for the average scaled MSE, henceforth called the *Scott bound* [244].

There is a close relation between rank-one tight IC measurements and weighted 2-designs: A rank-one measurement with outcomes  $\Pi_j = |\psi_j\rangle w_j \langle \psi_j|$  is tight IC if and only if the weighted set  $\{|\psi_j\rangle, w_j\}$  forms a weighted 2-design [244] (see Appendix B for a brief introduction of weighted  $t$ -designs). For example, SIC POMs and complete sets of mutually unbiased measurements are rank-one tight IC measurements according to this connection, which can also be verified directly. More examples of tight IC measurements can be found in Ref. [234].

### 3.3 Applications of random-matrix theory to quantum state tomography

In this section, we apply random-matrix theory [196] to studying the tomographic efficiency of tight IC measurements and illustrate the general result with SIC POMs. In particular, we derive analytical formulas for the mean trace distance and the mean HS distance between the estimator and the true state, thereby giving a simple picture of the resources required to achieve a given accuracy as quantified by either of the two distances. Our study clearly shows different scaling behaviors of the two error measures with the dimension of the Hilbert space. The idea of computing the mean trace distance using random-matrix theory may also be extended to investigate other figures of merit that depend on only the difference between the estimator and the true state.

#### 3.3.1 A simple idea

Here is a simple idea of computing the mean trace distance with random-matrix theory: In each experiment, after measuring  $N$  copies of the unknown true state  $\rho$ , we can construct an estimator  $\hat{\rho}$  for the true state according to the procedure described in Sec. 3.2.1. Once a basis is fixed, the scaled deviation  $\Delta\rho$  can be represented by a  $d \times d$

### 3.3. Applications of random-matrix theory to quantum state tomography

matrix, which varies from one experiment to another. After a large number of repeated experiments, the set of matrices  $\Delta\rho$  form an ensemble of random matrices that obeys a multidimensional Gaussian distribution

$$p(\Delta\rho) \propto \exp\left(-\frac{1}{2}\langle\langle\Delta\rho|\mathcal{C}(\rho)^{-1}|\Delta\rho\rangle\rangle\right). \quad (3.13)$$

Since  $\mathcal{C}(\rho)$  is supported on the space of traceless Hermitian operators, the distribution of  $\Delta\rho$  is restricted to the hyperplane satisfying  $\text{tr}(\Delta\rho) = 0$ . Suppose  $f(x)$  is the level-density function of this ensemble of matrices with the normalization convention  $\int dx f(x) = d$ . Then the scaled mean trace distance between the estimator and the true state is proportional to the first absolute moment of  $f(x)$ ,

$$\mathcal{E}_{\text{tr}}(\rho) := \frac{1}{2}\text{E}(\text{tr}|\Delta\rho|) = \frac{1}{2}\int dx |x|f(x). \quad (3.14)$$

If  $\mathcal{C}(\rho)$  is (approximately) proportional to the identity superoperator  $\mathbf{I}$ , then the ensemble of matrices  $\Delta\rho' = \sqrt{d^2/2\mathcal{E}(\rho)}\Delta\rho$  is (approximately) a standard Gaussian unitary ensemble. According to random-matrix theory, for sufficiently large  $d$ , the level-density  $f_{\text{G}}(x)$  of the Gaussian unitary ensemble is specified by the famous Wigner semicircle law [196]:

$$f_{\text{G}}(x) = \begin{cases} \frac{1}{\pi}(2d - x^2)^{1/2} & \text{if } |x| \leq \sqrt{2d}, \\ 0 & \text{otherwise.} \end{cases} \quad (3.15)$$

We can derive  $f(x)$  from  $f_{\text{G}}(x)$  by a scale transformation and then compute the scaled mean trace distance between the estimator and the true state, with the result

$$\mathcal{E}_{\text{tr}}(\rho) \approx \frac{4}{3\pi}\sqrt{d\mathcal{E}(\rho)}. \quad (3.16)$$

This equation is still quite accurate if  $\mathcal{C}(\rho)$  is approximately proportional to  $\bar{\mathbf{I}}$  instead of  $\mathbf{I}$ , especially when  $d$  is large. Therefore, the feasibility of our approach is not limited by the fact that  $\mathcal{C}(\rho)$  is supported on the space of traceless Hermitian operators.

When  $\mathcal{C}(\rho)$  is proportional to  $\bar{\mathbf{I}}$ , the scaled deviation  $\Delta\rho$  follows a  $(d^2 - 1)$ -

dimensional isotropic Gaussian distribution, and  $\|\Delta\rho\|_{\text{HS}}^2$  obeys a  $\chi^2$  distribution with  $d^2 - 1$  degrees of freedom. The scaled mean HS distance can thus be computed with the result

$$\mathcal{E}_{\text{HS}}(\rho) := \text{E}(\|\Delta\rho\|_{\text{HS}}) = \sqrt{\frac{\mathcal{E}(\rho)}{d^2 - 1} \frac{\sqrt{2}\Gamma(\frac{d^2}{2})}{\Gamma(\frac{d^2-1}{2})}}. \quad (3.17)$$

As a consequence of the central-limit theorem,  $\mathcal{E}_{\text{HS}}(\rho)$  is almost equal to the square root of  $\mathcal{E}(\rho)$  when  $d$  is large, and with a high probability the estimator  $\hat{\rho}$  is distributed within a thin spherical shell of radius  $\mathcal{E}_{\text{HS}}(\rho)$  that is centered at the true state.

In general, the accuracy of Eqs. (3.16) and (3.17) may depend on the dimension of the Hilbert space and the degree of anisotropy of the uncertainty ellipsoid as determined by  $\mathcal{C}(\rho)$ . However, it turns out that the mean trace distance and the mean HS distance are not so sensitive to the degree of anisotropy of the uncertainty ellipsoid. As we shall see shortly, the two equations are surprisingly accurate for a large family of measurements, especially tight IC measurements, even if  $d$  is very small (see Fig. 3.1).

Although we have started our analysis from linear state tomography, the idea of computing the mean trace distance with random-matrix theory has a wider applicability. We may apply this approach to study the tomographic efficiencies of other reconstruction schemes, such as the ML method. We may also consider other figures of merit that depend on only the deviation between the estimator and the true state.

### 3.3.2 Isotropic measurements

In this section we single out those rank-one IC measurements for which the uncertainty ellipsoid is the most isotropic, so that Eqs. (3.16) and (3.17) are best justified. These measurements turn out to be a special class of tight IC measurements. In addition to minimizing the MSE, they also minimize the fluctuation of the reconstruction error over repeated experiments. Moreover, these IC measurements have the nice property that the mean reconstruction error is almost independent of the true state.

Suppose we have a rank-one IC measurement with outcomes  $\Pi_j = |\psi_j\rangle w_j \langle \psi_j|$ . According to Sec. 3.2.2, the MSE matrix for the completely mixed state is proportional to  $\bar{\mathbf{I}}$  if and only if the measurement is tight IC. For a generic true state, the degree

### 3.3. Applications of random-matrix theory to quantum state tomography

of anisotropy of the MSE matrix can be quantified by  $\overline{\text{Tr}\{\mathcal{C}(\rho)^2\}} - [\overline{\text{Tr}\{\mathcal{C}(\rho)\}}]^2$ , where the over-line denotes the average over unitarily equivalent density operators. Since the scaled MSE  $\text{Tr}\{\mathcal{C}(\rho)\}$  is the same for all rank-one tight IC measurements according to Eq. (3.12), it is advisable to focus on  $\overline{\text{Tr}\{\mathcal{C}(\rho)^2\}}$ . Note that  $\overline{\text{Tr}\{\mathcal{C}(\rho)^2\}}$  also quantifies the fluctuation in  $\|\Delta\rho\|_{\text{HS}}^2$  over repeated experiments according to Eq. (3.7). We find

$$\begin{aligned} \overline{\text{Tr}\{\mathcal{C}(\rho)^2\}} &= d^2 + 2d - \frac{2}{d} + \frac{(d+1)^3\Phi_3 - 2(2d^2 + 3d - 1)}{(d-1)} \left[ \text{tr}(\rho^2) - \frac{1}{d} \right] \\ &\quad + [\text{tr}(\rho^2)]^2 - 2 \left[ \frac{2(d+1)}{d+2} \text{tr}(\rho^3) + \frac{d-1}{d+2} \text{tr}(\rho^2) - \frac{1}{d+2} \right] \end{aligned} \quad (3.18)$$

$$\begin{aligned} &\geq d^2 + 2d - \frac{2}{d} + 2 \frac{d^2 - 2}{d+2} \left[ \text{tr}(\rho^2) - \frac{1}{d} \right] + [\text{tr}(\rho^2)]^2 \\ &\quad - 2 \left[ \frac{2(d+1)}{d+2} \text{tr}(\rho^3) + \frac{d-1}{d+2} \text{tr}(\rho^2) - \frac{1}{d+2} \right], \end{aligned} \quad (3.19)$$

where  $\Phi_3$  is the order-3 frame potential defined in Eq. (B.1), and we have applied the inequality  $\Phi_3 \geq 6d/(d+1)(d+2)$  in deriving Eq. (3.19). The lower bound is saturated if and only if the set  $\{|\psi_j\rangle, w_j\}$  forms a weighted 3-design.

An IC measurement derived from a weighted 3-design is called an *isotropic measurement* for reasons that will become clear shortly (see Sec. 3.3.4 for some concrete examples in the case of a qubit). According to the properties of weighted 3-designs, the scaled MSE matrix  $\mathcal{C}(\rho)$  is the same for any IC measurement derived from a weighted 3-design, including the covariant measurement composed of all pure states weighted by the Haar measure. In other words,  $\mathcal{C}(\rho)$  is invariant under any unitary transformation of the measurement outcomes. As a consequence, the mean reconstruction error is unitarily invariant as long as the figure of merit is unitarily invariant, such as the mean trace distance, the mean HS distance, or the mean fidelity.

In linear state tomography, in addition to achieving the minimal MSE, an isotropic measurement also minimizes the fluctuation of the statistical error over repeated experiments or, equivalently, the degree of anisotropy in the distribution of  $\Delta\rho$ . Calculation shows that the scaled MSE matrix  $\mathcal{C}(\rho)$  for a pure true state has only four (three if  $d=2$ ) distinct eigenvalues,  $(d+1)/(d+2)$ ,  $2(d+1)/(d+2)$ ,  $2d/(d+2)$ ,  $0$  with multiplicities  $d(d-2)$ ,  $2(d-1)$ ,  $1$ ,  $1$ , respectively. The degree of anisotropy is even lower if the

### Chapter 3. Fully symmetric measurements and product measurements

---

true state is mixed since the leading contribution to  $\mathcal{C}(\rho)$  is linear in  $\rho$  [see Eq. (3.4)].

In conclusion, Eqs. (3.16) and (3.17) are good approximations for computing the scaled mean trace distance and the mean HS distance under isotropic measurements. After inserting Eq. (3.12) into Eqs. (3.16) and (3.17), we get

$$\mathcal{E}_{\text{tr}}(\rho) \approx \frac{4}{3\pi} \sqrt{d[d^2 + d - 1 - \text{tr}(\rho^2)]} \sim \frac{4}{3\pi} d^{3/2}, \quad (3.20)$$

$$\mathcal{E}_{\text{HS}}(\rho) \approx \sqrt{\frac{d^2 + d - 1 - \text{tr}(\rho^2)}{d^2 - 1}} \frac{\sqrt{2} \Gamma(\frac{d^2}{2})}{\Gamma(\frac{d^2-1}{2})} \sim d. \quad (3.21)$$

The two equations clearly show the difference in the scaling behaviors of the two error measures with the dimension of the Hilbert space.

An isotropic measurement is, in a sense, the most symmetric measurement allowed by quantum mechanics. Remarkably, such a measurement can be realized with only a finite number of outcomes, and its tomographic efficiency can be characterized by simple formulas. On the other hand, an isotropic measurement comprises at least  $d^2(d+1)/2$  outcomes, which are much more than the minimum  $d^2$  required for an IC measurement. Recall that a weighted 3-design comprises at least  $d^2(d+1)/2$  elements [see Eq. (B.2)]. Therefore, tight IC measurements with fewer outcomes, such as SIC POMs, are of more practical interest.

#### 3.3.3 Tight IC POMs and SIC POMs

In this section we consider generic rank-one tight IC measurements, paying particular attention to SIC POMs [6, 232, 245, 275]. When the weighted set  $\{|\psi_j\rangle, w_j\}$  forms a weighted 2-design but not necessarily a weighted 3-design, the inequality  $\Phi_3 \leq \Phi_2 = 2d/(d+1)$  (see Appendix B) applied to Eq. (3.18) implies that

$$\overline{\text{Tr}\{\mathcal{C}(\rho)^2\}} \leq d^2 + 2d + 2d(d+1) \left[ \text{tr}(\rho^2) - \frac{1}{d} \right]. \quad (3.22)$$

In conjunction with Eqs. (3.7) and (3.12), this equation provides two important pieces of information. First, the relative deviation  $\sqrt{\mathcal{V}(\rho)}/\mathcal{E}(\rho)$  is approximately inversely proportional to  $d$ ; hence,  $\mathcal{E}_{\text{HS}}(\rho)$  is approximately equal to the square root of  $\mathcal{E}(\rho)$ ,



### 3.3. Applications of random-matrix theory to quantum state tomography

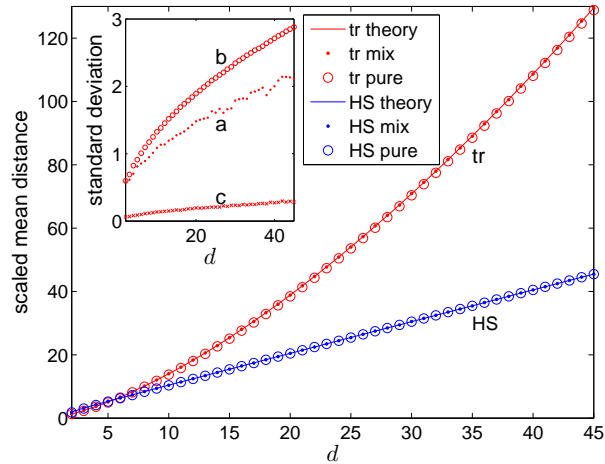


Figure 3.1: Theoretical and numerical simulation results on the scaled mean trace distances and the scaled mean HS distances in state tomography with SIC POMs for dimensions from 2 to 45. The theoretical values are computed according to Eqs. (3.20) and (3.21) with  $\rho = 1/d$ . In the numerical simulation,  $N = 1000 + 20d^2$ . Each data point for the completely mixed state is the average over 1000 repeated experiments, and that for pure states is the average over 1000 randomly generated pure states, each averaged over 100 repeated experiments. The inset shows three kinds of standard deviations of the scaled mean trace distances in the numerical simulation: (a) over repeated experiments for the completely mixed state; (b) over repeated experiments averaged over pure states; (c) over the randomly generated pure states, including a partial contribution over repeated experiments because of the finite number of repetitions.

and Eq. (3.21) is a good approximation for computing the scaled mean HS distance, especially when  $d$  is large. Second, the degree of anisotropy in the distribution of  $\Delta\rho$  cannot be too high as long as the measurement is rank one tight IC. Given that the level-density function  $f(x)$  and, especially, its first absolute moment are not so sensitive to slight variations in the degree of anisotropy, it is reasonable to expect that the scaled mean trace distance can be computed approximately with Eq. (3.20). This expectation is supported by extensive numerical simulations.

Figure 3.1 shows the results of theoretical calculations and numerical simulations on state tomography with SIC POMs. As mentioned before, all SIC POMs are generated by Heisenberg–Weyl groups from the fiducial states of Ref. [232]. The scaled mean trace distance and the scaled mean HS distance from the numerical simulations agree perfectly with the theoretical formulas in Eqs. (3.20) and (3.21); in fact, they agree much better than we expected. Figure 3.1 also clearly illustrates different scaling

### Chapter 3. Fully symmetric measurements and product measurements

---

behaviors of the two error measures with the dimension of the Hilbert space. According to the inset in Fig. 3.1, the fluctuation in the mean trace distance over different pure states is much smaller than the fluctuation over repeated experiments on the same state. Actually, the former is so small that it is difficult to separate out the partial contribution of the latter with a limited number of repeated experiments. Therefore, the reconstruction error is not sensitive to the identity of the true state.

We emphasize that the results on SIC POMs are representative of typical rank-one tight IC measurements. Since the order-3 frame potential  $\Phi_3 = (d^2 + 3d)/(d + 1)^2$  for a SIC POM is much larger than the value  $6d/(d + 1)(d + 2)$  required for a 3-design, a SIC POM is a very poor approximation of a 3-design, for which Eqs. (3.20) and (3.21) are best justified. Alternatively, we can see this point from the value of  $\text{Tr}\{\mathcal{C}(\rho)^2\}$  for a SIC POM, which follows from Eq. (3.4),

$$\text{Tr}\{\mathcal{C}(\rho)^2\} = (d^2 + d + 2)[1 + \text{tr}(\rho^2)] - 1 + [\text{tr}(\rho^2)]^2 - 2(d^2 + d)^2 \sum_j p_j^3. \quad (3.23)$$

When  $d \gg 1$ , the term  $|\rho\rangle\rangle\langle\langle\rho|$  in the expression of  $\mathcal{C}(\rho)$  can be neglected, and we have

$$\text{Tr}\{\mathcal{C}(\rho)^2\} \approx (d^2 + d)[1 + \text{tr}(\rho^2)]. \quad (3.24)$$

Comparison with Eqs. (3.19) and (3.22) shows that the value for a SIC POM is roughly half way between the lower bound and the upper bound for tight IC measurements.

In the rest of this section, we briefly examine tight IC measurements that are not rank-one and that can arise in practice. In realistic experiments on quantum state tomography with a SIC POM, there always exists noise associated with detector inefficiency, dark counts, and other imperfections. It is important to understand how the noise affects the tomographic efficiency. We investigate these effects by means of a simple white-noise model, in which the outcomes of the SIC POM are modified as follows:

$$\Pi_j(\alpha) = \frac{\alpha \frac{1}{d} + |\psi_j\rangle\langle\psi_j|}{d\alpha + d}, \quad (3.25)$$

where the parameter  $\alpha$  ( $\alpha \geq 0$ ) quantifies the strength of the noise. This model is

### 3.3. Applications of random-matrix theory to quantum state tomography

natural when there is no prior knowledge about the noise. Incidentally, measurements of this form have been considered for entanglement detection with witness operators [282].

The measurement introduced above is still tight IC, and the scaled MSE can be calculated according to the procedure presented in Sec. 3.2.1, with the result

$$\mathcal{E}(\rho) = \frac{1}{d} [1 + (d+1)^2(d-1)(\alpha+1)^2] - \text{tr}(\rho^2). \quad (3.26)$$

Compared with Eq. (3.12), the scaled MSE is roughly  $(\alpha+1)^2$  times as large as in the ideal case. The scaled mean trace distance and the scaled mean HS distance can still be computed according to Eqs. (3.16) and (3.17), with the result

$$\mathcal{E}_{\text{tr}}(\rho) \approx \frac{4}{3\pi}(\alpha+1)d^{3/2}, \quad \mathcal{E}_{\text{HS}}(\rho) \approx (\alpha+1)d, \quad (3.27)$$

which are roughly  $\alpha+1$  times the values for the ideal case. Owing to the noise,  $(\alpha+1)^2$  times as many measurements are needed to reach the same accuracy as in the ideal case. Similar analysis also applies to tight IC measurements derived from other 2-designs, such as complete sets of MUB.

#### 3.3.4 Qubit tomography

In this section we show that any measurement constructed out of a platonic solid other than the tetrahedron is an isotropic measurement in the case of a qubit. The similarities and differences between isotropic measurements and the SIC POM are discussed in detail. We then derive exact formulas for the mean trace distances for both isotropic measurements and the SIC POM and explain the dependence of the reconstruction error on the Bloch vector of the true state (see Refs. [51, 183, 226] for earlier accounts). Our study confirms that the earlier result based on random-matrix theory is already quite accurate for  $d = 2$ , although it is best justified when  $d$  is large. As a simple application, we make contact with the experimental result given by Ling et al. [183].

Given a platonic solid with  $n$  vertices inscribed on the Bloch sphere, the unit vectors  $\mathbf{v}_k$  representing the vertices define a measurement with outcomes  $\Pi_k = (1 + \mathbf{v}_k \cdot \boldsymbol{\sigma})/n$ ,

### Chapter 3. Fully symmetric measurements and product measurements

---

where  $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$  are the Pauli matrices. Since the measurement corresponding to any platonic solid is tight IC, the reconstruction operators assume the form  $\Theta_k = (1 + 3\mathbf{v}_k \cdot \boldsymbol{\sigma})/2$  according to Eq. (3.11). Reconstructing the true state  $\rho$  is equivalent to reconstructing its Bloch vector  $\mathbf{s}$ ,

$$\rho = \sum_{k=1}^n p_k \Theta_k = \frac{1}{2} \left( 1 + 3 \sum_{k=1}^n p_k \mathbf{v}_k \cdot \boldsymbol{\sigma} \right), \quad \mathbf{s} = 3 \sum_{k=1}^n p_k \mathbf{v}_k. \quad (3.28)$$

This expression reduces to the one given in Ref. [226] when the platonic solid is a regular tetrahedron and the corresponding measurement is SIC. Incidentally, both the HS norm  $\|\Delta\rho\|_{\text{HS}}$  and the trace norm  $\|\Delta\rho\|_{\text{tr}}$  are proportional to the length of  $\Delta\mathbf{s} = \sqrt{N}(\hat{\mathbf{s}} - \mathbf{s})$ , where  $\hat{\mathbf{s}}$  is an estimator of  $\mathbf{s}$ ; namely,  $\|\Delta\rho\|_{\text{HS}}^2 = (\Delta\mathbf{s})^2/2$  and  $\|\Delta\rho\|_{\text{tr}} = |\Delta\mathbf{s}|/2$ .

According to Eq. (3.4), the scaled MSE matrix of the estimator  $\hat{\rho}$  assumes the form

$$\mathcal{C}(\rho) = \frac{3}{4} \left( \sum_{j=x,y,z} |\sigma_j\rangle\langle\sigma_j| \right) - \frac{1}{4} |\mathbf{s} \cdot \boldsymbol{\sigma}\rangle\langle\mathbf{s} \cdot \boldsymbol{\sigma}| + \frac{9}{4n} \sum_{k=1}^n |\mathbf{v}_k \cdot \boldsymbol{\sigma}\rangle\langle\mathbf{v}_k \cdot \boldsymbol{\sigma}|. \quad (3.29)$$

To get a concrete geometric picture, we had better work with the scaled MSE matrix of the estimator  $\hat{\mathbf{s}}$  of the Bloch vector,

$$C(\mathbf{s}) = 3\mathbf{I}_3 - \mathbf{s}\mathbf{s} + \frac{9}{n} \sum_{k=1}^n (\mathbf{v}_k \cdot \mathbf{s}) \mathbf{v}_k \mathbf{v}_k, \quad (3.30)$$

where  $\mathbf{I}_3$  is the  $3 \times 3$  identity dyadic. The scaled MSE of the estimator  $\hat{\mathbf{s}}$  reads

$$E(|\Delta\mathbf{s}|^2) = 2\mathcal{E}(\rho) = 9 - s^2, \quad (3.31)$$

which is independent of the orientation of the Bloch vector of the true state, as expected for any rank-one tight IC measurement.

When the platonic solid is a cube, octahedron, dodecahedron, or icosahedron, the last term in Eq. (3.30) vanishes owing to their symmetries and, as a consequence, the MSE matrix is independent of the orientation of the platonic solid. In other words, the measurement corresponding to any platonic solid other than the tetrahedron is an isotropic measurement (see Sec. 3.3.2). A particular appealing isotropic measurement

### 3.3. Applications of random-matrix theory to quantum state tomography

is the one corresponding to an octahedron, where the six outcomes form a complete set of MUB, which is a 3-design in dimension 2 (see Appendix B). The MSE matrix for any isotropic measurement is covariant in the sense that the MSE matrices for any two true states with the same purity can be turned into each other by the same rotations that turn their Bloch vectors into each other, which is clearly reflected in the uncertainty ellipsoids, as illustrated in the left plot of Fig. 3.2. Therefore, the mean trace distance is independent of the orientation of the Bloch vector of the true state, and the same is true for any other figure of merit that is unitarily invariant. This is not the case for the SIC POM.

Suppose  $a, b, c$  are the square roots of the three eigenvalues of the scaled MSE matrix  $C(\mathbf{s})$ ; then the scaled mean error is determined by the integral

$$\begin{aligned} E(|\Delta \mathbf{s}|) &= \int dx dy dz \frac{\sqrt{x^2 + y^2 + z^2}}{(2\pi)^{3/2} abc} \exp \left[ - \left( \frac{x^2}{2a^2} + \frac{y^2}{2b^2} + \frac{z^2}{2c^2} \right) \right] \\ &= \sqrt{\frac{2}{\pi}} \int_0^1 dt \frac{c^2 [a^2 c^2 + b^2 c^2 + (2a^2 b^2 - a^2 c^2 - b^2 c^2) t^2]}{[c^4 (1-t)^2 + a^2 b^2 t^4 + (a^2 + b^2) c^2 t^2 (1-t^2)]^{3/2}}. \end{aligned} \quad (3.32)$$

If at least two of the standard deviations are equal, say,  $b = a$ , then we have

$$E(|\Delta \mathbf{s}|) = \begin{cases} \sqrt{\frac{2}{\pi}} c & \text{if } a = 0, \\ \sqrt{\frac{\pi}{2}} a & \text{if } c = 0, \\ 2\sqrt{\frac{2}{\pi}} a & \text{if } c = a, \\ \sqrt{\frac{2}{\pi}} \left( \frac{a^2 \arctan \sqrt{\frac{a^2 - c^2}{c^2}}}{\sqrt{a^2 - c^2}} + c \right) & \text{if } a > c, \\ \sqrt{\frac{2}{\pi}} \left( \frac{a^2 \operatorname{arctanh} \sqrt{\frac{c^2 - a^2}{c^2}}}{\sqrt{c^2 - a^2}} + c \right) & \text{if } a < c. \end{cases} \quad (3.33)$$

If the uncertainty ellipsoid is isotropic; that is,  $a = b = c$ , then Eq. (3.33) implies

$$\mathcal{E}_{\text{tr}}(\rho) = \frac{1}{2} E(|\Delta \mathbf{s}|) = \sqrt{\frac{2}{3\pi}} \sqrt{9 - s^2}. \quad (3.34)$$

For the completely mixed state, this formula is exact; by contrast, the alternative

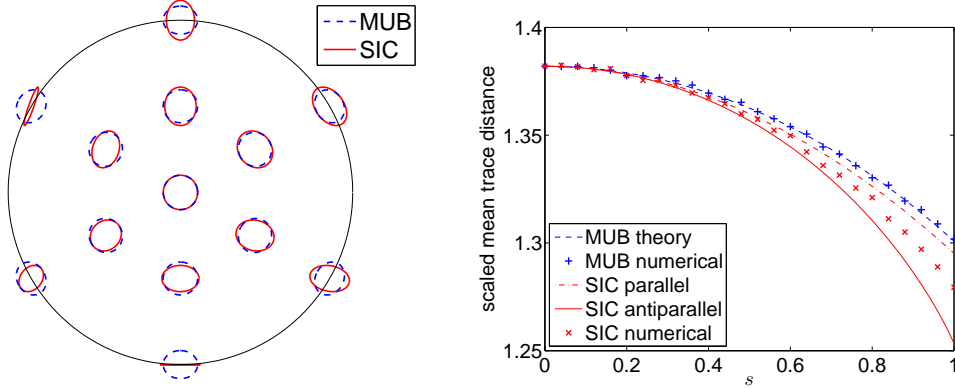


Figure 3.2: Uncertainty ellipses and tomographic efficiencies in linear state tomography on a qubit with the MUB and the SIC POM. (left) Uncertainty ellipses of the marginal distributions on the  $x$ - $z$  plane of the Bloch ball corresponding to 300 measurements. For the MUB, the result is independent of the orientations of the outcomes. For the SIC POM, one outcome is aligned with the  $z$  axis (vertical direction), and another one lies on the  $x$ - $z$  plane with positive  $x$  component. (right) Theoretical and numerical scaled mean trace distances. In theoretical calculation for the SIC POM, the Bloch vector of the true state is either parallel or antiparallel to one of the measurement outcomes. In numerical simulation,  $N$  is set at 1000, and every data point is averaged over 1000 randomly generated states, each averaged over 400 repeated experiments.

$\mathcal{E}_{\text{tr}}(\rho) \approx 4\sqrt{9-s^2}/3\pi$  based on random-matrix theory [see Eq. (3.16)] is about 8% smaller. The disparity is much smaller than the relative deviation of  $\|\Delta\rho\|_{\text{tr}}$  over repeated experiments, which is about 42%, and it is even smaller for other states. This observation shows that the random-matrix approximation is already quite accurate even when  $d = 2$ .

For isotropic measurements, Eq. (3.30) implies that  $a^2 = b^2 = 3$  and  $c^2 = 3 - s^2$ ; the uncertainty ellipsoid is rotationally symmetric and oblate whenever  $s > 0$ . According to Eq. (3.33), the scaled mean trace distance reads

$$\mathcal{E}_{\text{tr}}^{\text{iso}}(\rho) = \frac{1}{\sqrt{2\pi}} \left( \frac{3}{s} \arctan \frac{s}{\sqrt{3-s^2}} + \sqrt{3-s^2} \right), \quad (3.35)$$

which is very close to the value under isotropic approximation since the degree of anisotropy of the uncertainty ellipsoid is low for isotropic measurements.

For the SIC POM, the reconstruction error depends on not only the purity of the

### 3.3. Applications of random-matrix theory to quantum state tomography

true state but also the orientation of the Bloch vector. Those states whose Bloch vectors are either parallel or antiparallel to one of the measurement outcomes have attracted considerable attention from both theoretician [226] and experimentalists [183] since they represent two extreme cases. We shall compute the mean trace distances for those states and discuss this dependence.

If  $\mathbf{v}_1$  is chosen as the  $z$  axis, then the Bloch vectors of those extreme states can be parameterized as  $\mathbf{s} = z\mathbf{v}_1$  with  $-1 \leq z \leq 1$ . According to Eq. (3.30), we have

$$C(\mathbf{s}) = (3 - z)\mathbf{I}_3 + (3z - z^2)\mathbf{v}_1\mathbf{v}_1, \quad (3.36)$$

whose eigenvalues are given by

$$a^2 = b^2 = 3 - z, \quad c^2 = (3 - z)(1 + z). \quad (3.37)$$

The corresponding uncertainty ellipsoid is rotationally symmetric. As  $z$  decreases from 1 to  $-1$ , it evolves from a prolate to an oblate, and finally to a singular ellipsoid. The scaled mean trace distance of those states can be calculated according to Eq. (3.33).

The right plot of Fig. 3.2 shows the scaled mean trace distances in linear state tomography using the MUB and the SIC POM, respectively. Although the MUB and the SIC POM are equally efficient with respect to the MSE, the SIC POM is slightly more efficient with respect to the mean trace distance (the situation can be different with other reconstruction methods; see Sec. 4.5). For the SIC POM, the mean trace distance is slightly smaller for states whose Bloch vectors are antiparallel to one of the outcomes of the measurement than states whose Bloch vectors are parallel. The average of the mean trace distance over randomly generated states with a given purity sits roughly in the middle of the two extreme cases. In all the cases considered, there is a slight decrease in the mean trace distances as the purity of the true state increases, which can be attributed to two reasons: the decrease in the MSEs and the increase in the degrees of anisotropy of the uncertainty ellipsoids.

Ling et al. studied the tomographic efficiency of the qubit SIC POM experimentally

and determined the scaled mean trace distances for the three states with  $z = 0, -1, 1$ , respectively, with the result 1.417, 1.288, 1.323 [183]. For comparison, our theoretical calculation yields the result 1.382, 1.259, 1.295. The experimental and the theoretical values reflect the same dependence of the reconstruction error on the Bloch vector of the true state. The former are slightly larger than the latter, but the differences are very small. Note that the relative fluctuation of the reconstruction error over repeated experiments is larger than 40%, and the experimental data are averaged over only 40 runs. In addition, any imperfection inevitable in real experiments can also affect the accuracy of the estimator.

### 3.4 Joint SIC POMs and Product SIC POMs

In bipartite or multipartite settings, it is technologically much more challenging and sometimes even impossible to perform full joint measurements such as SIC POMs on the whole systems. It is thus of paramount practical interest to determine the optimal product measurements as well as the efficiency gap between product measurements and joint measurements.

#### 3.4.1 Bipartite scenarios

Consider a product measurement on a bipartite system whose parts have dimensions  $d_1$  and  $d_2$ , respectively, and the total dimension is  $d = d_1 d_2$ . To show the optimality of the product SIC POM, we shall use the same strategy described in Sec. 3.2.2. More generally, we show that if the product measurement minimizes the MSE averaged over unitarily equivalent states, then the measurement on each subsystem is rank one tight IC, and vice versa. As a consequence, the product SIC POM is optimal and, furthermore, any minimal optimal product measurement must be a product SIC POM since SIC POMs are the only minimal rank-one tight IC measurements [244].

As in the case of joint measurements (see Sec. 3.2.2), it suffices to demonstrate our claim when  $\rho = 1/d$ . Suppose  $\Pi_{j_1}$  are the outcomes of the measurement on the first subsystem and  $\Pi_{j_2}$  on the second subsystem; then each outcome in the product



### 3.4. Joint SIC POMs and Product SIC POMs

---

measurement has a tensor-product form,  $\Pi_{j_1 j_2} = \Pi_{j_1} \otimes \Pi_{j_2}$ . The same is true for the frame superoperator  $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$  and the reconstruction operators  $\Theta_{j_1 j_2} = \Theta_{j_1} \otimes \Theta_{j_2}$ . According to Eq. (3.8),

$$\mathcal{C}\left(\frac{1}{d}\right) = \mathcal{F}_1^{-1} \otimes \mathcal{F}_2^{-1} - \frac{|1\rangle\rangle\langle\langle 1|}{d^2}, \quad \mathcal{E}\left(\frac{1}{d}\right) = \text{Tr}(\mathcal{F}_1^{-1}) \text{Tr}(\mathcal{F}_2^{-1}) - \frac{1}{d}. \quad (3.38)$$

The MSE is minimized if and only if both  $\text{Tr}(\mathcal{F}_1^{-1})$  and  $\text{Tr}(\mathcal{F}_2^{-1})$  are minimized, that is, if the measurement on each subsystem is rank one tight IC (see Sec. 3.2.2).

Now, let us focus on the tomographic efficiency of the optimal product measurements. If the product measurement is composed of two rank-one tight IC measurements, as in the case of the product SIC POM, then each factor in the reconstruction operator  $\Theta_{j_1 j_2} = \Theta_{j_1} \otimes \Theta_{j_2}$  is given by Eq. (3.11). The scaled MSE can be computed according to Eq. (3.5),

$$\mathcal{E}^{\text{prod}}(\rho) = (d_1^2 + d_1 - 1)(d_2^2 + d_2 - 1) - \text{tr}(\rho^2) \quad (3.39)$$

Surprisingly, the MSE is almost independent of the true state, as in the case of the SIC POM. Meanwhile, it is approximately equal to the product of the MSEs for the two subsystems, respectively. The variance  $\mathcal{V}^{\text{prod}}(\rho)$  of the square error can depend on the specific choice of the product measurement according to Eq. (3.7). For the product SIC POM, it is approximately given by

$$\begin{aligned} \mathcal{V}^{\text{prod}}(\rho) \approx & (d_1^2 + d_1 - 2)(d_2^2 + d_2 - 2)[1 + \text{tr}(\rho^2) + \text{tr}(\rho_1^2) + \text{tr}(\rho_2^2)] \\ & + (d_1^2 + d_1 - 2)[1 + \text{tr}(\rho_1^2)] + (d_2^2 + d_2 - 2)[1 + \text{tr}(\rho_2^2)]. \end{aligned} \quad (3.40)$$

The variance depends on not only the purity of the global state but also the purities of the reduced states, which means that it usually depends on the entanglement of the global state. When the true state is pure, for example, the variance is approximately maximized for product states and minimized for the maximally entangled state.

Compared with the MSE associated with the joint SIC POM given in Eq. (3.12),

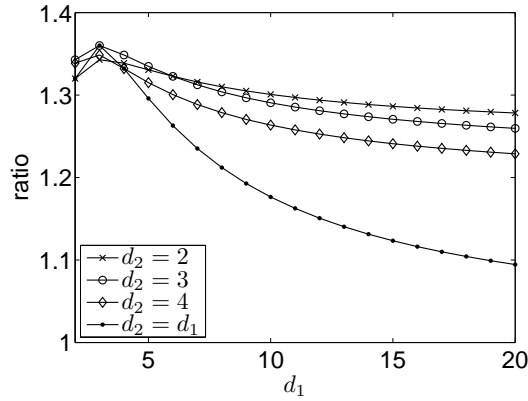


Figure 3.3: The ratio of the MSE associated with the product SIC POM to that with the joint SIC POM when the true state is completely mixed. It is maximized when  $d_1 = d_2 = 3$ . Note that the ratio is almost independent of the true state.

the MSE associated with the product SIC POM is slightly larger, but the difference is generally very small, especially when both  $d_1$  and  $d_2$  are large. On the other hand, the fluctuation over repeated experiments is stronger by a bigger margin for the product SIC POM. Figure 3.3 shows the ratio of the MSEs when the true state is the completely mixed state; the ratios for other states are almost the same. The maximal ratio 1.36 is attained at  $d_1 = d_2 = 3$ . When  $d_1, d_2 \geq 3$ , the ratio decreases monotonically with  $d_1$  and  $d_2$ ; when  $d_2 = 2$  and  $d_1 \geq 3$ , the ratio decreases monotonically with  $d_1$ . For sufficiently large  $d_1$  and  $d_2$ , the ratio is about  $1 + 1/d_1 + 1/d_2$ . In conclusion, there is only a marginal efficiency advantage of using the joint SIC POM over the product SIC POM. The latter is more appealing for practical applications since it is much easier to implement.

Although the product SIC POM is not even a tight IC measurement, comparison of Eqs. (3.39) and (3.40) shows that the relative deviation of the square error is quite small, especially when  $d_1$  and  $d_2$  are large. Hence, Eq. (3.17) is still a good approximation for computing the scaled mean HS distance. The scaled mean trace distance can be calculated approximately according to Eq. (3.16), with the result

$$\mathcal{E}_{\text{tr}}(\rho) \approx \frac{4\sqrt{d_1 d_2}}{3\pi} \sqrt{(d_1^2 + d_1 - 1)(d_2^2 + d_2 - 1) - \text{tr}(\rho^2)}. \quad (3.41)$$

### 3.4. Joint SIC POMs and Product SIC POMs

---

Table 3.1: Theoretical and numerical scaled mean trace distances in two-qubit state estimation with the joint SIC POM (Joint) and the product SIC POM (Prod). In the numerical simulation,  $N$  is set at 1000. For the completely mixed state, each data point is averaged over 1000 repeated experiments. For pure states, it is averaged over 1000 randomly generated states, each averaged over 1000 repeated experiments. The standard deviations of the scaled trace distances over the 1000 randomly generated pure states (including a partial contribution of the fluctuation over repeated experiments for each state owing to the finite number of experiments) are 0.033 and 0.027 for the product SIC POM and the joint SIC POM, respectively, both of which are very small.

POM	Completely mixed state			Average over pure states		
	Theory	Numerical	Error%	Theory	Numerical	Error%
Prod	4.223	4.255	-0.8	4.158	4.162	-0.1
Joint	3.676	3.716	-1.1	3.601	3.575	+0.7
Ratio	1.149	1.145	—	1.155	1.164	—

Generally speaking, the larger the values of  $d_1$  and  $d_2$  are, the more accurate this formula is. The ratio of the mean trace distance for the product SIC POM to that for the joint SIC POM is approximately equal to the square root of the ratio of the MSEs.

Table 3.1 shows the theoretical and numerical simulation results of the scaled mean trace distances for the two-qubit product SIC POM and the joint SIC POM. There is quite a good agreement between theoretical calculations and numerical simulations although  $d_1$  and  $d_2$  are so small. The mean trace distances achieved by the product SIC POM are roughly 15% larger than that by the joint SIC POM. As a consequence, with the product SIC POM, we need about 32% more copies of the true states to reach the same accuracy achieved by the joint SIC POM. Despite its slightly lower efficiency, the product SIC POM is more appealing than the joint SIC POM owing to its relatively easier implementation in real experiments. In the case of two qubits, the same conclusion was reached in Ref. [255], where the ML method was adopted for state reconstruction.

#### 3.4.2 Multipartite scenarios

Suppose  $k$  parties want to reconstruct a quantum state shared among them with a product measurement, and  $d_j$  for  $j = 1, 2, \dots, k$  is the dimension of the Hilbert space

of the  $j$ th party. According to the same analysis as in the bipartite setting, the product SIC POM is optimal among all product measurements in linear state tomography. The scaled MSE achieved by the product SIC POM can be calculated in the same manner, with the result

$$\mathcal{E}^{\text{prod}}(\rho) = \prod_{j=1}^k (d_j^2 + d_j - 1) - \text{tr}(\rho^2), \quad (3.42)$$

which is almost independent of the true state. When the true state is completely mixed, the variance of the square error is

$$\mathcal{V}^{\text{prod}}(\rho) = 2 \left( \prod_{j=1}^k \frac{(d_j^3 + 2d_j^2 - 2)}{d_j} - \frac{1}{d^2} \right). \quad (3.43)$$

For a generic state, the variance depends on the purity of the global state as well as the purities of various reduced states and can be much larger than the value given above.

If the dimension of the Hilbert space of each party is equal to  $d_1$ , then the ratio of the MSE for the product SIC POM to that for the joint SIC POM grows exponentially with the number of parties  $k$ ,

$$\frac{\mathcal{E}^{\text{prod}}(\rho)}{\mathcal{E}^{\text{joint}}(\rho)} \approx \left( 1 + \frac{1}{d} - \frac{1}{d^2} \right)^{-1} \left( 1 + \frac{1}{d_1} - \frac{1}{d_1^2} \right)^k. \quad (3.44)$$

The ratio of the variances grows with an even higher rate, and its specific value can heavily depend on the true state. Therefore, the efficiency advantage of the joint SIC POM over the product SIC POM grows exponentially with the number of parties.

Although the fluctuation in the reconstruction error over repeated experiments is stronger in the product SIC POM as compared with the joint SIC POM, the relative fluctuation is still weak. So Eq. (3.17) is still a good approximation for computing the scaled mean HS distance. When  $k$  is not too large, according to Eq. (3.16), the scaled mean trace distance is approximately given by

$$\mathcal{E}_{\text{tr}}(\rho) \approx \frac{4\sqrt{d}}{3\pi} \sqrt{\prod_{j=1}^k (d_j^2 + d_j - 1) - \text{tr}(\rho^2)}. \quad (3.45)$$

### 3.5. Summary

---

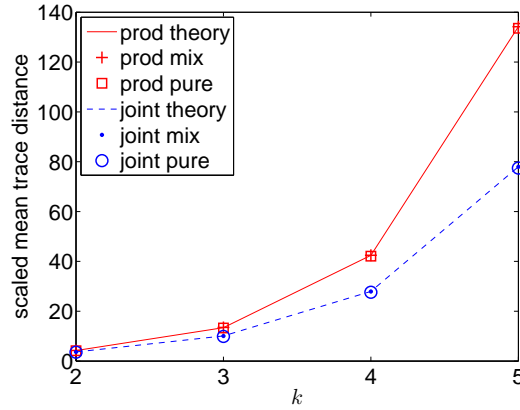


Figure 3.4: Results of theoretical calculation and numerical simulation of the scaled mean trace distances for the joint SIC POMs and the product SIC POMs on multiqubit systems, where  $k$  is the number of qubits. The theoretical values derive from Eqs. (3.20) and (3.45) with  $\rho = 1/d$ . In the numerical simulation,  $N$  is set at  $1000 + 20d^2$ . For the completely mixed state, each data point is averaged over 1000 repeated experiments. For pure states, it is averaged over 1000 randomly generated states, each averaged over 100 repeated experiments.

Since the ratio of the mean trace distance achieved by the product SIC POM to that by the joint SIC POM is approximately equal to the square root of the ratio of the MSEs, it also increases exponentially with the number of parties; the same is true for the ratio of the mean HS distances. Figure 3.4 shows the results of theoretical calculation and numerical simulation of the scaled mean trace distances for the product SIC POMs and the joint SIC POMs on multiqubit systems. There is a pretty good agreement between theoretical prediction and numerical simulation for  $k$  up to 5. Figure 3.4 further confirms that the efficiency advantage of using the joint SIC POM over the product SIC POM increases exponentially with the number of parties.

### 3.5 Summary

We have introduced random-matrix theory for studying the tomographic efficiency of tight IC measurements, which include SIC POMs as a special example. In particular, we derived analytical formulas for the mean trace distance and the mean HS distance between the estimator and the true state and showed different scaling behaviors of the two error measures with the dimension of the Hilbert space. The accuracy of these

### Chapter 3. Fully symmetric measurements and product measurements

---

formulas was confirmed by extensive numerical simulations on state tomography with SIC POMs. As a byproduct, we also discovered a special class of tight IC measurements, called isotropic measurements, which feature exceptionally symmetric outcome statistics and low fluctuation over repeated experiments. In the case of a qubit, we provided several concrete examples of isotropic measurements that are constructed out of platonic solids other than the tetrahedron, and explicated the similarities and differences between isotropic measurements and the SIC POM. We also derived exact formulas of the mean trace distances for both isotropic measurements and the SIC POM, followed by a detailed explanation of the dependence of the reconstruction error on the Bloch vector of the true state.

In bipartite and multipartite settings, we showed that the product SIC POMs are optimal among all product measurements in the same sense as the joint SIC POMs among all joint measurements. For bipartite systems, there is only a marginal efficiency advantage of using the joint SIC POMs over the product SIC POMs, which vanishes in the large-dimension limit. Hence, it is not worth the trouble of performing joint measurements at the current stage. For multipartite systems, however, this efficiency advantage increases exponentially with the number of parties.

Our study furnished a simple picture of the scaling behavior of resource requirement in state tomography with the dimension of the Hilbert space and of the efficiency gap between product measurements and joint measurements. The idea of applying random-matrix theory to studying tomographic efficiencies may also find wider applications in other state-estimation problems.

# The power of informationally overcomplete measurements

---

## 4.1 Introduction

A central problem in quantum state estimation is to determine the state of a quantum system as efficiently as possible with suitable measurements and data processing. In practice, the set of accessible measurements is usually determined by experimental settings, which are not easy to modify. In contrast, there is much more freedom in choosing the reconstruction methods, and making a good choice is the first step towards an efficient estimator. In Chapter 3, we focused on state estimation with linear state reconstruction, whose main merit is simplicity. It is a good starting point in theoretical analysis, but not a good choice in practice owing to several obvious defects (see Sec. 2.4.1). The problem of nonpositivity gets less serious as the sample size increases, but the ambiguity in choosing the reconstruction operators persists as long as the measurements are informationally overcomplete, which is the case we are mainly concerned here. The set of canonical reconstruction operators is optimal in linear state tomography, in which the set, once chosen, is independent of the measurement statistics. However, this constraint is by no means a necessity.

To extract maximal information, we need to consider the reconstruction operators that are optimal in the pointwise sense, which generally depend on the measurement results. This problem has been addressed by D'Ariano and Perinotti [73], who derived the set of optimal reconstruction operators with respect to the MSE. The situation is still not clear when the figure of merit is the WMSE corresponding to a generic weight

matrix, such as the MSB. Meanwhile, several basic questions are not well understood. For example, by how much can the efficiency be improved with the optimal reconstruction operators instead of the canonical reconstruction operators? What relations exist between this reconstruction scheme and other schemes, such as the ML method (see Sec. 2.4.2)? What is the efficiency gap between minimal IC measurements and informationally overcomplete measurements?

As a special example of informationally overcomplete measurements, measurements constructed from MUB [83, 155, 272], known as mutually unbiased measurements, are of great interest. The idea of determining quantum states with mutually unbiased measurements was first explored by Ivanović [155], who constructed complete sets of mutually unbiased bases for prime dimensions. Later, Wootters and Fields [272] generalized Ivanović's construction to prime-power dimensions and proved that, among all choices of  $d + 1$  projective measurements, such measurements are optimal in state estimation with respect to a kind of information measure. Recently, Baier and Petz [21] further proved that they are optimal in minimizing the determinant of the MSE matrix. The MSE itself for mutually unbiased measurements was first investigated by Řeháček and Hradil [227], who also explored its connection with an information measure proposed by Brukner and Zeilinger [46]. However, their conclusion that the MSE loses unitary invariance under the ML reconstruction is not well founded. The correct formula for the MSE was later derived by Embacher and Narnhofer [88]. Are mutually unbiased measurements optimal in minimizing the MSE? A definite answer to this persistent open question is highly desirable since the MSE is one of the most popular figures of merit used in practice. Note that the analogous question in the case of minimal IC measurements was settled by Scott [244], who showed that SIC POMs are optimal among all minimal IC measurements (see Sec. 3.2.2).

In this chapter, we determine the set of optimal reconstruction operators in the pointwise sense, using the MSE matrix as a benchmark. Our approach is applicable to studying the WMSE based on any weight matrix and is much simpler than the one in Ref. [73]. It turns out that the resulting reconstruction scheme is equivalent



## 4.2. Optimal state reconstruction

---

to the ML method in the asymptotic limit. Compared with the latter approach, a main merit of our approach is that it is parametrization independent and is thus often much easier to work with. Also, it is illustrative of the differences between linear state reconstruction and optimal state reconstruction since it treats the two alternatives in a unified framework. In addition, our approach is well suited for studying adaptive measurements, which is the main topic of Chapter 5.

As a first application of the above approach, we show that, among all choices of  $d+1$  projective measurements, mutually unbiased measurements are optimal in minimizing the MSE averaged over unitarily equivalent true states, thereby answering the question posed above. Coincidentally, our study leads to a conjecture that singles out SIC POMs and MUB as the only solutions to a state-estimation problem.

Next, we show that covariant measurements are optimal among all nonadaptive measurements in minimizing the WMSE based on any unitarily invariant distance, including the MSE and the MSB. Compared with minimal IC measurements, covariant measurements can improve the tomographic efficiency significantly when the true state has a high purity. Nevertheless, the average scaled MSB diverges at the boundary of the state space in the large-sample limit. This divergence is also present for any WMSE based on a monotone Riemannian metric [30, 216, 219] as long as the measurement is nonadaptive, in sharp contrast with the intuitive belief that states with high purity are easier to estimate than states with low purity.

## 4.2 Optimal state reconstruction for informationally overcomplete measurements

According to Sec. 3.2.1, given an IC POM with outcomes  $\Pi_\xi$ , there exists a set of reconstruction operators  $\Theta_\xi$ , with which an estimator can be constructed from the set of frequencies,  $\hat{\rho} = \sum_\xi f_\xi \Theta_\xi$ . The set of reconstruction operators is not unique when the POM is informationally overcomplete, that is, when the number of outcomes is larger than  $d^2$ . What is the optimal choice? Here we shall determine the set of optimal reconstruction operators in the pointwise sense and show that the resulting

reconstruction scheme is equivalent to the ML method in the asymptotic limit.

#### 4.2.1 Optimal reconstruction in the perspective of frame theory

To derive the set of optimal reconstruction operators, we shall make use of the following lemma, whose proof is relegated to Appendix C.

**Lemma 4.1** *Suppose  $A$  and  $B$  are two  $m \times n$  matrices such that  $AB^\dagger$  is a projector. Then  $AA^\dagger \geq (BB^\dagger)^+$ , and the inequality is saturated if and only if  $A = B^{\dagger+} = (BB^\dagger)^+B$ . If, in addition,  $AB^\dagger = 1$ , then  $AA^\dagger \geq (BB^\dagger)^{-1}$ , and the inequality is saturated if and only if  $A = (BB^\dagger)^{-1}B$ .*

Here  $A^+$  denotes the (Moore-Penrose) pseudoinverse of  $A$  (the arithmetics of pseudoinverses can be found in Ref. [34]).

According to Eq. (3.4), given a set of reconstruction operators  $\Theta_\xi$ , the scaled MSE matrix of the estimator  $\hat{\rho}$  is given by

$$\mathcal{C}(\rho) = \sum_{\xi} |\Theta_\xi\rangle\rangle p_\xi \langle\langle \Theta_\xi | - |\rho\rangle\rangle \langle\langle \rho|. \quad (4.1)$$

Lemma 4.1 applied to  $(|\Theta_1\rangle\rangle p_1^{1/2}, |\Theta_2\rangle\rangle p_2^{1/2}, \dots)$  and  $(|\Pi_1\rangle\rangle p_1^{-1/2}, |\Pi_2\rangle\rangle p_2^{-1/2}, \dots)$  yields

$$\mathcal{C}(\rho) \geq \mathcal{F}(\rho)^{-1} - |\rho\rangle\rangle \langle\langle \rho|, \quad (4.2)$$

where

$$\mathcal{F}(\rho) = \sum_{\xi} |\Pi_\xi\rangle\rangle \frac{1}{p_\xi} \langle\langle \Pi_\xi | \quad (4.3)$$

is also called the frame superoperator, which generalizes the definition in Eq. (3.1).

The inequality is saturated if and only if the reconstruction operators are of the form

$$|\Theta_\xi\rangle\rangle = p_\xi^{-1} \mathcal{F}(\rho)^{-1} |\Pi_\xi\rangle\rangle, \quad (4.4)$$

## 4.2. Optimal state reconstruction

---

in which case we have

$$\mathcal{C}(\rho) = \mathcal{F}(\rho)^{-1} - |\rho\rangle\rangle\langle\langle\rho|, \quad \mathcal{E}(\rho) = \text{Tr}\{\mathcal{F}(\rho)^{-1}\} - \text{tr}(\rho^2). \quad (4.5)$$

Meticulous readers may have noticed that the optimal reconstruction operators depend on the true state, which is usually unknown. To remedy this problem, we may replace the true state in the above formulas with an estimator that derives from another reconstruction scheme, canonical reconstruction for instance.

When  $\rho$  is the completely mixed state, Eq. (4.3) reduces to Eq. (3.1), and it follows that the set of canonical reconstruction operators is optimal, as are the MSE matrix and the MSE associated with the canonical reconstruction. The above analysis implies that the canonical reconstruction is optimal in minimizing the WMSE averaged over unitarily equivalently true states as long as the weight matrix is independent of the true state. When the weight matrix is a constant, our study reproduces the conclusion of Scott [244] (see Sec. 3.2.1).

For the convenience of subsequent discussions, several basic properties of the frame superoperator and the optimal reconstruction operators are listed below,

$$\mathcal{F}(\rho)|\rho\rangle\rangle = |1\rangle\rangle, \quad \mathcal{F}(\rho)^{-1}|1\rangle\rangle = |\rho\rangle\rangle, \quad (4.6a)$$

$$\text{tr}(\Theta_\xi) = 1, \quad (4.6b)$$

$$\sum_{\xi} \text{tr}(\Pi_\xi)\Theta_\xi = 1. \quad (4.6c)$$

Equation (4.6a) follows from the definition of  $\mathcal{F}(\rho)$ ; Eq. (4.6b) can be derived by multiplying both sides of Eq. (4.4) with  $\langle\langle 1|$  and applying Eq. (4.6a); Eq. (4.6c) follows from the assumption  $\sum_{\xi} |\Theta_\xi\rangle\rangle\langle\langle\Pi_\xi| = \mathbf{I}$  and holds for any set of reconstruction operators, regardless whether it is optimal or not.

## Chapter 4. The power of informationally overcomplete measurements

---

According to Eqs. (4.5) and (4.6a),  $|1\rangle\rangle$  is a null eigenvector of  $\mathcal{C}(\rho)$ ; that is,  $\mathcal{C}(\rho)$  is supported on the space of traceless Hermitian operators. Define

$$\bar{\mathcal{F}}(\rho) := \bar{\mathbf{I}}\mathcal{F}(\rho)\bar{\mathbf{I}} = \sum_{\xi} |\bar{\Pi}_{\xi}\rangle\rangle \frac{1}{p_{\xi}} \langle\langle \bar{\Pi}_{\xi}|. \quad (4.7)$$

Then one can show by virtue of Eq. (4.6) that  $\mathcal{C}(\rho)\bar{\mathcal{F}}(\rho) = \bar{\mathbf{I}}$ , which implies that  $\mathcal{C}(\rho)$  is the inverse of  $\bar{\mathcal{F}}(\rho)$  in the space of traceless Hermitian operators,

$$\mathcal{C}(\rho) = \bar{\mathcal{F}}(\rho)^{-1}, \quad \mathcal{E}(\rho) = \text{Tr}\{\bar{\mathcal{F}}(\rho)^{-1}\}. \quad (4.8)$$

Comparison with Eq. (4.5) yields a simple but useful formula,

$$\bar{\mathcal{F}}(\rho)^{-1} = \mathcal{F}(\rho)^{-1} - |\rho\rangle\rangle\langle\langle\rho|. \quad (4.9)$$

In the rest of this section, we briefly discuss the problem of state reconstruction when the measurement is not IC. This problem is also relevant to studying informationally overcomplete measurements, such as mutually unbiased measurements, since many of them are convex combinations of informationally incomplete measurements.

For an informationally incomplete measurement, it is generally impossible to infer the true state accurately even if the sample size is arbitrarily large. Nevertheless, the projection of the true state onto the reconstruction subspace, the space spanned by the  $\Pi_{\xi}$ s, can be determined exactly in the asymptotic limit. Let  $\rho_{\text{R}}$  and  $\mathcal{C}_{\text{R}}(\rho)$  be the projections of the true state and the MSE matrix onto the reconstruction subspace. Then

$$\mathcal{C}_{\text{R}}(\rho) \geq \mathcal{F}(\rho)^+ - |\rho_{\text{R}}\rangle\rangle\langle\langle\rho_{\text{R}}| = \bar{\mathcal{F}}(\rho)^+. \quad (4.10)$$

The inequality is saturated if and only if the reconstruction operators are given by

$$|\Theta_{\xi}\rangle\rangle = p_{\xi}^{-1} \mathcal{F}(\rho)^+ |\Pi_{\xi}\rangle\rangle, \quad (4.11)$$

when restricted to the reconstruction subspace.

## 4.2. Optimal state reconstruction

---

To illustrate the above idea, let us consider a rank-one projective measurement for example. Noticing that the  $\Pi_\xi$ s are orthogonal projectors and  $\rho_R = \sum_\xi p_\xi \Pi_\xi$ , we have

$$\mathcal{C}_R(\rho) = \sum_\xi |\Pi_\xi\rangle\rangle p_\xi \langle\langle \Pi_\xi| - \sum_{\xi, \zeta} |\Pi_\xi\rangle\rangle p_\xi p_\zeta \langle\langle \Pi_\zeta|, \quad \mathcal{E}_R(\rho) = 1 - \sum_\xi p_\xi^2. \quad (4.12)$$

### 4.2.2 Connection with the maximum-likelihood method

To see the connection between the optimal reconstruction scheme presented in the previous section and the ML method [152, 208] (see Sec. 2.4.2), it is convenient to adopt the affine parametrization in Eq. (2.19). According to Sec. 2.5, the Fisher information matrix takes on the form

$$I_{jk} = \sum_\xi \frac{\langle\langle E_j | \Pi_\xi \rangle\rangle \langle\langle \Pi_\xi | E_k \rangle\rangle}{p_\xi} = \langle\langle E_j | \mathcal{F}(\rho) | E_k \rangle\rangle = \langle\langle E_j | \bar{\mathcal{F}}(\rho) | E_k \rangle\rangle. \quad (4.13)$$

Therefore, the frame superoperator  $\bar{\mathcal{F}}(\rho)$  is essentially the Fisher information matrix in disguise, and state reconstruction with optimal reconstruction operators is equivalent to the ML method in the large- $N$  limit. Recall that the MSE matrix of any unbiased estimator is lower bounded by the inverse of the Fisher information matrix and that the lower bound can be saturated asymptotically with the ML estimator [68, 93, 94, 224] (see Sec. 2.5).

Alternatively, we can elucidate this point by inspecting the likelihood functional in the large- $N$  limit. According to Eq. (2.6),

$$\begin{aligned} \frac{1}{N} \ln \mathcal{L}(\rho) &= \sum_\xi f_\xi \ln p_\xi \approx \sum_\xi f_\xi \ln f_\xi - \frac{1}{2} \sum_\xi \frac{(p_\xi - f_\xi)^2}{f_\xi} \\ &\approx \sum_\xi f_\xi \ln f_\xi - \frac{1}{2} \sum_\xi \frac{(p_\xi - f_\xi)^2}{p_\xi}. \end{aligned} \quad (4.14)$$

Suppose the likelihood functional is maximized at  $\tilde{\theta}$ . Let  $\Delta\theta = \theta - \tilde{\theta}$ ; then

$$\frac{1}{N} \ln \mathcal{L}(\rho) \approx c - \frac{1}{2} \sum_{j,k} \Delta\theta_j \Delta\theta_k \langle\langle E_j | \bar{\mathcal{F}}(\rho) | E_k \rangle\rangle, \quad (4.15)$$

where  $c$  is a constant. Again,  $\bar{\mathcal{F}}(\rho)$  plays the role of the Fisher information matrix.

Compared with the ML method, our approach is free from the distraction due to the parametrization, which is somehow arbitrary, and is thus often easier to work with. It is also illustrative of the differences between linear state reconstruction and optimal state reconstruction. In addition, it is well suited for studying adaptive measurements, as we shall see in Chapter 5. The drawback of our approach is that the optimal reconstruction operators need to be chosen adaptively, and it is not easy to take into account naturally the positivity constraint on the density operators. Depending on the situation, one alternative may be more suitable than the other, and a judicious choice may greatly simplify the discussion.

### 4.3 Quantum state estimation with mutually unbiased measurements

Two bases  $\{|\psi_j\rangle\}$  and  $\{|\phi_j\rangle\}$  are mutually unbiased if all the transition probabilities  $|\langle\psi_j|\phi_k\rangle|^2$  across their basis elements are equal to  $1/d$  [83, 155, 272]. In a  $d$ -dimensional Hilbert space, there exist at most  $d+1$  MUB; such a maximal set, if it exists, is called a complete set of MUB. When  $d$  is a prime power, a complete set of MUB can be constructed explicitly [83, 155, 272]; all known constructions rely on the existence of Galois fields, which admit no generalization to any other dimension. It is believed that no complete set of MUB can exist for any dimension that is not a prime power, although no rigorous proof is known even for dimension 6, the smallest candidate. Since their discovery, MUB have found numerous applications, such as in the determination of quantum states, in the study of quantum kinematics, and in the construction of generalized Bell states (see Ref. [83] for a review).

Two (rank-one) projective measurements are mutually unbiased if their measurement bases are mutually unbiased. Such measurements are particularly interesting because of their optimality properties for quantum state estimation. According to parameter counting,  $d+1$  projective measurements are needed for a complete determination of a  $d$ -level quantum system. Wootters and Fields showed that complete sets of

### 4.3. Quantum state estimation with mutually unbiased measurements

mutually unbiased measurements are optimal in the sense of maximizing the information gain [272]. Recently, Baier and Petz further demonstrated that such measurements are optimal in minimizing the determinant of the MSE matrix [21]. However, it is not known whether they are optimal in minimizing the MSE, which is much more prevalent as a figure of merit, although the MSE itself has been determined by Embacher and Narnhofer [88] (see also Ref. [227]).

In this section, we show that mutually unbiased measurements are optimal in minimizing the MSE averaged over unitarily equivalent true states. We then reveal an interesting connection between SIC POMs and MUB with a state-estimation problem.

Consider state estimation using  $d + 1$  projective measurements  $\{\Pi_{jk}\}_k$ , each with probability  $p_j$ , where the  $\Pi_{jk}$ s ( $j = 0, 1, \dots, d; k = 1, 2, \dots, d$ ) for given  $j$  are normalized projectors of the  $j$ th measurement. The Fisher information matrix can be written as

$$\bar{\mathcal{F}}(\rho) = \sum_{j=0}^d p_j \bar{\mathcal{F}}_j(\rho), \quad \bar{\mathcal{F}}_j(\rho) = \sum_{k=1}^d |\bar{\Pi}_{jk}\rangle\rangle \frac{1}{p_{jk}} \langle\langle \bar{\Pi}_{jk}|, \quad (4.16)$$

where  $p_{jk} = \text{tr}(\rho \Pi_{jk})$ . Let  $\mathbf{v}$  denote the row vector of the  $d^2 - 1$  eigenvalues of  $\bar{\mathcal{F}}(\rho)^{\dagger}$  and  $\mathbf{v}_j$  of the  $d - 1$  nonzero eigenvalues of  $\bar{\mathcal{F}}_j(\rho)$ . Then the vector  $\mathbf{v}' := (p_0 \mathbf{v}_0, p_1 \mathbf{v}_1, \dots, p_d \mathbf{v}_d)$  derived from the  $\mathbf{v}_j$ s is majorized by  $\mathbf{v}$  according to Theorem G.1.b in page 242 of Ref. [188]; moreover, the equality  $\mathbf{v}'^{\downarrow} = \mathbf{v}^{\downarrow}$  is attained if and only if all the Fisher information matrices  $\bar{\mathcal{F}}_j$  have mutually orthogonal supports or, equivalently, if the  $d + 1$  measurement bases are mutually unbiased. Consequently, we have

$$\mathcal{E}(\rho) = \text{Tr}\{\bar{\mathcal{F}}(\rho)^{-1}\} \geq \sum_{j=0}^d \frac{1}{p_j} \text{Tr}\{\bar{\mathcal{F}}_j(\rho)^{\dagger}\} = \sum_{j=0}^d \frac{1}{p_j} \left(1 - \sum_{k=1}^d p_{jk}^2\right), \quad (4.17)$$

and the bound is saturated if and only if the measurements are mutually unbiased since the trace of the inverse of a matrix is strictly Schur convex in its eigenvalues [30, 188].

For mutually unbiased measurements, the scaled MSE matrix and the reconstruc-

---

<sup>1</sup>For operators that are supported on the space of traceless Hermitian operators, the null eigenvector  $|1\rangle\rangle$  and the corresponding eigenvalue are omitted for simplicity.

tion operators are respectively given by

$$\begin{aligned}\mathcal{C}(\rho) &= \sum_{j=0}^d \frac{1}{p_j} \bar{\mathcal{F}}_j(\rho)^+ = \sum_{j=0}^d \frac{1}{p_j} \left( \sum_{k=1}^d |\Pi_{jk}\rangle\rangle p_{jk} \langle\langle \Pi_{jk}| - |\rho_j\rangle\rangle \langle\langle \rho_j| \right), \\ \Theta_j &= \frac{1}{p_j} (\Pi_{jk} - \rho_j) + \rho,\end{aligned}\tag{4.18}$$

where  $\rho_j = \sum_{k=1}^d p_{jk} \Pi_{jk}$  is the projection of  $\rho$  onto the  $j$ th reconstruction subspace.

Taking the average of Eq. (4.17) over unitarily equivalent states yields

$$\overline{\mathcal{E}(\rho)} \geq \sum_{j=0}^d \frac{1}{p_j} \left( 1 - \sum_{k=1}^d \overline{p_{jk}^2} \right) = \frac{d - \text{tr}(\rho^2)}{d+1} \sum_{j=0}^d \frac{1}{p_j} \geq (d+1)[d - \text{tr}(\rho^2)],\tag{4.19}$$

where we have applied the formula  $\overline{p_{jk}^2} = [1 + \text{tr}(\rho^2)]/[d(d+1)]$ . The lower bound is saturated if and only if all  $d+1$  bases are selected with the same probability  $1/(d+1)$ .

In that case, the scaled MSE is unitarily invariant [88],

$$\mathcal{E}(\rho) = \overline{\mathcal{E}(\rho)} = (d+1)[d - \text{tr}(\rho^2)],\tag{4.20}$$

where we have applied the formula  $\sum_{j=0}^d \sum_{k=1}^d p_{jk}^2 = 1 + \text{tr}(\rho^2)$ , noting that a complete set of MUB forms a 2-design [167, 232, 244] (see Appendix B). Therefore, mutually unbiased measurements are optimal not only in minimizing the average MSE but also in the minimax sense. Similar analysis also applies to other figures of merit that are either Schur convex or Schur concave in the eigenvalues of the MSE matrix. For example, the average volume of the uncertainty ellipsoid, as quantified by  $\overline{\ln \det \mathcal{C}(\rho)}$ , is minimized by mutually unbiased measurements.

A lower MSE than Eq. (4.20) can be achieved if adaptive measurements are accessible. Suppose we can rotate the set of MUB simultaneously and vary the probability with which each basis is measured. Then the minimal scaled MSE is given by [88]

$$\mathcal{E}(\rho) = \sum_{j=0}^d \frac{a_j}{p_j} \geq \left( \sum_{j=0}^d \sqrt{a_j} \right)^2 \geq (\sqrt{1 - \text{tr}(\rho^2)} + \sqrt{d^2 - d})^2,\tag{4.21}$$



### 4.3. Quantum state estimation with mutually unbiased measurements

---

where  $a_j := 1 - \sum_{k=1}^d p_{jk}^2$  satisfy the set of constraints  $1 - \text{tr}(\rho^2) \leq a_j \leq 1 - 1/d$  and  $\sum_{j=0}^d a_j = d - \text{tr}(\rho^2)$ . The first inequality in Eq. (4.21) is saturated if  $p_j = \sqrt{a_j}/(\sum_{l=0}^d \sqrt{a_l})$ ; the second one if  $a_0 = 1 - \text{tr}(\rho^2)$  and  $a_1 = a_2 \cdots = a_d = 1 - 1/d$ . In the optimal measurement scheme, the eigenbasis of  $\rho$  is measured with the least probability, and the remaining  $d$  bases are measured with an equal probability.

Compared with the MSE achieved by a SIC POM [244] (see Sec. 3.2.2), the MSE achieved by mutually unbiased measurements with equal probability is slightly smaller. With optimal state reconstruction, MUB are more efficient than SIC POMs, in contrast with the scenario in linear state reconstruction. A common nice feature of the two measurement schemes is that their MSEs are unitarily invariant; a measurement with this property is called a *balanced* measurement. The property of balance of mutually unbiased measurements is crucial to the introduction of the operationally invariant information by Brukner and Zeilinger [46]. It is also a desirable feature in representing quantum states with probabilities [269, 270].

Surprisingly, SIC measurements and complete sets of mutually unbiased measurements are the only known balanced measurements with finite rank-one outcomes, assuming that no two outcomes are proportional to each other. The covariant measurement is another example of balanced measurements if we allow an infinite number of outcomes, but this example is not so interesting. To appreciate the difficulty of constructing balanced measurements, it is worth noting that a convex combination of two balanced measurements is generally not balanced (see Sec. 4.5 for examples), contrary to the intuition of many people.

**Conjecture 4.2** *A rank-one IC measurement with finite number of outcomes is balanced if and only if it is a SIC POM or it is composed of a complete set of mutually unbiased measurements with equal probability.*

This conjecture picks SIC POMs and MUB as the only solutions to a state-estimation problem. It would be really remarkable if they can be connected in such a peculiar manner.

## 4.4 Efficiency of covariant measurements

In this section we investigate the efficiency gap between minimal IC measurements and informationally overcomplete measurements, as well as the limitation of nonadaptive measurements. As we shall see shortly, covariant measurements play a crucial role in understanding the efficiencies of informationally overcomplete measurements, although it is not practical to implement them. Previously, most studies on covariant measurements focused on pure-state models [130].

Suppose  $\bar{\mathcal{F}}_1(\rho)$  and  $\bar{\mathcal{F}}_2(\rho)$  are the Fisher information matrices of two given IC measurements. If the two measurements are performed with probabilities  $p_1$  and  $p_2 = 1 - p_1$ , then the Fisher information matrix is a convex combination,

$$\bar{\mathcal{F}}(\rho) = p_1 \bar{\mathcal{F}}_1(\rho) + p_2 \bar{\mathcal{F}}_2(\rho). \quad (4.22)$$

Noticing the operator convexity of the function  $1/x$  over the interval  $(0, \infty)$ , we have

$$\mathcal{C}(\rho) \leq p_1 \mathcal{C}_1(\rho) + p_2 \mathcal{C}_2(\rho), \quad \mathcal{E}(\rho) \leq p_1 \mathcal{E}_1(\rho) + p_2 \mathcal{E}_2(\rho), \quad (4.23)$$

which implies that  $\overline{\mathcal{E}(\rho)} \leq p_1 \overline{\mathcal{E}_1(\rho)} + p_2 \overline{\mathcal{E}_2(\rho)}$ . In particular,  $\overline{\mathcal{E}(\rho)} \leq \overline{\mathcal{E}_1(\rho)} = \overline{\mathcal{E}_2(\rho)}$  if the two given measurements are unitarily equivalent. In other words, the average MSE never increases by combining unitarily equivalent measurements. Therefore, it is minimized by the covariant measurement. By the same token, so is the average WMSE based on any unitarily invariant distance, such as the Bures distance.

The frame superoperator for the covariant measurement is given by

$$\mathcal{F}(\rho) = d \int d\mu(\psi) \frac{1}{\langle \psi | \rho | \psi \rangle} |\Pi_\psi\rangle\rangle \langle\langle \Pi_\psi|, \quad (4.24)$$

where  $\Pi_\psi = |\psi\rangle\langle\psi|$  and  $d\mu(\psi)$  is the normalized Haar measure. To illustrate the dependence of  $\mathcal{F}(\rho)$  and  $\mathcal{E}(\rho)$  on the true state, we shall consider those states that are

#### 4.4. Efficiency of covariant measurements

---

convex combinations of the completely mixed state and a projector state of rank  $r$ ,

$$\rho_r(s) = \frac{s}{r} \sum_{j=1}^r |j\rangle\langle j| + (1-s)\frac{1}{d}, \quad 1 \leq r \leq d-1, \quad 0 \leq s \leq 1. \quad (4.25)$$

Calculation shows that  $\mathcal{F}(\rho_r(s))$  has the form

$$\begin{aligned} \mathcal{F}(\rho_r(s)) = & a \sum_{j \neq k=1}^r |E_{jk}\rangle\rangle\langle\langle E_{jk}| + b \sum_{j=1}^r \sum_{k=r+1}^d (|E_{jk}\rangle\rangle\langle\langle E_{jk}| + |E_{kj}\rangle\rangle\langle\langle E_{kj}|) \\ & + c \sum_{j \neq k=r+1}^d |E_{jk}\rangle\rangle\langle\langle E_{jk}| + \sum_{j,k=1}^d \mathcal{M}_{jk} |E_{jj}\rangle\rangle\langle\langle E_{kk}|, \end{aligned} \quad (4.26)$$

where  $E_{jk} = |j\rangle\langle k|$  [see Eq. (2.20)] and

$$\mathcal{M}_{jk} = \begin{cases} (1 + \delta_{jk})a & \text{if } 1 \leq j, k \leq r, \\ (1 + \delta_{jk})c & \text{if } r+1 \leq j, k \leq d, \\ b & \text{otherwise.} \end{cases} \quad (4.27)$$

The three parameters  $a = g_{20}$ ,  $b = g_{11}$ , and  $c = g_{02}$  are determined by the integral

$$g_{jk} = \frac{2dr\Gamma(d+1)}{\Gamma(r+j)\Gamma(d-r+k)} \int_0^{\pi/2} d\alpha \frac{(\cos \alpha)^{2r-1+2j} (\sin \alpha)^{2d-2r-1+2k}}{ds(\cos \alpha)^2 + r(1-s)}, \quad (4.28)$$

which can be evaluated by applying the formula

$$\int_0^{\pi/2} d\alpha \frac{\cos \alpha (\sin \alpha)^{2m+1}}{(\cos \alpha)^2 + u} = \frac{1}{2} (1+u)^m \ln \frac{1+u}{u} - \frac{1}{2} \sum_{n=0}^{m-1} \frac{(1+u)^n}{m-n}, \quad u > 0 \quad (4.29)$$

after replacing  $(\cos \alpha)^2$  with  $1 - (\sin \alpha)^2$ . The Fisher information matrix  $\bar{\mathcal{F}}(\rho_r(s))$  has the same form as  $\mathcal{F}(\rho_r(s))$ , except that  $\mathcal{M}$  is replaced by  $\bar{\mathcal{M}} := \bar{\mathbf{I}}\mathcal{M}\bar{\mathbf{I}}$ .

Calculation shows that  $\bar{\mathcal{M}}$  has  $r-1$  eigenvalues equal to  $a$ ,  $d-r-1$  eigenvalues equal to  $c$ , and one eigenvalue equal to

$$\beta = \frac{(r+1)(d-r)a + r(d-r+1)c - 2r(d-r)b}{d}. \quad (4.30)$$

Note that  $E_{jk}$  for  $j \neq k$  is an eigenvector of  $\mathcal{F}$  and  $\bar{\mathcal{F}}$ , and that the common eigenvalue

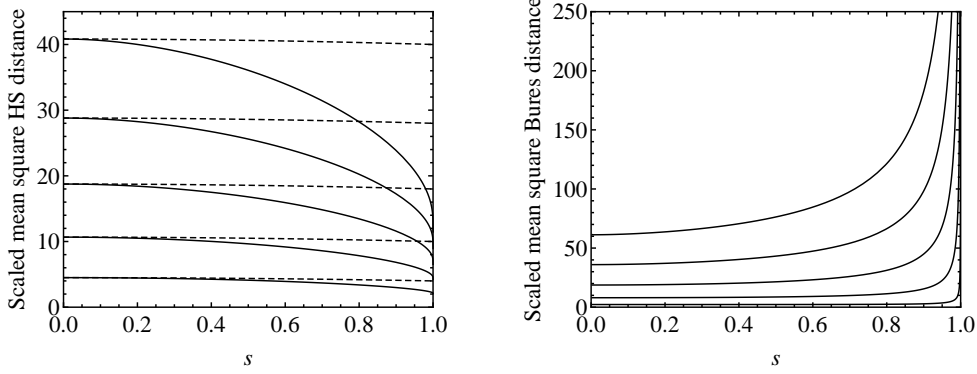


Figure 4.1: The scaled MSE (with respect to the HS distance, left plot) and the scaled MSB (right plot) of the covariant measurements when the true states have the form in Eq. (4.25) with  $r = 1$  and  $d = 2, \dots, 6$  (from bottom to top). For comparison, the dashed lines in the left plot show the scaled MSE of the optimal linear or minimal tomography.

is one of the three choices  $a, b, c$  depending on the values of  $j$  and  $k$ . Therefore,  $\bar{\mathcal{F}}$  has four distinct eigenvalues  $a, b, c$ , and  $\beta$  with multiplicities  $r^2 - 1$ ,  $2r(d - r)$ ,  $(d - r)^2 - 1$ , and 1, respectively.

According to Eq. (4.8), the scaled MSE reads

$$\mathcal{E}(\rho_r(s)) = \frac{r^2 - 1}{a} + \frac{2r(d - r)}{b} + \frac{(d - r)^2 - 1}{c} + \frac{1}{\beta}. \quad (4.31)$$

The scaled MSB can be calculated by means of Eq. (A.9), with the result

$$\mathcal{E}_{\text{SB}}(\rho_r(s)) = \frac{1}{4} \left( \frac{r^2 - 1}{a\lambda_1} + \frac{4r(d - r)}{b(\lambda_1 + \lambda_2)} + \frac{(d - r)^2 - 1}{c\lambda_2} + \frac{d - r}{d\beta\lambda_1} + \frac{r}{d\beta\lambda_2} \right), \quad (4.32)$$

where  $\lambda_1 = (s/r) + (1 - s)/d$  and  $\lambda_2 = (1 - s)/d$  are the two distinct eigenvalues of  $\rho$ . Figure 4.1 illustrates the scaled MSE and MSB in the case  $r = 1$  and  $d = 2, 3, 4, 5, 6$ . Compared with linear state tomography or minimal state tomography, optimal state estimation with covariant measurements can improve the efficiency significantly when the true state has a high purity. Nevertheless, the efficiency is still too limited to be satisfactory when the scaled MSB is chosen as the figure of merit.

As  $s$  approaches unity,  $\rho_r(s)$  turns into a subnormalized projector of rank  $r$ . When

#### 4.4. Efficiency of covariant measurements

---

$r \geq 2$ , the three parameters  $a, b, c$  have well-defined limits  $a = r/(r+1), b = 1, c = r/(r-1)$ , and so does the MSE,

$$\mathcal{E}(\rho_r(1)) = d^2 + 2d - 1 - \frac{d^2}{r} - \frac{1}{r}. \quad (4.33)$$

When  $r = 1$ , the parameters  $a$  and  $b$  still have well-defined limits, whereas  $c$  diverges as  $\ln[d/(1-s)]$ . The formula for the MSE is still applicable, except that the derivative of  $\mathcal{E}(\rho_r(s))$  with respect to  $s$  can diverge. In the pure-state limit, the scaled MSE  $2(d-1)$  achieved by the covariant measurement is equal to the corresponding value for the pure-state model. Furthermore, it is minimal not only in the Bayesian sense but also in the pointwise sense since it saturates a quantum analog of the CR bound [130, 193]. Compared with the scaled MSE  $d^2 + d - 2$  [see Eq. (3.12)] that is achievable with the optimal linear state tomography, it is smaller by  $(d+2)/2$  times.

In sharp contrast, the scaled MSB diverges in the limit  $s \rightarrow 1$ . This seemingly surprising phenomenon can be explained as follows: The entries of  $\bar{\mathcal{F}}$  are either finite or logarithmically divergent in this limit, while the entries of the weight matrix diverges much more quickly according to Eq. (A.9). Recalling that the covariant measurement minimizes the average scaled MSB among all nonadaptive measurements, we conclude that the average scaled MSB diverges at the boundary of the state space for all nonadaptive measurements. From the Bayesian perspective, our analysis implies that the MSB generally decreases more slowly than the scaling law  $1/N$  that is expected from common statistical consideration. For single qubit, this phenomenon was noticed in Ref. [20]. The same conclusion also holds for any WMSE based on a monotone Riemannian metric since the Bures metric is minimal among all such metrics [30, 216, 219]. This observation reveals a severe limitation of nonadaptive measurements and motivates us to study adaptive measurements, which is the main subject matter of Chapter 5.

In the pure-state limit, the scaled MSE matrix can be determined based on

Eqs. (4.8) and (4.26), with the result

$$\mathcal{C}(|1\rangle\langle 1|) = \sum_{j=2}^d (|E_{1j}^+\rangle\langle E_{1j}^+| + |E_{1j}^-\rangle\langle E_{1j}^-|), \quad (4.34)$$

which is a rank- $2(d-1)$  projector. Accordingly, the scaled deviation  $\Delta\rho$  has the form

$$\Delta\rho = \frac{1}{\sqrt{2}} \sum_{j=2}^d [(x_j - iy_j)|1\rangle\langle j| + (x_j + iy_j)|j\rangle\langle 1|], \quad (4.35)$$

where  $x_j, y_j$  follow a  $2(d-1)$ -dimensional standard isotropic Gaussian distribution. Since  $\Delta\rho$  has only two nonzero eigenvalues  $\pm\sqrt{\sum_{j=2}^d (x_j^2 + y_j^2)}/2$ , its trace norm is proportional to the HS norm; namely,  $\|\Delta\rho\|_{\text{tr}} = \|\Delta\rho\|_{\text{HS}}/\sqrt{2}$ . The scaled mean trace distance reads

$$\mathcal{E}_{\text{tr}}(\rho) = \frac{1}{\sqrt{2}} \mathcal{E}_{\text{HS}}(\rho) = \frac{\Gamma(d - \frac{1}{2})}{\Gamma(d-1)} \approx \sqrt{d-1}. \quad (4.36)$$

In contrast to the result achievable with linear or minimal tomography [see Eq. (3.20)], it is approximately smaller by a factor of  $4d/3\pi$  when  $d \gg 2$ . The improvement of informationally overcomplete measurements is more dramatic compared with the scenario in which the MSE serves as the figure of merit.

## 4.5 Informationally overcomplete measurements on the two-level system

In this section, we study the efficiencies of the covariant measurement and measurements constructed out of platonic solids in qubit state estimation. There are already many studies on this subject [51, 183, 226], but most theoretical works are based on numerical calculations. We have derived several analytical results for linear state tomography in Sec. 3.3.4 (see also Ref. [281]). Here we shall focus on optimal state reconstruction.

Suppose the qubit state is parameterized by the Bloch vector  $\mathbf{s} = (x, y, z)$ . Then

#### 4.5. Informationally overcomplete measurements on the two-level system

the parameters  $a, b, c$  in Eq. (4.26) and  $\beta$  in Eq. (4.30) are given by

$$\begin{aligned} a &= \frac{2s(-1 + 2s) + (1 - s)^2 \ln\left(\frac{1+s}{1-s}\right)}{4s^3}, & b &= \frac{2s - (1 - s^2) \ln\left(\frac{1+s}{1-s}\right)}{2s^3}, \\ c &= \frac{-2s(1 + 2s) + (1 + s)^2 \ln\left(\frac{1+s}{1-s}\right)}{4s^3}, & \beta &= \frac{-2s + \ln\left(\frac{1+s}{1-s}\right)}{s^3}. \end{aligned} \quad (4.37)$$

The Fisher information matrix of the covariant measurement takes on the form

$$\bar{\mathcal{F}}(\rho) = b\bar{\mathbf{I}} + \frac{1}{2}(\beta - b)|\mathbf{s} \cdot \boldsymbol{\sigma}\rangle\langle\mathbf{s} \cdot \boldsymbol{\sigma}|. \quad (4.38)$$

In terms of the Bloch vector, it reads

$$I(\mathbf{s}) = \frac{1}{2}[b\mathbf{I}_3 + (\beta - b)\mathbf{s}\mathbf{s}]. \quad (4.39)$$

The scaled MSE (with respect to the HS distance) and MSB follows from Eqs. (4.31) and (4.32),

$$\begin{aligned} \mathcal{E}(\rho) &= \frac{s^3}{-2s + \ln\left(\frac{1+s}{1-s}\right)} + \frac{4s^3}{2s + (-1 + s^2) \ln\left(\frac{1+s}{1-s}\right)}, \\ \mathcal{E}_{\text{SB}}(\rho) &= \frac{s^3}{2(1 - s^2)[-2s + \ln\left(\frac{1+s}{1-s}\right)]} + \frac{2s^3}{2s + (-1 + s^2) \ln\left(\frac{1+s}{1-s}\right)}. \end{aligned} \quad (4.40)$$

The scaled MSB diverges in the pure-state limit, as explained in Sec. 4.4.

Following the convention in Sec. 3.3.4, to each platonic solid inscribed on the Bloch sphere, there corresponds a measurement. The scaled MSEs of the measurements constructed from the tetrahedron, octahedron, and cube are respectively given by

$$\begin{aligned} \mathcal{E}^{\text{SIC}}(\rho) &= \frac{9 - s^2}{2}, \\ \mathcal{E}^{\text{MUB}}(\rho) &= \frac{3(3 - s^2)}{2}, \\ \mathcal{E}^{\text{Cube}}(\rho) &= \frac{27 - 18s^2 + s^4 + 2(x^4 + y^4 + z^4)}{2(3 - s^2)}. \end{aligned} \quad (4.41)$$

Here we assume that the octahedron and the cube take the standard orientation, and the tetrahedron is inscribed on the cube. The MSE is unitarily invariant for the SIC

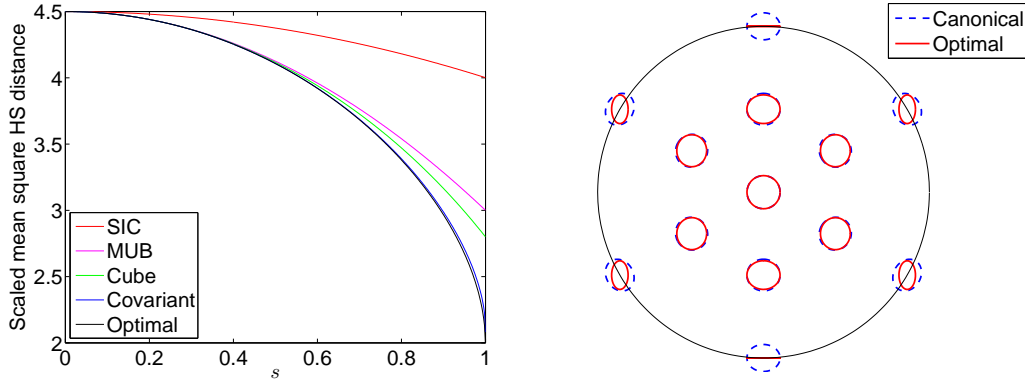


Figure 4.2: Left plot: The average scaled MSE (with respect to the HS distance) in qubit state estimation with the SIC, MUB, cube, and covariant measurements (from top to bottom). The scaled MSE of the optimal adaptive strategy [88, 107, 129, 137] is also shown for comparison. Right plot: Uncertainty ellipses of the marginal distributions on the  $x$ - $z$  plane of the Bloch ball associated with mutually unbiased measurements on 300 copies of the true states. The canonical reconstruction and the optimal reconstruction are compared. The optimal reconstruction reduces the size of the uncertainty ellipses at the prize of losing the covariance property.

(tetrahedron) measurement and the MUB (octahedron) measurement, as mentioned in Sec. 4.3. This is not the case for the cube measurement, although it is a combination of two tetrahedron measurements and is seemingly more symmetric than a single tetrahedron measurement. This observation provides some evidence in favor of Conjecture 4.2. For given  $s$ , the minimal scaled MSE  $(9 - s^2)(9 - 5s^2)/6(3 - s^2)$  is attained when  $\mathbf{s}$  is parallel to one of the diagonals of the cube, and the maximum  $3(3 - s^2)/2$  is attained when  $\mathbf{s}$  is parallel to one of the axes. The average is  $(135 - 90s^2 + 11s^4)/10(3 - s^2)$ . The formulas for the MSEs of the dodecahedron measurement and icosahedron measurement are too complicated to convey a clear meaning; suffice it to mention that the MSEs are not unitarily invariant in both cases, as in the case of the cube measurement.

The left plot of Fig. 4.2 shows the average scaled MSEs in qubit state estimation with the SIC, MUB, cube, and covariant measurements in conjunction with the optimal state reconstruction (without considering the correction due to the boundary). The efficiencies of the MUB, cube, and covariant measurements are higher than that of the SIC measurement, in contrast to the scenario in linear state reconstruction, in which they are equally efficient. Comparison with the MSE achieved by the optimal adaptive



## 4.6. Summary

---

strategy [88, 107, 129, 137] shows that the covariant measurement is almost optimal in the pointwise sense. However, it should be noted that this is generally not the case with respect to other figures of merit, such as the MSB. Also, the situation can be very different beyond the two-level system, as we shall see in Chapter 5.

To visualize the difference between the canonical reconstruction and the optimal reconstruction, let us take the MUB measurement as an example. The scaled Fisher information matrix assumes the form

$$I(\mathbf{s}) = \frac{1}{3} \operatorname{diag}\left(\frac{1}{1-x^2}, \frac{1}{1-y^2}, \frac{1}{1-z^2}\right). \quad (4.42)$$

The scaled MSE matrix for the optimal reconstruction is given by

$$C(\mathbf{s}) = 3 \operatorname{diag}(1-x^2, 1-y^2, 1-z^2). \quad (4.43)$$

It is smaller than the MSE matrix  $3\mathbf{I}_3 - \mathbf{s}\mathbf{s}$  for the canonical reconstruction [see Eq. (3.30)], but is no longer invariant under unitary transformations of the measurement outcomes. The differences between the two reconstruction methods are clearly reflected in the uncertainty ellipses, as illustrated in Fig. 4.2. The situation is quite similar for measurements constructed from other platonic solids, except for the tetrahedron.

## 4.6 Summary

We have studied optimal state reconstruction in the case of informationally overcomplete measurements from the perspective of frame theory and determined the set of optimal reconstruction operators in the pointwise sense. The resulting reconstruction scheme was shown to be equivalent to the ML method in the asymptotic limit.

Based on this approach, we proved that, among all choices of  $d + 1$  projective measurements, mutually unbiased measurements are optimal not only in minimizing the average MSE but also in minimizing the maximal MSE over unitarily equivalent true states. In addition, we introduced the concept of balanced measurements, thereby connecting SIC POMs and MUB in a peculiar way.

## Chapter 4. The power of informationally overcomplete measurements

---

Furthermore, we showed that the covariant measurement is optimal among all non-adaptive measurements in minimizing the WMSE based on any unitarily invariant distance, including the MSE and the MSB. Informationally overcomplete measurements can improve the tomographic efficiency significantly when the states of interest have high purity. Nevertheless, the average scaled MSB diverges at the boundary of the state space in the large-sample limit. And the same is true for the WMSE based on any monotone Riemannian metric as long as the measurement is nonadaptive. On the one hand, this observation breaks down the intuitive belief that states with high purity are easier to estimate than those with low purity. On the other hand, it motivates the study of more sophisticated estimation strategies based on adaptive measurements and collective measurements, which are the highlights of the next two chapters.

# Optimal state estimation with adaptive measurements

---

## 5.1 Introduction

A good state-estimation strategy entails judicial choices on both measurement schemes and data processing methods. Given the measurement results, the optimization of data processing is basically a subject of classical statistical inference, although due modifications are necessary to account for additional constraints, such as the positivity of the density matrices. When the sample is reasonably large, the quality of the estimator is usually quantified by the MSE matrix, which is determined by the Fisher information matrix [94] through the Cramér–Rao (CR) bound [68, 224].

The main challenge in quantum state estimation is to devise the measurements that yield the most information. The set of accessible measurements is usually determined by experimental settings as well as basic principles of quantum mechanics. The simplest choices are independent and identical measurements studied in Chapters 3 and 4; more sophisticated choices, such as adaptive measurements and collective measurements, are the focus of this chapter and the next chapter. The importance of studying these alternatives can be explicated in three aspects. First, as we have seen in Chapter 4, nonadaptive measurements are quite inefficient in many scenarios; adaptive measurements and collective measurements are generally much more efficient, as we shall see later. Thanks to the advance of technology, adaptive measurements have already been realized in experiments [15, 143], and certain collective measurements are also accessible to present experimentalists. Therefore, these measurements are promising alternatives

for reducing quantum resources in practice. Second, the choice of measurements is the main difference between quantum state estimation and classical state estimation, which underlies the difference between quantum information processing and classical information processing. A better understanding of these measurements can help elucidate the peculiar features of quantum information processing. Third, the efficiencies of these more sophisticated measurements embody the characteristics of quantum mechanics, and thus can serve as a window for inspecting foundational issues, such as the complementarity principle, the uncertainty relations, and the geometry of quantum states.

The development of quantum estimation theory has had a convoluted journey. In the late 1960s, Helstrom [139, 140, 141, 142] derived quantum analogs of the Fisher information matrix and the CR bound based on the symmetric logarithmic derivative (SLD) and solved the optimization problem in the one-parameter setting, in which case the bound is tight. It turns out that the local optimal estimation strategies can be realized with only individual measurements. To achieve the global optimal performance, it suffices to implement the local optimal measurements after a localization procedure, following the spirit of two-step adaptive schemes [28, 107, 202]. Therefore, collective measurements do not help in the one-parameter setting in the asymptotic limit. Note that the quantum Fisher information matrix is additive. Incidentally, Braunstein and Caves [45] later showed that the optimal estimation strategy defines a statistical metric in the state space that is equivalent to the Bures metric, in the same sense as the optimal strategy in classical statistical inference defines the Fisher-Rao metric in the probability simplex [63, 94, 224, 268]. Similar ideas also played a crucial role in studying general monotone Riemannian metrics on the state space [30, 216, 219].

The problem in the multiparameter setting turned out to be much more challenging. The SLD bound generally cannot be saturated since the optimal measurements corresponding to different parameters are usually incompatible. As a consequence, it is quite difficult to determine the optimal estimation efficiency with either separable measurements or collective measurements, not to say their efficiency gaps. Great ef-

## 5.1. Introduction

---

forts have been directed to find better lower bounds for the MSE, notable examples including the RLD bound [274], the Holevo bound [147], and the Nagaoka bound [201]. Unfortunately, these bounds generally cannot be saturated under separable measurements; actually they are often quite loose, as we shall see later. The optimal estimation strategies are known only for a few special examples, such as estimating the complex amplitude of a coherent signal in the Gaussian noise [274], the mean values of Gaussian states [147], and the states of the two-level system [107, 129].

To devise a good estimation strategy, it is indispensable to take into account the information trade-off among different parameters. About a decade ago, a promising step along this direction was initiated by Gill and Massar [107], who introduced an inequality about the Fisher information matrix, which succinctly summarizes such trade-off. In a sense, the Gill–Massar (GM) inequality generalizes the one derived by Englert [90] concerning the fringe visibility and the which-way information, which is a quantitative manifestation of the complementarity principle [41]. By means of this inequality, they derived a general lower bound, the GM bound, for the WMSE that is applicable to all separable measurements on a  $d$ -level system. In the case of a two-level system, the GM bound was shown to be tight [107], in agreement with the earlier analysis of Hayashi [129]. In general, however, little is known whether the bound is attainable or not, which is the main motivation behind the present study.

In this chapter, we investigate the optimal estimation strategies and the optimal efficiency with adaptive measurements. Since we are concerned with the large-sample scenario, the optimal estimation strategies usually can be realized with two-step adaptive schemes [28, 107, 202]. Therefore, our main task is to devise the local optimal measurements and determine the corresponding efficiency.

We first give an alternative derivation of the GM inequality, which is much simpler than the original one. Explicit formulas of the GM bounds for the MSH and the MSB are also calculated, followed by a detailed discussion about their general properties. We then introduce a new optimization paradigm for minimizing the WMSE based on a unitarily invariant distance, which reduces the optimization domain from the set of

POMs to the set of Fisher information matrices. In this way, the dimension of the parameter space decreases considerably, and the nonconvexity involved in traditional optimization procedures is avoided. Based on this approach, we prove that the GM bound for the MSB can be saturated approximately within a factor of two by constructing an explicit measurement scheme. We also show by numerics that the GM bounds for both the MSB and the MSH, especially for the latter, are nearly tight. Finally, we compare the tomographic efficiencies of adaptive strategies with that of nonadaptive ones and discuss the implications of our study<sup>1</sup>.

## 5.2 Quantum Fisher information and quantum CR bound

Suppose the state  $\rho(\theta)$  of a given quantum system is characterized by a set of parameters  $\theta_1, \theta_2, \dots, \theta_g$ . To determine the values of these parameters, we may perform generalized measurements and construct an estimator based on the outcome statistics. Once a measurement with outcomes  $\Pi_\xi$  is chosen, it is well known in statistical inference that the MSE matrix of any unbiased estimator is lower bounded by the inverse of the Fisher information matrix and that the bound can be saturated asymptotically with the ML estimator [68, 93, 94, 224]. This lower bound induces a lower bound for the WMSE given any weight matrix. To achieve the minimal WMSE, we need to optimize the Fisher information over all possible measurements, which is generally very difficult. A major achievement in quantum estimation theory is the introduction of quantum analogs of the Fisher information matrix and the CR bound, which set a lower bound for the WMSE of any unbiased estimator [139, 141, 147].

### 5.2.1 One-parameter setting

Let  $\rho'(\theta) = d\rho(\theta)/d\theta$ , a Hermitian operator  $L(\theta)$  satisfying the equation

$$\rho'(\theta) = \frac{1}{2}[\rho(\theta)L(\theta) + L(\theta)\rho(\theta)] \quad (5.1)$$

---

<sup>1</sup>We are grateful to Masahito Hayashi for stimulating discussions on quantum estimation theory.

## 5.2. Quantum Fisher information and quantum CR bound

---

is called the SLD of  $\rho(\theta)$  with respect to  $\theta$  [141, 147]. By definition, the SLD satisfies  $\text{tr}\{\rho(\theta)L(\theta)\} = 0$  and

$$\text{tr}\{\rho'(\theta)A\} = \Re \text{tr}\{\rho(\theta)L(\theta)A\} = \Re \text{tr}\{\rho(\theta)AL(\theta)\} \quad (5.2)$$

for any Hermitian operator  $A$ .

The *SLD (quantum) Fisher information* is defined as [141, 147]

$$J(\theta) = \text{tr}\{\rho(\theta)L(\theta)^2\}. \quad (5.3)$$

It is a quantum analog of and, meanwhile, a tight upper bound for the Fisher information  $I(\theta)$ , as first demonstrated by Helstrom [139, 141]. In conjunction with the classical CR bound [68, 224], the inverse SLD Fisher information sets a lower bound for the MSE of any unbiased estimator, which is known as the SLD quantum CR bound, or SLD bound in short [45, 141, 147]. The inequality  $I(\theta) \leq J(\theta)$  can be shown as follows,

$$\begin{aligned} I(\theta) &= \sum_{\xi} \frac{[\text{tr}(\rho'\Pi_{\xi})]^2}{\text{tr}(\rho\Pi_{\xi})} \leq \sum_{\xi} \frac{|\text{tr}(\rho\Pi_{\xi}L)|^2}{\text{tr}(\rho\Pi_{\xi})} = \sum_{\xi} \frac{|\text{tr}\{(\Pi_{\xi}^{1/2}\rho^{1/2})^{\dagger}\Pi_{\xi}^{1/2}L\rho^{1/2}\}|^2}{\text{tr}(\rho\Pi_{\xi})} \\ &\leq \sum_{\xi} \text{tr}\{\rho L\Pi_{\xi}L\} = \text{tr}(\rho L^2) = J(\theta), \end{aligned} \quad (5.4)$$

where the first inequality follows from Eq. (5.2), and the second one from the Cauchy–Schwarz inequality. The first inequality is saturated if each  $\Pi_{\xi}$  commutes with  $L$ , and the second one if each  $\Pi_{\xi}^{1/2}$  is proportional to  $\Pi_{\xi}^{1/2}L(\theta)$ . The two inequalities are saturated simultaneously by measuring the observable  $L(\theta)$ . Therefore, the SLD bound in the one-parameter setting can be saturated locally by optimal individual measurements in conjunction with MLE [45, 141, 147].

The optimal measurements corresponding to different parameter values are generally incompatible since the corresponding SLDs are not commutative. As a consequence, it is generally impossible to devise a measurement that is optimal for all parameter values. Nevertheless, this goal can be achieved in the large- $N$  limit with a

simple two-step adaptive strategy [28, 132, 136, 202]. The basic idea can be sketched as follows. In the first step, we can perform a generic IC measurement on  $c\sqrt{N}$  copies of the true state  $\rho(\theta)$  for some constant  $c$  and compute the ML estimator according to the measurement statistics. In the second step, we perform the optimal measurement with respect to the estimator on the remaining  $N - c\sqrt{N}$  copies and compute the ML estimator again. The final estimator thus obtained can approximately saturate the SLD bound at each point.

In addition to its application in quantum estimation theory, the quantum Fisher information also plays an important role in studying the geometry of quantum states [30, 45, 216, 219]. For example, Braunstein and Caves [45] showed that the SLD Fisher information allows defining a statistical distance in the state space that is equivalent to the Bures distance [see Eq. (A.9)],

$$D_{\text{B}}^2(\rho(\theta), \rho(\theta + d\theta)) = \frac{1}{4}J(\theta)d\theta^2. \quad (5.5)$$

This equation endows the infinitesimal Bures distance with a clear operational meaning. The basic idea of their approach had been applied to studying the geometry on the probability simplex [63, 94, 224, 268].

In the rest of this section, we present an alternative formulation of the SLD bound in terms of superoperators, whose merit will become more obvious in the multiparameter setting. Let  $A$  be an arbitrary Hermitian operator, define superoperator  $\mathcal{R}(\rho)$  by the equation [45, 216, 219]

$$\mathcal{R}(\rho)|A\rangle\rangle = \frac{1}{2}|A\rho + \rho A\rangle\rangle; \quad (5.6)$$

note that the definition is independent of the parametrization of  $\rho$ . Alternatively,  $\mathcal{R}(\rho)$  can be spelled out in terms of the operator basis specified in Eq. (2.20),

$$\mathcal{R}(\rho) = \frac{1}{2} \sum_{j,k=1}^d (|E_{jl}\rangle\rangle \rho_{jk} \langle\langle E_{kl}| + |E_{lk}\rangle\rangle \rho_{jk} \langle\langle E_{lj}|). \quad (5.7)$$



## 5.2. Quantum Fisher information and quantum CR bound

---

Define

$$\mathcal{J}(\rho) = \mathcal{R}^{-1}(\rho), \quad \bar{\mathcal{J}}(\rho) = \bar{\mathbf{I}}\mathcal{J}(\rho)\bar{\mathbf{I}}. \quad (5.8)$$

Then  $\bar{\mathcal{J}}(\rho)$  and  $\mathcal{J}(\rho)$  satisfy a similar relation as do  $\bar{\mathcal{F}}(\rho)$  and  $\mathcal{F}(\rho)$  [see Eq. (4.9)],

$$\bar{\mathcal{J}}^{-1}(\rho) = \mathcal{J}^{-1}(\rho) - |\rho\rangle\rangle\langle\langle\rho| = \mathcal{R}(\rho) - |\rho\rangle\rangle\langle\langle\rho|. \quad (5.9)$$

This equation implies the concavity of  $\bar{\mathcal{J}}^{-1}(\rho)$  in  $\rho$ , that is,

$$\bar{\mathcal{J}}^{-1}(x\rho_1 + (1-x)\rho_2) \geq x\bar{\mathcal{J}}^{-1}(\rho_1) + (1-x)\bar{\mathcal{J}}^{-1}(\rho_2), \quad 0 \leq x \leq 1. \quad (5.10)$$

Note that  $\mathcal{J}^{-1}(\rho)$  is linear in  $\rho$  and  $|\rho\rangle\rangle\langle\langle\rho|$  is convex.

In terms of the superoperators introduced above, the SLD and the SLD quantum Fisher information can be written as

$$|L\rangle\rangle = \mathcal{J}(\rho)|\rho'\rangle\rangle, \quad J = \langle\langle\rho'|\mathcal{J}(\rho)|\rho'\rangle\rangle = \langle\langle\rho'|\bar{\mathcal{J}}(\rho)|\rho'\rangle\rangle. \quad (5.11)$$

The SLD bound for the Fisher information [see Eq. (5.4)] now reads

$$\langle\langle\rho'|\bar{\mathcal{F}}(\rho)|\rho'\rangle\rangle \leq \langle\langle\rho'|\bar{\mathcal{J}}(\rho)|\rho'\rangle\rangle. \quad (5.12)$$

Since the inequality holds for arbitrary traceless Hermitian operator  $\rho'$ , it follows that  $\bar{\mathcal{F}}(\rho) \leq \bar{\mathcal{J}}(\rho)$  or, equivalently,  $\mathcal{F}(\rho) \leq \mathcal{J}(\rho)$ . On the other hand, either of the two inequalities implies the inequality in Eq. (5.4). Therefore, the SLD bound for the Fisher information has three equivalent formulations:

$$I(\theta) \leq J(\theta), \quad \bar{\mathcal{F}}(\rho) \leq \bar{\mathcal{J}}(\rho), \quad \mathcal{F}(\rho) \leq \mathcal{J}(\rho). \quad (5.13)$$

A judicious choice from these formulations can greatly simplify the discussion, as we shall see later.

### 5.2.2 Multiparameter setting

When the quantum state is characterized by a set of parameters  $\theta_1, \theta_2, \dots, \theta_g$ , the quantum Fisher information takes on a matrix form,

$$J_{jk} = J_{kj} = \frac{1}{2} \text{tr} \{ \rho (L_j L_k + L_k L_j) \} = \langle \langle \rho_{,j} | \mathcal{J} | \rho_{,k} \rangle \rangle, \quad (5.14)$$

where  $L_j$  is the SLD associated with the parameter  $\theta_j$  and  $\rho_{,j} := \partial \rho / \partial \theta_j$ . To simplify the notation, we have suppressed the explicit dependence on the parameters. As an immediate consequence of Eq. (5.13), we have  $I \leq J$  as in the one-parameter setting. However, there is a crucial difference: The upper bound generally cannot be saturated except when the  $L_j$ s commute with each other. Saturating the upper bound means that the equality  $\mathbf{u}^T I \mathbf{u} = \mathbf{u}^T J \mathbf{u}$  holds for all  $g$ -dimensional real vectors  $\mathbf{u}$ . To this end, we need to measure all observables of the form  $\sum_{j=1}^g L_j u_j$  simultaneously. As a consequence of the complementarity principle [41], however, it is impossible to measure two noncommutative sharp observables simultaneously [204].

Given a weight matrix  $W$ , the inequality  $I \leq J$  sets a lower bound for the scaled WMSE  $\text{tr}(WC)$  for any unbiased estimator,

$$\text{tr}(WC) \geq \text{tr}(WI^{-1}) \geq \text{tr}(WJ^{-1}). \quad (5.15)$$

The first inequality can be saturated asymptotically with the ML estimator, but the second one generally cannot be saturated unless the  $L_j$ s commute with each other. In terms of superoperators, the SLD bound reads

$$\text{Tr}(\mathcal{W}C) \geq \text{Tr}(\mathcal{W}\bar{\mathcal{F}}^{-1}) \geq \text{Tr}(\mathcal{W}\bar{\mathcal{J}}^{-1}). \quad (5.16)$$

This formulation is parametrization independent and is particularly convenient to work with when the figure of merit is parametrization independent, such as the MSH and the MSB. In addition, by a suitable choice of  $\mathcal{W}$ , we may assume without loss of generality that the number  $g$  of parameters is equal to the dimension  $d^2 - 1$  of the state space.

### 5.3. Gill–Massar trace and Gill–Massar bound

---

For example, Eq. (5.16) gives rise to the SLD bound for the scaled MSH when  $\mathcal{W} = \bar{\mathbf{I}}$ ,

$$\mathcal{E}_{\text{SH}}^{\text{SLD}}(\rho) = \text{Tr}(\bar{\mathcal{J}}^{-1}) = \text{Tr}\{\mathcal{R}(\rho)\} - \text{tr}(\rho^2) = d - \text{tr}(\rho^2). \quad (5.17)$$

According to Eq. (5.5), the bound for the scaled MSB results from setting  $\mathcal{W} = \bar{\mathcal{J}}/4$ ,

$$\mathcal{E}_{\text{SB}}^{\text{SLD}}(\rho) = \frac{1}{4} \text{Tr}(\bar{\mathcal{J}}\bar{\mathcal{J}}^{-1}) = \frac{d^2 - 1}{4}. \quad (5.18)$$

The SLD bound in the multiparameter setting can be very loose. When  $\rho = 1/d$ , for example, the bound for either figure of merit is  $d + 1$  times smaller than the value achievable with optimal individual measurements (see Sec. 4.2). This result should be anticipated much earlier if we notice that the bound builds on a linear matrix inequality that is applicable to arbitrary individual measurements. Such an inequality does not account for the information trade-off among noncommutative observables.

### 5.3 Gill–Massar trace and Gill–Massar bound

The complementarity principle states that quantum systems possess properties that are equally real but mutually exclusive [41, 247]. In the quintessential example of the double-slit experiment, the photons (or electrons) can exhibit either particle behavior or wave behavior, but the sharpening of the particle behavior is necessarily accompanied with the blurring of the wave behavior, and vice versa. From the information-theoretic perspective, this means that an increase in the path information necessarily comes with a decrease in the fringe visibility, which is precisely quantified by Englert’s duality inequality [90]. Such information trade-off is not limited to the double-slit experiment. It presents itself whenever we are trying to extract information about noncommutative observables, thereby imposing a fundamental limit on the efficiency of quantum state estimation in the multiparameter setting. In a seminal work, Gill and Massar [107] derived a simple inequality on the Fisher information matrix that succinctly summarizes such information trade-off. Based on this inequality, they derived another quantum CR bound, known as the GM bound, which is applicable to all separable measurements

on a  $d$ -level system and is often much tighter than those bounds known previously. However, their derivation was quite involved and might leave the impression that the result is merely a coincidence. Maybe, this is one of the reasons why the importance of their work has not been fully recognized.

In this section, we propose a concise derivation of Gill–Massar’s result. The GM bounds for the MSH and the MSB are then calculated explicitly followed by a detailed explanation about their properties. The discussion in this section paved the way for constructing optimal measurements to be presented in Sec. 5.4.

### 5.3.1 Reexamination of the Gill–Massar inequality

The *Gill–Massar trace* (GMT) [107] is defined as the trace of the product of the Fisher information matrix and the inverse quantum Fisher information matrix, that is,  $t(\theta) := \text{tr}\{J^{-1}(\theta)I(\theta)\}$ . Note that it is independent of the parametrization as long as the space spanned by the  $\rho_{,j}$ s remains the same. Consider a measurement on a single copy of the true state. In the one-parameter setting, the GMT is the ratio of the Fisher information to the maximal Fisher information over all possible measurements, so its maximum is 1. In general, we can ensure that  $J(\theta)$  be diagonal with suitable parametrization, then the trace is the sum of the ratios for respective parameters. If we could perform the optimal measurements for all parameters simultaneously, then the maximum of the GMT would equal the number of parameters  $g$ , and the absolute maximum would equal  $d^2 - 1$ , the dimension of the state space. Surprisingly, Gill and Massar [107] showed that the trace is bounded from above by  $d - 1$ .

**Theorem 5.1** (Gill–Massar) *The scaled GMT of any separable measurement on  $N$  identically prepared  $d$ -level systems is bounded from above by  $d - 1$ ; that is,*

$$\text{tr}\{J^{-1}(\theta)I(\theta)\} \leq d - 1. \tag{5.19}$$

*The upper bound is saturated for any rank-one separable measurement when the number of parameters to be estimated is equal to  $d^2 - 1$ .*

### 5.3. Gill–Massar trace and Gill–Massar bound

---

The inequality in Theorem 5.1, henceforth called the GM inequality, succinctly summarizes the information trade-off in quantum state estimation in the multiparameter setting, which implies the general impossibility of constructing a measurement that is optimal for all parameters.

The original proof of Theorem 5.1 was quite involved. Based on the observation in Sec. 5.2, we can now provide a much simpler proof. Since adding auxiliary parameters does not decrease the GMT, we can assume that  $g = d^2 - 1$  without loss of generality [107]. In addition, locally, the Fisher information matrix achievable with a separable measurement can be achieved by individual measurements [107, 134]. So it suffices to prove the theorem for measurements on a single copy of the true state. In terms of superoperators, the GM inequality amounts to

$$d - 1 \geq \text{Tr}\{\bar{\mathcal{J}}^{-1}(\rho)\bar{\mathcal{F}}(\rho)\} = \text{Tr}\{\bar{\mathcal{J}}^{-1}(\rho)\mathcal{F}(\rho)\} = \text{Tr}\{\mathcal{J}^{-1}(\rho)\mathcal{F}(\rho)\} - 1, \quad (5.20)$$

where we have applied Eqs. (4.6) and (5.9) in deriving the last equality. Now Theorem 5.1 is an immediate consequence of the following equation:

$$\text{Tr}\{\mathcal{J}^{-1}(\rho)\mathcal{F}(\rho)\} = \sum_{\xi} \frac{\langle\langle \Pi_{\xi} | \mathcal{J}^{-1}(\rho) | \Pi_{\xi} \rangle\rangle}{\langle\langle \rho | \Pi_{\xi} \rangle\rangle} = \sum_{\xi} \frac{\text{tr}(\rho \Pi_{\xi}^2)}{\text{tr}(\rho \Pi_{\xi})} \leq \sum_{\xi} \text{tr}(\Pi_{\xi}) = d. \quad (5.21)$$

The inequality is saturated if the measurement is rank one.

#### 5.3.2 Gill–Massar bound for the scaled WMSE

Theorem 5.1 imposes a fundamental limit on the efficiency of quantum state estimation with individual measurements. Let  $C(\theta)$  be the scaled MSE matrix of any locally unbiased estimator, then Theorem 5.1 and the classical CR bound imply that

$$\text{tr}\{J^{-1}(\theta)C^{-1}(\theta)\} \leq d - 1. \quad (5.22)$$

This inequality sets a lower bound for the scaled WMSE  $\text{tr}(WC)$  according to Appendix D.1 [107],

$$\mathcal{E}_W^{\text{GM}} = \frac{(\text{tr} \sqrt{J^{-1/2} W J^{-1/2}})^2}{d-1} = \frac{(\text{tr} \sqrt{W^{1/2} J^{-1} W^{1/2}})^2}{d-1}. \quad (5.23)$$

If the lower bound is saturated, the scaled MSE matrix and the Fisher information matrix are given by

$$C_W^{-1} = I_W = (d-1) J^{1/2} \frac{\sqrt{J^{-1/2} W J^{-1/2}}}{\text{tr} \sqrt{J^{-1/2} W J^{-1/2}}} J^{1/2}. \quad (5.24)$$

In terms of superoperators, we have

$$\begin{aligned} \mathcal{E}_W^{\text{GM}} &= \frac{1}{d-1} \left( \text{Tr} \sqrt{\bar{\mathcal{J}}^{-1/2} \mathcal{W} \bar{\mathcal{J}}^{-1/2}} \right)^2, \\ C_W^{-1} = \bar{\mathcal{F}}_W &= (d-1) \frac{\bar{\mathcal{J}}^{1/2} \sqrt{\bar{\mathcal{J}}^{-1/2} \mathcal{W} \bar{\mathcal{J}}^{-1/2}} \bar{\mathcal{J}}^{1/2}}{\text{Tr} \sqrt{\bar{\mathcal{J}}^{-1/2} \mathcal{W} \bar{\mathcal{J}}^{-1/2}}}. \end{aligned} \quad (5.25)$$

Since  $\bar{\mathcal{J}}$  is supported on the space of traceless Hermitian operators, these formulas do not change with  $\mathcal{W}$  as long as its restriction on this space does not. This freedom may be exploited to simplify calculations. When  $W$  and  $J$  commute, Eqs. (5.23) and (5.24) can be simplified considerably,

$$\mathcal{E}_W^{\text{GM}} = \frac{(\text{tr} \sqrt{W J^{-1}})^2}{d-1}, \quad C_W^{-1} = I_W = \frac{(d-1) \sqrt{W J}}{\text{tr} \sqrt{W J^{-1}}}, \quad (5.26)$$

and so can Eq. (5.25).

In the case of a qubit, Theorem 5.1 implies that  $I(\theta) \leq J(\theta)$  or, equivalently,  $\bar{\mathcal{F}}(\rho) \leq \bar{\mathcal{J}}(\rho)$ . Therefore, the GM bound for the WMSE is at least as strong as the SLD bound; in fact, it can always be saturated [107]. To see this, suppose that the eigenbasis of  $\bar{\mathcal{F}}_W$  is composed of the three operators  $\sigma_j/\sqrt{2}$  for  $j = 1, 2, 3$ , where  $\sigma_j := \mathbf{r}_j \cdot \boldsymbol{\sigma}$ , and  $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$  are orthonormal vectors. Let  $a_1, a_2, a_3$  be the corresponding eigenvalues and  $s_1, s_2, s_3$  the three components of the Bloch vector of the true state in this basis. Then the GM bound can be saturated by measuring each observable  $\sigma_j$  with probability  $a_j(1 - s_j^2)/2$ . Note that the desired measurement is composed of

### 5.3. Gill–Massar trace and Gill–Massar bound

---

a complete set of mutually unbiased measurements. The normalization condition is ensured by Theorem 5.1, given that

$$\bar{\mathcal{J}}^{-1}(\rho) \hat{=} \frac{1}{2} \begin{pmatrix} 1 - s_1^2 & 1 - s_1 s_2 & 1 - s_1 s_3 \\ -s_1 s_2 & 1 - s_2^2 & -s_2 s_3 \\ -s_1 s_3 & -s_2 s_3 & 1 - s_3^2 \end{pmatrix}. \quad (5.27)$$

Our claim follows from the equation

$$\begin{aligned} \bar{\mathcal{F}}(\rho) &= \sum_{j=1,2,3} \frac{a_j(1 - s_j^2)}{2} \bar{\mathbf{I}} \left( \frac{|1 + \sigma_j\rangle\langle 1 + \sigma_j|}{2(1 + s_j)} + \frac{|1 - \sigma_j\rangle\langle 1 - \sigma_j|}{2(1 - s_j)} \right) \bar{\mathbf{I}} \\ &= \sum_{j=1,2,3} \frac{a_j}{2} |\sigma_j\rangle\langle \sigma_j| = \bar{\mathcal{F}}_{\mathcal{W}}. \end{aligned} \quad (5.28)$$

In general, little is known whether the GM bound can be saturated or not. In addition, neither the GM bound nor the SLD bound implies the other; their relative strengths are determined by the weight matrix. For many figures of merit commonly adopted in practice, such as the MSH and the MSB, it turns out that the GM bound is usually much tighter, as we shall see shortly. Incidentally, every rank-one measurement minimizes the WMSE for certain weight matrix, a simple example being  $\mathcal{W} = \bar{\mathcal{F}}\bar{\mathcal{J}}^{-1}\bar{\mathcal{F}}$ . In other words, every rank-one measurement is optimal for some purpose.

#### 5.3.3 Gill–Massar bounds for the mean square Bures distance and the mean square HS distance

The GM bound for the scaled MSB derives from Eq. (5.25) with  $\mathcal{W}(\rho) = \mathcal{J}(\rho)/4$  or  $\mathcal{W}(\rho) = \bar{\mathcal{J}}(\rho)/4$ ,

$$\mathcal{E}_{\text{SB}}^{\text{GM}}(\rho) = \frac{1}{4(d-1)} \left( \text{Tr} \sqrt{\bar{\mathcal{J}}(\rho)\bar{\mathcal{J}}^{-1}(\rho)} \right)^2 = \frac{1}{4}(d+1)^2(d-1). \quad (5.29)$$

Interestingly, the bound is independent of the true state. It is saturated if and only if there exists a measurement such that the Fisher information matrix is equal to

$$\bar{\mathcal{F}}_{\text{SB}}(\rho) = \frac{\bar{\mathcal{J}}(\rho)}{d+1}. \quad (5.30)$$

For a rank-one measurement, this condition amounts to the requirement that  $\bar{\mathcal{F}}(\rho)$  be proportional to  $\bar{\mathcal{J}}(\rho)$ . Literally, this means that all parameters are estimated equally well (or equally badly) compared with the optimal performance when each parameter is estimated separately. When  $\rho$  is the completely mixed state, the bound can be saturated according to Sec. 3.2.2; the general situation will be discussed in Sec. 5.4.

The GM bound for the scaled MSH follows from Eq. (5.25) with  $\mathcal{W} = \mathbf{I}$  or  $\bar{\mathbf{I}}$ ,

$$\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho) = \frac{1}{d-1} [\text{Tr}\{\bar{\mathcal{J}}^{-1/2}(\rho)\}]^2 = \frac{1}{d-1} \left[ \sum_{j \neq k=1}^d \sqrt{\frac{\lambda_j + \lambda_k}{2}} + \text{tr}(\sqrt{\Lambda}) \right]^2, \quad (5.31)$$

where the  $\lambda_j$ s are the eigenvalues of  $\rho$ , and  $\Lambda$  is a  $d \times d$  matrix with  $\Lambda_{jk} = \lambda_j \delta_{jk} - \lambda_j \lambda_k$ . It is saturated if and only if there exists a measurement that yields the Fisher information matrix

$$\bar{\mathcal{F}}_{\text{SH}} = \frac{(d-1)\sqrt{\bar{\mathcal{J}}(\rho)}}{\text{Tr}\{\bar{\mathcal{J}}^{-1/2}(\rho)\}}. \quad (5.32)$$

For a rank-one measurement, this condition amounts to the requirement that  $\bar{\mathcal{F}}(\rho)$  be proportional to  $\sqrt{\bar{\mathcal{J}}(\rho)}$ . When the dimension is large, Eq. (5.31) can be simplified by approximating  $\text{tr} \sqrt{\Lambda}$  with  $\sum_{j=1}^d \sqrt{\lambda_j} - \sqrt{\text{tr} \rho^2}$  or simply with  $\sum_{j=1}^d \sqrt{\lambda_j}$ ,

$$\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho) \approx \frac{1}{d-1} \left( \sum_{j,k=1}^d \sqrt{\frac{\lambda_j + \lambda_k}{2}} - \sqrt{\text{tr} \rho^2} \right)^2 \approx \frac{1}{d-1} \left( \sum_{j,k=1}^d \sqrt{\frac{\lambda_j + \lambda_k}{2}} \right)^2. \quad (5.33)$$

Following the concavity of  $\bar{\mathcal{J}}(\rho)$ , the GM bound  $\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho)$  is also concave in  $\rho$ . To see this, let  $\bar{\mathcal{J}}_1 = \bar{\mathcal{J}}(\rho_1)$  and  $\bar{\mathcal{J}}_2 = \bar{\mathcal{J}}(\rho_2)$ ; then we have

$$\begin{aligned} (d-1)\mathcal{E}_{\text{SH}}^{\text{GM}}(x\rho_1 + (1-x)\rho_2) &\geq [\text{Tr}\{x\bar{\mathcal{J}}_1^{-1} + (1-x)\bar{\mathcal{J}}_2^{-1}\}^{1/2}]^2 \\ &\geq x[\text{Tr}(\bar{\mathcal{J}}_1^{-1/2})]^2 + (1-x)[\text{Tr}(\bar{\mathcal{J}}_2^{-1/2})]^2 \\ &= (d-1)[x\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho_1) + (1-x)\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho_2)], \quad 0 \leq x \leq 1. \end{aligned} \quad (5.34)$$

As a consequence,  $\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho)$  is also Schur-concave in  $\rho$ . In particular, it reaches its maximum  $(d+1)^2(d-1)/d$  at the completely mixed state and its minimum  $2(d-1)$  at pure states. In both cases, the bounds can be saturated with the covariant measurement



### 5.3. Gill–Massar trace and Gill–Massar bound

---

according to Secs. 3.2.2 and 4.4. Compared with the SLD bound for the scaled MSH [see Eq. (5.17)], the GM bound is always tighter; for example, it is two times the SLD bound for pure states and  $d + 1$  times for the completely mixed state.

In certain scenarios it is more convenient to group identical eigenvalues of  $\rho$  together. Suppose  $\rho$  has  $n$  distinct eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$  with multiplicities  $d_1, d_2, \dots, d_n$ , respectively, where  $\sum_{j=1}^n d_j = d$ . Then Eq. (5.31) reduces to

$$\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho) = \frac{1}{d-1} \left[ \sum_{j,k=1}^n (d_j d_k - \delta_{jk}) \sqrt{\frac{\lambda_j + \lambda_k}{2}} + \text{tr}(\tilde{\Lambda}^{1/2}) \right]^2, \quad (5.35)$$

where  $\tilde{\Lambda}$  is an  $n \times n$  matrix with  $\tilde{\Lambda}_{jk} = \lambda_j \delta_{jk} - \sqrt{d_j d_k} \lambda_j \lambda_k$ . When  $n = 2$ , we have

$$\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho) = \frac{1}{d-1} [(d_1^2 - 1)\sqrt{\lambda_1} + (d_2^2 - 1)\sqrt{\lambda_2} + d_1 d_2 \sqrt{2(\lambda_1 + \lambda_2)} + \sqrt{d \lambda_1 \lambda_2}]^2. \quad (5.36)$$

Here the two distinct eigenvalues are determined by the purity  $\wp := \text{tr} \rho^2$  (assuming  $\lambda_1 \geq \lambda_2$ ),

$$\lambda_1 = \frac{1}{d} \left( 1 + \sqrt{\frac{d_2(d\wp - 1)}{d_1}} \right), \quad \lambda_2 = \frac{1}{d} \left( 1 - \sqrt{\frac{d_1(d\wp - 1)}{d_2}} \right). \quad (5.37)$$

Any state of this form is unitarily equivalent to a convex combination of the completely mixed state and a projector state, such as

$$\rho(s) = \frac{s P_{d_1}}{d_1} + \frac{(1-s)}{d}, \quad P_{d_1} = \sum_{j=1}^{d_1} |j\rangle\langle j|, \quad 0 \leq s \leq 1. \quad (5.38)$$

Our main interest in these states stems from the extremal property in the case  $d_1 = 1$  and  $d_2 = d - 1$ , as suggested by numerical calculation: For given purity,  $\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho)$  is maximized when all eigenvalues of  $\rho$  except the largest one are equal [it can be proved under the approximation Eq. (5.33)]. These states are, in a sense, the most difficult to estimate, assuming that the GM bound can be saturated. When  $d$  is very large and  $s$  is not very close to 0 or 1, calculation shows that  $\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho(s))$  decreases almost linearly with the increase of  $\sqrt{\wp}$  or  $s$  (see Fig. 5.2).

Numerical calculation also shows that, for a given purity  $\wp$ , the bound  $\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho)$  is minimized when all nonzero eigenvalues of  $\rho$  are equal except for the smallest one. Such a state assumes the following form up to unitary transformations,

$$\rho = \lambda_1 \sum_{j=1}^{r-1} |1\rangle\langle 1| + \lambda_2 |r\rangle\langle r|, \quad 2 \leq r \leq d, \quad \frac{1}{r} \leq \wp \leq \frac{1}{r-1}, \quad (5.39)$$

$$\lambda_1 = \frac{1}{r} \left( 1 + \sqrt{\frac{r\wp - 1}{r-1}} \right), \quad \lambda_2 = \frac{1}{r} \left( 1 - \sqrt{(r-1)(r\wp - 1)} \right).$$

Its rank is approximately inversely proportional to its purity. The GM bound reads

$$\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho) = \frac{1}{d-1} \left\{ [r(r-2) + \sqrt{2}(r-1)(d-r)]\sqrt{\lambda_1} + \sqrt{2}(r-1)\sqrt{\lambda_1 + \lambda_2} + (d-r)\sqrt{2\lambda_2} + \sqrt{r\lambda_1\lambda_2} \right\}^2. \quad (5.40)$$

When  $d$  is large, the rank  $r$  may be replaced by the effective rank  $1/\wp$ , and we have

$$\mathcal{E}_{\text{SH}}^{\text{GM}}(\rho) = \frac{(\sqrt{2}d\wp + 1 - \sqrt{2} - \wp^2)^2}{(d-1)\wp^3}. \quad (5.41)$$

This bound is roughly proportional to  $\wp^{-3}$  for very low purity and to  $\wp^{-1}$  for intermediate purity. For given purity, it is generally much smaller than the bound associated with the state in Eq. (5.38). It turns out that the bounds for both type of states can be saturated approximately, as we shall see later. Therefore, the GM bounds are crucial to understanding the tomographic efficiencies of individual measurements.

## 5.4 Optimal quantum state estimation with adaptive measurements

The problem of determining the optimal or nearly optimal estimation strategy has been a central problem in quantum estimation theory since the seminal works of Helstrom [139, 141] and of Holevo [147]. Although the problem in the one-parameter setting was solved nearly half a century ago [139], the problem in the multiparameter setting has largely remained open up to now. The difficulty is deeply rooted in the complementar-

## 5.4. Optimal quantum state estimation with adaptive measurements

---

ity principle, which imposes a fundamental limit on the information trade-off among noncommutative observables. A concise description of such trade-off was proposed by Gill and Massar [107], as discussed in the previous section, still little is known about the optimal estimation strategy as well as its efficiency gap from the GM bound.

In this section, we propose a general recipe for constructing optimal measurements with respect to the WMSE based on any unitarily invariant distance, such as the HS distance or the Bures distance. In contrast with traditional approaches, our solution does not need to optimize over the set of POMs directly, which is generally neither reliable nor efficient because of the nonconvexity and high-dimensionality of the optimization problem. Instead, it builds on a convex optimization procedure over the set of Fisher information matrices, which can be implemented reliably and efficiently in many cases of practical interest. Meanwhile, our approach provides a simple framework for understanding the information trade-off among different parameters. Based on this approach, we show that the GM bounds for the MSH and the MSB can be saturated approximately, although not exactly in general. In addition, adaptive strategies can improve the tomographic efficiency significantly over all nonadaptive ones, especially when the dimension of the Hilbert space is large or the states of interest have high purities.

### 5.4.1 A general recipe

In this section, we consider the problem of minimizing the scaled WMSE  $\text{Tr}\{\mathcal{W}(\rho)\bar{\mathcal{F}}^{-1}(\rho)\}$  or, equivalently,  $\text{Tr}\{\mathcal{W}(\rho)\mathcal{F}^{-1}(\rho)\}$ , assuming that  $\mathcal{W}(\rho)$  is covariant. In light of the two-step adaptive scheme described in Sec. 5.2.1, it suffices to look for the measurement that is optimal locally.

#### 5.4.1.1 $U_\rho$ -invariant Fisher information matrices

Let  $\mathcal{H}_j$  denote the eigenspace of  $\rho$  corresponding to the eigenvalue  $\lambda_j$  and  $U(\mathcal{H}_j)$  the unitary group acting on  $\mathcal{H}_j$  with normalized Haar measure  $d\mu_j$ . The stabilizer  $U_\rho$  of  $\rho$  under the action of  $U(\mathcal{H})$  is the direct product of the  $U(\mathcal{H}_j)$ s and has normalized

Haar measure  $d\mu_\rho = \prod_{j=1}^n d\mu_j$ . Since  $\mathcal{W}(\rho)$  is invariant under the action of  $U_\rho$  by assumption, the Fisher information  $\mathcal{F}(\rho)$  that minimizes the WMSE can also be chosen to be invariant; it is necessarily invariant if  $\mathcal{W}(\rho)$  is positive definite, although the measurement itself need not be invariant.

Consider a POM with outcomes  $\Pi_\xi = a_\xi |\psi_\xi\rangle\langle\psi_\xi|$  (assuming  $a_\xi > 0$ ). Each ket  $\sqrt{a_\xi} |\psi_\xi\rangle$  can be decomposed according to the eigenspaces of  $\rho$ ,

$$\sqrt{a_\xi} |\psi_\xi\rangle = \sum_{j=1}^n \sqrt{a_{j\xi}} |\psi_{j\xi}\rangle, \quad |\psi_{j\xi}\rangle \in \mathcal{H}_j, \quad a_{j\xi} \geq 0. \quad (5.42)$$

The normalization condition  $\sum_\xi \Pi_\xi = 1$  implies that

$$\sum_\xi a_{j\xi} = d_j \quad \text{for all } j. \quad (5.43)$$

Denote by  $1_j$  the identity operator on  $\mathcal{H}_j$  and let  $p_\xi = \text{tr}(\rho \Pi_\xi) = \sum_{j=1}^n \lambda_j a_{j\xi}$ . The invariance of  $\mathcal{F}(\rho)$  under the action of  $U_\rho$  implies that

$$\mathcal{F}(\rho) = \sum_\xi \frac{1}{p_\xi} \int d\mu_\rho |U \Pi_\xi U^\dagger\rangle\langle U \Pi_\xi U^\dagger| = \sum_{j,k=1}^n F_{jk} (\mathbf{I}_{jk} + |1_j\rangle\langle 1_k|), \quad (5.44)$$

where

$$\mathbf{I}_{jk} = \sum_{r_j, s_k} |E_{r_j, s_k}\rangle\langle E_{r_j, s_k}|, \quad F_{jk} = F_{kj} = \sum_\xi \frac{a_{j\xi} a_{k\xi}}{(d_j d_k + d_j \delta_{jk}) p_\xi}. \quad (5.45)$$

Here  $E_{r_j, s_k} = |r_j\rangle\langle s_k|$  [see Eq. (2.20)], the  $|r_j\rangle$ s form an orthonormal basis of  $\mathcal{H}_j$ , and the  $|s_k\rangle$ s of  $\mathcal{H}_k$ .

The matrix  $F$  plays the role of the Fisher information matrix and is our focus in this section. The diagonal entries of  $F$  represent the information gain on each eigenspace of  $\rho$ , while the off-diagonal ones on the coherence among different eigenspaces. The normalization condition Eq. (5.43) imposes  $n$  constraints on these entries,

$$\sum_{j=1}^n \lambda_j d_j F_{jk} + \lambda_k F_{kk} = 1 \quad \text{for } k = 1, 2, \dots, n, \quad (5.46)$$

## 5.4. Optimal quantum state estimation with adaptive measurements

---

which imply, among others, the GM equality

$$\mathrm{Tr}(\mathcal{J}^{-1}\mathcal{F}) = \frac{1}{2} \sum_{j,k=1}^n (d_j d_k + \delta_{jk} d_j) (\lambda_j + \lambda_k) F_{jk} = d. \quad (5.47)$$

Compared with Eq. (5.47), Eq. (5.46) imposes a more stringent information trade-off among different parameters. As a consequence, only  $n(n-1)/2$  entries of  $F$  are independent, and  $U_\rho$ -invariant Fisher information matrices form a convex set of dimension  $n(n-1)/2$ .

### 5.4.1.2 Extremal Fisher information matrices

The extremal points of the set of  $U_\rho$ -invariant Fisher information matrices are particularly interesting both for foundational studies and for practical calculations. On the one hand, they generalize the extremal setting in the double-slit experiment in which either the maximal path information or the maximal fringe visibility is attained [90]. On the other hand, they are crucial to determining the optimal measurement strategy in quantum state estimation, as we shall see shortly. In this section we uncover a series of extremal points that have a clear operational meaning. We believe that they exhaust all the extremal points and that the set of  $U_\rho$ -invariant Fisher information matrices forms a convex polytope.

Let  $B$  be a subset of the set of numbers  $1, 2, \dots, n$  and  $\bar{B}$  its complement. Define  $\mathcal{H}_B = \bigoplus_{j \in B} \mathcal{H}_j$  and  $d_B = \mathrm{Dim}(\mathcal{H}_B) = \sum_{j \in B} d_j$ . Then Eq. (5.46) implies that

$$\frac{1}{2} \sum_{j,k \in B} (d_j d_k + \delta_{jk} d_j) (\lambda_j + \lambda_k) F_{jk} \leq d_B. \quad (5.48)$$

The upper bound is saturated if and only if

$$\lambda_j F_{jk} = 0 \quad \text{for all } j \in \bar{B}, k \in B. \quad (5.49)$$

Whenever a measurement yields the maximal information on the subspace  $\mathcal{H}_B$ , it provides no information on the coherence between  $\mathcal{H}_B$  and its orthogonal complement.

Such a measurement can be decomposed into a measurement on the subspace  $\mathcal{H}_B$  and another one on its orthogonal complement.

When  $B$  consists of a single element  $j$ , we have

$$\frac{d_j}{d_j + 1} \leq F_{jj} \leq \frac{1}{(d_j + 1)\lambda_j} \quad \text{for } j = 1, 2, \dots, n. \quad (5.50)$$

The upper bound follows from Eq. (5.48); the lower bound follows from the convexity of the function  $a_j^2 / (\sum_{k=1}^n \lambda_k a_k)$  and is saturated if and only if

$$\frac{a_j \xi}{d_j} = \sum_{k=1}^n \lambda_k a_{k\xi} \quad \text{for all } \xi. \quad (5.51)$$

Multiplying Eq. (5.50) by  $d_j(d_j + 1)\lambda_j$  and summing over  $j$  yields

$$\sum_{j=1}^n d_j^2 \lambda_j \leq \sum_{j=1}^n d_j(d_j + 1)\lambda_j F_{jj} \leq d, \quad (5.52)$$

which, together with Eq. (5.47), implies that

$$0 \leq \frac{1}{2} \sum_{j \neq k=1}^n d_j d_k (\lambda_j + \lambda_k) F_{jk} \leq d - \sum_{j=1}^n d_j^2 \lambda_j. \quad (5.53)$$

The upper bound in Eq. (5.52) or the lower bound in Eq. (5.53) is saturated if and only if the measurement can decompose into independent measurements on the respective eigenspaces  $\mathcal{H}_j$ . Such a measurement provides the maximal information on each eigenspace but no information on the coherence among different eigenspaces. By contrast, the lower bound in Eq. (5.52) or the upper bound in Eq. (5.53) is saturated if and only if

$$\frac{a_j \xi}{d_j} = \frac{a_k \xi}{d_k} \quad \text{for all } j, k, \xi. \quad (5.54)$$

Literally, this equation means that all outcomes of the measurement are unbiased with respect to the eigenspaces of  $\rho$ , or the eigenbasis of  $\rho$  when the eigenvalues are nondegenerate. Such a measurement yields the maximal information on the coherence among different eigenspaces but the least information on each eigenspace.

#### 5.4. Optimal quantum state estimation with adaptive measurements

---

Define

$$t_B = \frac{1}{2} \sum_{j \neq k \in B} d_j d_k (\lambda_j + \lambda_k) F_{jk}. \quad (5.55)$$

Then similar reasoning as above yields

$$0 \leq t_B \leq d_B - \frac{1}{\lambda_B} \sum_{j \in B} d_j^2 \lambda_j, \quad (5.56)$$

where  $\lambda_B = \sum_{j \in B} d_j \lambda_j$ . The upper bound is saturated if and only if the measurement can decompose into a measurement on  $\mathcal{H}_B$  and a measurement on  $\mathcal{H}_{\overline{B}}$  and, in addition,

$$\frac{a_{j\xi}}{d_j} = \frac{a_{k\xi}}{d_k} \quad \text{for all } j, k \in B \quad \text{and for all } \xi. \quad (5.57)$$

Such a measurement is called  $B$ -unbiased.

Let  $\mathcal{P} = \{B_1, B_2, \dots, B_k\}$  be a partition of the set of numbers  $1, 2, \dots, n$  and define

$$t_{\mathcal{P}} = \sum_{B \in \mathcal{P}} t_B. \quad (5.58)$$

Then we have

$$0 \leq t_{\mathcal{P}} \leq d - \sum_{B \in \mathcal{P}} \left( \frac{1}{\lambda_B} \sum_{j \in B} d_j^2 \lambda_j \right). \quad (5.59)$$

The upper bound is saturated if and only if the measurement is  $B$ -unbiased for all  $B \in \mathcal{P}$ , in which case it is called  $\mathcal{P}$ -unbiased. Remarkably, the Fisher information matrix  $F$  does not depend on the specific measurement as long as it is  $\mathcal{P}$ -unbiased,

$$F_{jk} = \begin{cases} \frac{d_j}{(d_j+1)\lambda_B} & \text{if } j, k \in B \text{ for some } B \in \mathcal{P} \text{ and } k = j, \\ \frac{1}{\lambda_B} & \text{if } j, k \in B \text{ for some } B \in \mathcal{P} \text{ and } k \neq j, \\ 0 & \text{otherwise.} \end{cases} \quad (5.60)$$

The Fisher information matrix of a  $\mathcal{P}$ -unbiased measurement is also called  $\mathcal{P}$ -unbiased and is denoted by  $F_{\mathcal{P}}$  or  $\mathcal{F}_{\mathcal{P}}$  in superoperator form. According to the above analysis, it is an extremal point of the convex set of  $U_{\rho}$ -invariant Fisher information matrices.

We believe that the converse is also true.

**Conjecture 5.2** *Any  $U_\rho$ -invariant Fisher information matrix is extremal if and only if it is  $\mathcal{P}$ -unbiased for some partition  $\mathcal{P}$ .*

This conjecture holds when  $n = 2$  since, in that case, the set of Fisher information matrices forms a line segment and has only two extremal points.

We emphasize that extremal Fisher information matrices do not necessarily correspond to extremal POMs and vice versa. For example, in the case of a qubit, assuming that  $\rho$  has a nondegenerate spectrum, the Fisher information matrix of any POM composed of the outcomes  $(1 + \mathbf{r}_\xi \cdot \boldsymbol{\sigma})/k$  is  $\{\{1, 2\}\}$ -unbiased as long as the  $\mathbf{r}_\xi$ s constitute a regular  $k$ -polygon on the equator of the Bloch sphere. However, the POM can be written as a convex combination of projective measurements whenever  $k$  is even. On the other hand, let

$$\mathbf{r}_1 = (0, 0, 1), \quad \mathbf{r}_2 = \frac{(\sqrt{3}, 0, -1)}{2}, \quad \mathbf{r}_3 = \frac{(-\sqrt{3}, 3, -2)}{4}, \quad \mathbf{r}_4 = -\frac{(\sqrt{3}, 3, 2)}{4}; \quad (5.61)$$

then the POM with the four outcomes

$$\frac{3(1 + \mathbf{r}_1 \cdot \boldsymbol{\sigma})}{9}, \quad \frac{2(1 + \mathbf{r}_2 \cdot \boldsymbol{\sigma})}{9}, \quad \frac{2(1 + \mathbf{r}_3 \cdot \boldsymbol{\sigma})}{9}, \quad \frac{2(1 + \mathbf{r}_4 \cdot \boldsymbol{\sigma})}{9} \quad (5.62)$$

is extremal, but the corresponding Fisher information matrix is not.

In the rest of this section, we show that any  $\mathcal{P}$ -unbiased measurement can be realized with only finite outcomes. To demonstrate this point, it suffices to consider the case in which  $\mathcal{P}$  is the trivial partition, the partition consisting of only one block  $B = \{1, 2, \dots, n\}$ . When  $n = 1$ , any rank-one POM constructed out of a weighted 2-design with a finite number of elements (see Appendix B) satisfies the requirement. Otherwise, suppose that the  $m_j$  states  $|\psi_{j\xi_j}\rangle$  form a 2-design on  $\mathcal{H}_j$ . Then a  $B$ -unbiased measurement with  $4^{n-1} \prod_{j=1}^n m_j$  outcomes can be constructed as follows,

$$\begin{aligned} \Pi_{\boldsymbol{\xi}, \mathbf{k}} &= \frac{d}{4^{n-1} \prod_{j=1}^n m_j} |\Psi_{\boldsymbol{\xi}, \mathbf{k}}\rangle \langle \Psi_{\boldsymbol{\xi}, \mathbf{k}}|, \\ |\Psi_{\boldsymbol{\xi}, \mathbf{k}}\rangle &= \frac{1}{\sqrt{d}} \left( |\psi_{1\xi_1}\rangle \sqrt{d_1} + \sum_{j=2}^n |\psi_{j\xi_j}\rangle \sqrt{d_j} i^{k_j} \right), \end{aligned} \quad (5.63)$$



## 5.4. Optimal quantum state estimation with adaptive measurements

---

where  $\boldsymbol{\xi} = (\xi_1, \xi_2, \dots, \xi_n)$ ,  $\mathbf{k} = (k_2, k_3, \dots, k_n)$ , and each  $k_j$  takes on four possible values 0, 1, 2, 3. A drawback of this construction is that the number of outcomes increases exponentially with the number of eigenspaces. It is desirable to devise an alternative with polynomial number of outcomes.

In the special case  $n = 2$  and  $d_1 = d_2 = 1$ , we can choose  $|\psi_1\rangle = |1\rangle$  and  $|\psi_2\rangle = |2\rangle$ , then the  $\{\{1, 2\}\}$ -unbiased measurement is composed of the four outcomes  $\Pi_k = (|\Psi_k\rangle\langle\Psi_k|)/2$  with  $|\Psi_k\rangle = (|1\rangle + i^k|2\rangle)/\sqrt{2}$  for  $k = 0, 1, 2, 3$ , which can be decomposed into the projective measurements associated with  $\sigma_x$  and  $\sigma_y$ , respectively. This construction is not minimal; the minimal candidate consists of three outcomes, which correspond to the vertices of an equilateral triangle on the equator of the Bloch sphere.

### 5.4.1.3 Minimizing the WMSE with convex optimization

If Conjecture 5.2 is true, then any  $U_\rho$ -invariant Fisher information matrix can be written as a convex combination of the  $\mathcal{F}_{\mathcal{P}}$ s,

$$\mathcal{F}(\{x_{\mathcal{P}}\}) = \sum_{\mathcal{P}} x_{\mathcal{P}} \mathcal{F}_{\mathcal{P}}, \quad (5.64)$$

where  $\{x_{\mathcal{P}}\}$  is a probability distribution on the set of partitions. In addition, according to Carathéodory Theorem [3], each Fisher information matrix can be decomposed into a convex combination of no more than  $n(n-1)/2 + 1$  terms. Since each  $\mathcal{P}$ -unbiased Fisher information matrix can be realized with a finite number of outcomes, so can any Fisher information matrix of the form in Eq. (5.64).

Given a weight matrix  $\mathcal{W}$ , then the scaled WMSE takes on the form

$$\begin{aligned} \mathcal{E}_{\mathcal{W}}(\{x_{\mathcal{P}}\}) &= \text{Tr}\{\mathcal{W}\bar{\mathcal{F}}^{-1}(\{x_{\mathcal{P}}\})\} = \text{Tr}\{\mathcal{W}\mathcal{F}^{-1}(\{x_{\mathcal{P}}\})\} - \langle\langle\rho|\mathcal{W}|\rho\rangle\rangle \\ &= \text{Tr}\left\{\mathcal{W}\left(\sum_{\mathcal{P}} x_{\mathcal{P}} \mathcal{F}_{\mathcal{P}}\right)^{-1}\right\} - \langle\langle\rho|\mathcal{W}|\rho\rangle\rangle. \end{aligned} \quad (5.65)$$

Since  $\mathcal{E}_{\mathcal{W}}(\{x_{\mathcal{P}}\})$  is convex in the  $x_{\mathcal{P}}$ s, its minimum can be determined reliably. If Conjecture 5.2 holds, then this minimum is also the minimum achievable by any separable measurement. It should be noted that Conjecture 5.2 is sufficient but not necessary to

guarantee this claim. As we shall see shortly, the minimum of  $\mathcal{E}_{\mathcal{W}}(\{x_{\mathcal{P}}\})$  is so close to the GM bound in many interesting scenarios that it is at least nearly optimal if not exactly.

When the number of eigenspaces of  $\rho$  is small, the minimum of  $\mathcal{E}_{\mathcal{W}}(\{x_{\mathcal{P}}\})$  can be computed efficiently. As the number of eigenspaces increases, however, the number of partitions increases exponentially, so exact minimization of  $\mathcal{E}_{\mathcal{W}}(\{x_{\mathcal{P}}\})$  gets inefficient. A simple recipe for addressing this problem is to group eigenspaces with similar eigenvalues together and to consider only those partitions with respect to coarse-grained eigenspaces. This strategy is particularly effective for most states of interest in current experiments, which are nearly pure or have low ranks.

### 5.4.2 Approximate saturation of the Gill–Massar bound for the MSB

In this section, based on the work of Embacher and Narnhofer [88], we show that the GM bound for the MSB can be saturated approximately within a factor of two by constructing an explicit measurement scheme. This scheme provides a dramatic improvement over any nonadaptive alternative, with which the average MSB always diverges in the pure-state limit (see Sec. 4.4). Remarkably, the optimal measurement can be realized with at most  $2d$  types of projective measurements.

Inspired by Ref. [88], we shall construct a measurement from a convex combination of two measurements: The first one is the projective measurement on the eigenbasis of  $\rho$ , and the second one consists of the following  $2d(d-1)$  outcomes:

$$\begin{aligned} \Pi_{jkl} &= \frac{1}{2(d-1)} |\psi_{jkl}\rangle \langle \psi_{jkl}|, & |\psi_{jkl}\rangle &= \frac{1}{\sqrt{2}} (|j\rangle + i^l |k\rangle), \\ & & & 1 \leq j < k \leq d, \quad l = 0, 1, 2, 3. \end{aligned} \tag{5.66}$$

#### 5.4. Optimal quantum state estimation with adaptive measurements

---

The Fisher information matrices for the two measurements are given by

$$\begin{aligned}\mathcal{F}_1 &= \sum_{j=1}^d \frac{|E_{jj}\rangle\rangle\langle\langle E_{jj}|}{\lambda_j}, \\ \mathcal{F}_2 &= \frac{1}{2(d-1)} \left( \sum_{k>j=1}^d \frac{2(|E_{jk}^+\rangle\rangle\langle\langle E_{jk}^+| + |E_{jk}^-\rangle\rangle\langle\langle E_{jk}^-|)}{\lambda_j + \lambda_k} \right. \\ &\quad \left. + \sum_{k>j=1}^d \frac{2(|E_{jj} + E_{kk}\rangle\rangle\langle\langle E_{jj} + E_{kk}|)}{\lambda_j + \lambda_k} \right),\end{aligned}\tag{5.67}$$

where  $E_{jj}$  and  $E_{jk}^\pm$  are defined in Eqs. (2.20) and (2.21).

If we perform the two measurements with probabilities  $p_1 = 1/(2d-1)$  and  $p_2 = 2(d-1)/(2d-1)$ , respectively, then the total Fisher information is given by

$$\begin{aligned}\mathcal{F} &= p_1\mathcal{F}_1 + p_2\mathcal{F}_2 = \frac{1}{2d-1} \left( \sum_{k>j=1}^d \frac{2(|E_{jk}^+\rangle\rangle\langle\langle E_{jk}^+| + |E_{jk}^-\rangle\rangle\langle\langle E_{jk}^-|)}{\lambda_j + \lambda_k} \right. \\ &\quad \left. + \sum_{j=1}^d \frac{|E_{jj}\rangle\rangle\langle\langle E_{jj}|}{\lambda_j} + \sum_{k>j=1}^d \frac{2(|E_{jj} + E_{kk}\rangle\rangle\langle\langle E_{jj} + E_{kk}|)}{\lambda_j + \lambda_k} \right) \geq \frac{1}{2d-1} \mathcal{J}.\end{aligned}\tag{5.68}$$

Note that the sum of the first two terms in the parentheses is equal to  $\mathcal{J}$ . Since  $\mathcal{F}$  coincides with  $\mathcal{J}/(2d-1)$  in the  $(d^2-d)$ -dimensional subspace spanned by the  $E_{jk}^+$ s and the  $E_{jk}^-$ s for  $j \neq k$  [and the same applies to  $\bar{\mathcal{F}}$  and  $\bar{\mathcal{J}}/(2d-1)$ ], the scaled MSB satisfies

$$\frac{1}{4}(2d-1)(d^2-d) \leq \mathcal{E}_{\text{SB}} \leq \frac{1}{4}(2d-1)(d^2-1),\tag{5.69}$$

which, together with Eq. (5.29), implies that

$$\frac{d(2d-1)}{(d+1)^2} \leq \frac{\mathcal{E}_{\text{SB}}}{\mathcal{E}_{\text{SB}}^{\text{GM}}} \leq \frac{2d-1}{d+1}.\tag{5.70}$$

The GM bound is saturated approximately within a factor of two as claimed. The above measurement scheme can thus improve the tomographic efficiency significantly over any nonadaptive measurement, for which the scaled MSB diverges in the pure-state limit. In the case  $d=2$ , the bound is actually saturated, so that the measurement constructed above is optimal, in agreement with the analysis in Sec. 5.3.2.

For comparison, the scaled MSH of the measurement satisfies

$$(2d - 1)(d - 1) \leq \mathcal{E}_{\text{SH}} \leq (2d - 1)(d - \text{tr } \rho^2). \quad (5.71)$$

In sharp contrast with the previous scenario, now the performance of the measurement is usually worse than that of the optimal linear tomography [244] (see Sec. 3.2.2). A similar phenomenon was also noticed in Ref. [88].

In the rest of this section, following an idea of Embacher and Narnhofer [88], we show that the same performance as the measurement introduced above can be achieved by performing  $2d - 1$  types of projective measurements when  $d$  is even and that similar performance can be achieved by performing  $2d$  types of projective measurements when  $d$  is odd. The analysis is closely related to the decomposition of the Fisher information matrix into  $\mathcal{P}$ -unbiased Fisher information matrices defined in Sec. 5.4.1.2.

Suppose the eigenvalues of  $\rho$  are nondegenerate; then  $\mathcal{F}_1$  is the unbiased Fisher information matrix corresponding to the complete partition. When  $d$  is even,  $\mathcal{F}_2$  is an equal-weight combination of all unbiased Fisher information matrices corresponding to partitions in which each block has two elements. Remarkably, it is possible to find a decomposition that contains only  $d - 1$  terms. Any such decomposition essentially amounts to a solution to the combinatoric problem: Find  $d - 1$  partitions of the set of numbers  $1, 2, \dots, d$  such that each block has two elements and that every pair of numbers  $j, k$  appears exactly once in a same block. Such a set of partitions is called *mutually exclusive*. It is known that there exists a set of mutually exclusive partitions whenever  $d$  is even [88]. When  $d = 4$ , for example, the three partitions  $\{\{1, 2\}, \{3, 4\}\}$ ,  $\{\{1, 3\}, \{2, 4\}\}$ ,  $\{\{1, 4\}, \{2, 3\}\}$  are mutually exclusive. Once  $d - 1$  mutually exclusive partitions  $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_{d-1}$  are found, we have

$$\mathcal{F}_2 = \frac{1}{d - 1} \sum_{j=1}^{d-1} \mathcal{F}_{\mathcal{P}_j}. \quad (5.72)$$

As a consequence,  $\mathcal{F}_2$  can be achieved by performing  $2d - 2$  projective measurements with equal probability, noting that each  $\mathcal{F}_{\mathcal{P}_j}$  can be achieved by performing two pro-

## 5.4. Optimal quantum state estimation with adaptive measurements

---

jective measurements with equal probability.

When  $d$  is odd, we cannot find a simple decomposition of  $\mathcal{F}_2$  or  $\mathcal{F}$  that is independent of the eigenvalues of  $\rho$ . Fortunately, the alternative  $\mathcal{F}' := (\mathcal{F}_1/d) + (d-1)\mathcal{F}_2/d$  also approximately saturates the GM bound for the MSB within a factor of two. It can be written as an equal-weight combination of all unbiased Fisher information matrices corresponding to partitions in which each block has two elements except for one block with only one element. Moreover, we can reduce the number of terms in the decomposition to  $d$ . To see this, let  $\mathcal{P}'_1, \mathcal{P}'_2, \dots, \mathcal{P}'_d$  be  $d$  mutually exclusive partitions of the set of numbers  $1, 2, \dots, d+1$ . Construct a partition  $\mathcal{P}_j$  of the set of numbers  $1, 2, \dots, d$  from  $\mathcal{P}'_j$  by deleting the element  $d+1$  from the block that contains it. Then we have

$$\mathcal{F}' = \frac{1}{d} \sum_{j=1}^d \mathcal{F}_{\mathcal{P}_j}. \quad (5.73)$$

Therefore,  $\mathcal{F}'$  can be achieved by performing  $2d$  projective measurements with equal probability.

### 5.4.3 Degenerate two-level systems

To illustrate the method described in Sec. 5.4.1, in this section we determine the optimal measurement scheme when the true state has two distinctive eigenvalues. As simple as it may appear, this example already exhibits many features not present in state estimation for the two-level system, which are instructive for understanding optimal state estimation in more complicated scenarios. For concreteness, we choose the MSH and the MSB as the main figures of merit, but our approach applies equally well to other figures of merit that are based on unitarily invariant distances. Our study shows that the GM bounds for the MSH and the MSB can be saturated approximately but not exactly in general.

When  $\rho$  has two distinct eigenvalues, the set of  $U_\rho$ -invariant Fisher information matrices has two extremal points, which correspond to the complete partition and the trivial partition, respectively (see Sec. 5.4.1). The corresponding Fisher information

matrices  $\mathcal{F}_1$  and  $\mathcal{F}_2$  are given by

$$\begin{aligned}\mathcal{F}_1 &= \frac{1}{\lambda_1(d_1 + 1)} (\mathbf{I}_1 + |1_1\rangle\rangle\langle\langle 1_1|) + \frac{1}{\lambda_2(d_2 + 1)} (\mathbf{I}_2 + |1_2\rangle\rangle\langle\langle 1_2|), \\ \mathcal{F}_2 &= \frac{d_1}{d_1 + 1} (\mathbf{I}_1 + |1_1\rangle\rangle\langle\langle 1_1|) + \frac{d_2}{d_2 + 1} (\mathbf{I}_2 + |1_2\rangle\rangle\langle\langle 1_2|) \\ &\quad + (\mathbf{I} - \mathbf{I}_1 - \mathbf{I}_2 + |1_2\rangle\rangle\langle\langle 1_1| + |1_1\rangle\rangle\langle\langle 1_2|).\end{aligned}\tag{5.74}$$

Any other  $U_\rho$ -invariant Fisher information matrix is their convex combination,

$$\begin{aligned}\mathcal{F}(x) &= (1 - x)\mathcal{F}_1 + x\mathcal{F}_2, \\ &= F_{11}(\mathbf{I}_1 + |1_1\rangle\rangle\langle\langle 1_1|) + F_{22}(\mathbf{I}_2 + |1_2\rangle\rangle\langle\langle 1_2|) \\ &\quad + F_{12}(\mathbf{I} - \mathbf{I}_1 - \mathbf{I}_2 + |1_2\rangle\rangle\langle\langle 1_1| + |1_1\rangle\rangle\langle\langle 1_2|),\end{aligned}\tag{5.75}$$

where

$$F_{11} = \frac{1 - \lambda_2 d_2 x}{\lambda_1(d_1 + 1)}, \quad F_{22} = \frac{1 - \lambda_1 d_1 x}{\lambda_2(d_2 + 1)}, \quad F_{12} = x.\tag{5.76}$$

According to Eqs. (5.50) and (5.53), we have

$$\frac{d_1}{d_1 + 1} \leq F_{11} \leq \frac{1}{\lambda_1(d_1 + 1)}, \quad \frac{d_2}{d_2 + 1} \leq F_{22} \leq \frac{1}{\lambda_2(d_2 + 1)}, \quad 0 \leq F_{12} \leq 1.\tag{5.77}$$

The Fisher information matrix  $\bar{\mathcal{F}}(x)$  has four distinct eigenvalues  $F_{11}, F_{22}, F_{12}$ , and  $(1 - x)/d\lambda_1\lambda_2$  with multiplicities  $d_1^2 - 1, d_2^2 - 1, 2d_1d_2$ , and 1; the first three eigenspaces correspond to the supports of  $\bar{\mathbf{I}}_1, \bar{\mathbf{I}}_2$ , and  $\mathbf{I} - \mathbf{I}_1 - \mathbf{I}_2$ , respectively, while the last one is spanned by  $\sqrt{d_2/d_1 d}|1_1\rangle\rangle - \sqrt{d_1/d_2 d}|1_2\rangle\rangle$ .

Given a cost matrix  $\mathcal{W}$ , to find the optimal estimation strategy, it suffices to minimize the cost function  $\text{Tr}\{\mathcal{W}\bar{\mathcal{F}}^{-1}(x)\}$  over the variable  $x$ , which is trivial numerically since the function is convex. The cost functions for the scaled MSH and the scaled MSB are obtained when  $\bar{\mathbf{I}}$  and  $\bar{\mathcal{J}}/4$  are chosen as cost matrices, respectively,

$$\begin{aligned}\mathcal{E}_{\text{SH}}(x) &= \frac{\lambda_1(d_1 + 1)^2(d_1 - 1)}{1 - \lambda_2 d_2 x} + \frac{\lambda_2(d_2 + 1)^2(d_2 - 1)}{1 - \lambda_1 d_1 x} + \frac{2d_1 d_2}{x} + \frac{d\lambda_1 \lambda_2}{1 - x}, \\ \mathcal{E}_{\text{SB}}(x) &= \frac{1}{4} \left( \frac{(d_1 + 1)^2(d_1 - 1)}{1 - \lambda_2 d_2 x} + \frac{(d_2 + 1)^2(d_2 - 1)}{1 - \lambda_1 d_1 x} + \frac{4d_1 d_2}{(\lambda_1 + \lambda_2)x} + \frac{1}{1 - x} \right).\end{aligned}\tag{5.78}$$

#### 5.4. Optimal quantum state estimation with adaptive measurements

---

For each figure of merit, the minimization of the cost function usually leads to an order-6 polynomial equation in  $x$ . When  $d_1 = 1$  (or  $d_2 = 1$ ), the polynomial has order 4 and can be solved analytically; however, the formula is not so informative. In the special case of a qubit, that is,  $d_1 = d_2 = 1$ , Eq. (5.78) reduces to

$$\mathcal{E}_{\text{SH}}(x) = \frac{2}{x} + \frac{2\lambda_1\lambda_2}{1-x}, \quad \mathcal{E}_{\text{SB}}(x) = \frac{1}{x} + \frac{1}{4(1-x)}. \quad (5.79)$$

The minimum  $2(1 + \sqrt{\lambda_1\lambda_2})^2$  of  $\mathcal{E}_{\text{SH}}(x)$  is attained when  $x = 1/(1 + \sqrt{\lambda_1\lambda_2})$ , and the minimum  $\frac{9}{4}$  of  $\mathcal{E}_{\text{SB}}(x)$  is attained when  $x = \frac{2}{3}$ . The GM bounds can be saturated for both figures of merit [107] (see Sec. 5.3.2).

In general, it turns out that the bounds cannot be saturated exactly. According to Sec. 5.3.3, when  $d_1, d_2 \geq 2$ , a necessary condition for saturating the GM bound for the scaled MSB is

$$\lambda_1 F_{11} = \frac{\lambda_1 + \lambda_2}{2} F_{12} = \lambda_2 F_{22} = \frac{1}{d+1}. \quad (5.80)$$

According to Eqs. (5.76) and (5.77), this equation cannot be satisfied except when  $\lambda_1 = \lambda_2$ , that is, when the true state is completely mixed. The same conclusion also holds when  $d_1 = 1$  and  $d_2 \geq 2$ , in which case the last two equalities in Eq. (5.80) are still applicable for saturating the GM bound. Therefore, it is generally impossible to estimate every parameter equally well as compared with the optimal performance in estimating each parameter independently, which reflects more subtle information trade-off beyond the qubit setting. To better understand this result, it is instructive to take a look at the implications of Eqs. (5.77) and (5.80),

$$\frac{\lambda_1 d_1}{d_1 + 1} \leq \frac{1}{d+1}, \quad \frac{\lambda_2 d_2}{d_2 + 1} \leq \frac{1}{d+1}, \quad \lambda_1 + \lambda_2 \geq \frac{2}{d+1}, \quad (5.81)$$

which in turn imply that

$$\frac{1}{d+1} \leq \lambda_j \leq \frac{d_j + 1}{d_j(d+1)} \quad \text{for } j = 1, 2. \quad (5.82)$$

There is only a narrow region of choice in which  $\lambda_1$  and  $\lambda_2$  can satisfy these constraints.

When  $\lambda_1 > (d_1 + 1)/d_1(d + 1)$ ,  $F_{11}$  is always larger than required for saturating the GM bound for the scaled MSB, which implies by complementarity that either  $F_{12}$  or  $F_{22}$  must be smaller than required.

By contrast, a necessary condition for saturating the GM bound for the scaled MSH [see Eq. (5.32)] is

$$\sqrt{\lambda_1}F_{11} = \sqrt{\frac{\lambda_1 + \lambda_2}{2}}F_{12} = \sqrt{\lambda_2}F_{22} = \sqrt{\frac{d-1}{\mathcal{E}_{\text{SH}}^{\text{GM}}}}. \quad (5.83)$$

This equation generally cannot be satisfied either according to Eqs. (5.76) and (5.77). Compared with the conditions for saturating the GM bound for the scaled MSB, a major difference is that the value of  $F_{11}$  ( $F_{22}$ ) required for saturating the bound for the scaled MSH is larger (smaller), assuming  $\lambda_1 > \lambda_2$ . As a consequence, it is easier to saturate the bound approximately. This intuition is confirmed by extensive numerical calculations (see Fig. 5.1) and is instructive to understanding the optimal measurement schemes with respect to the two figures of merit.

Although the GM bounds for the scaled MSH and the scaled MSB generally cannot be saturated exactly, numerical calculation shows that they can be saturated approximately, especially for the former. Figure 5.1 shows the minimal scaled MSH and the minimal scaled MSB when the true states have the form  $s|1\rangle\langle 1| + (1-s)/d$  with  $d = 2, 5, 10, 15, 20$ . The minimal MSH decreases monotonically with  $s$ , while its gap from the GM bound first increases and then decreases; the maximal gap is less than 2%. The value of  $x$  corresponding to the optimal measurement scheme first decreases and then increases, except when  $d = 2$ , in which case it increases monotonically. Although no simple formula is known for this optimal value, calculation shows that nearly optimal performance can be achieved with the simple choice

$$x = \frac{\sqrt{2}}{\sqrt{2\lambda_1 d_1 + (d_2 + 1)\sqrt{(\lambda_1 + \lambda_2)\lambda_2}}}, \quad (5.84)$$

which is the solution to the second equality in Eq. (5.83). In sharp contrast, the minimal scaled MSB and its gap from the GM bound increases monotonically with  $s$ ;



## 5.4. Optimal quantum state estimation with adaptive measurements

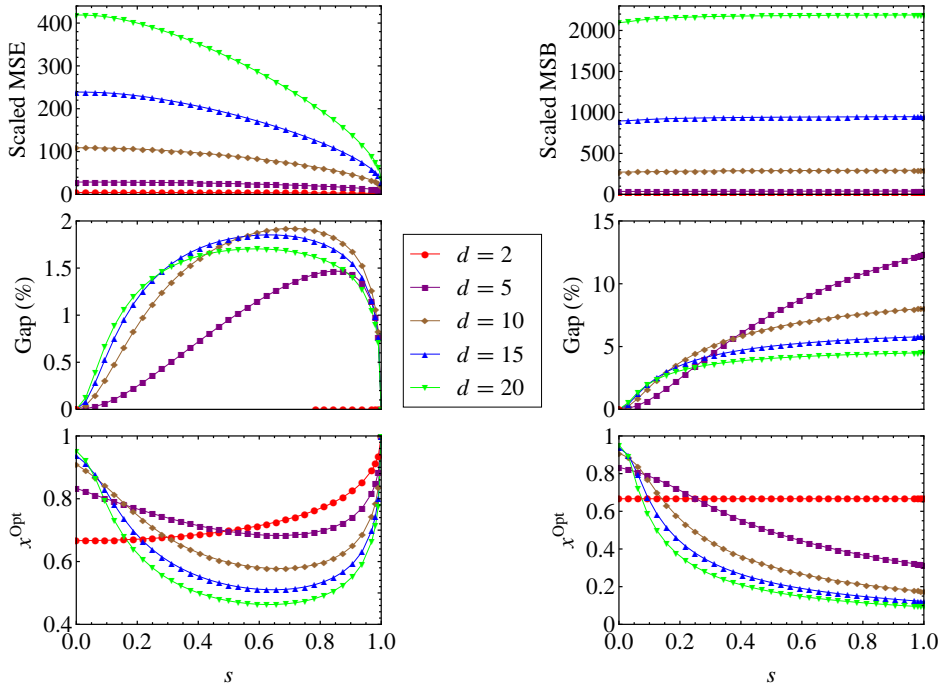


Figure 5.1: The scaled MSH (left) and the scaled MSB (right) in optimal state estimation with adaptive measurements for states of the form  $s|1\rangle\langle 1| + (1-s)/d$  with  $d = 2, 5, 10, 15, 20$ . Also shown are the relative gaps between the optimal values and the GM bounds as well as the values of  $x$  corresponding to the optimal measurements.

the maximal gap is less than 15%. The optimal value of  $x$  decreases monotonically with  $s$  except when  $d = 2$ , in which case it is a constant. In addition, nearly optimal performance can be achieved with the solution

$$x = \frac{2}{1 + (d+1)\lambda_1 + \lambda_2} \quad (5.85)$$

to the second equality in Eq. (5.80).

In the pure-state limit, the minimal scaled MSH and MSB, as well as the corresponding optimal values of  $x$  can be derived analytically; see Appendix D.2 for more details.

### 5.4.4 Comparison with nonadaptive schemes

In this section, we compare the performances of the optimal estimation strategies based on adaptive measurements and those based on nonadaptive measurements. Here we

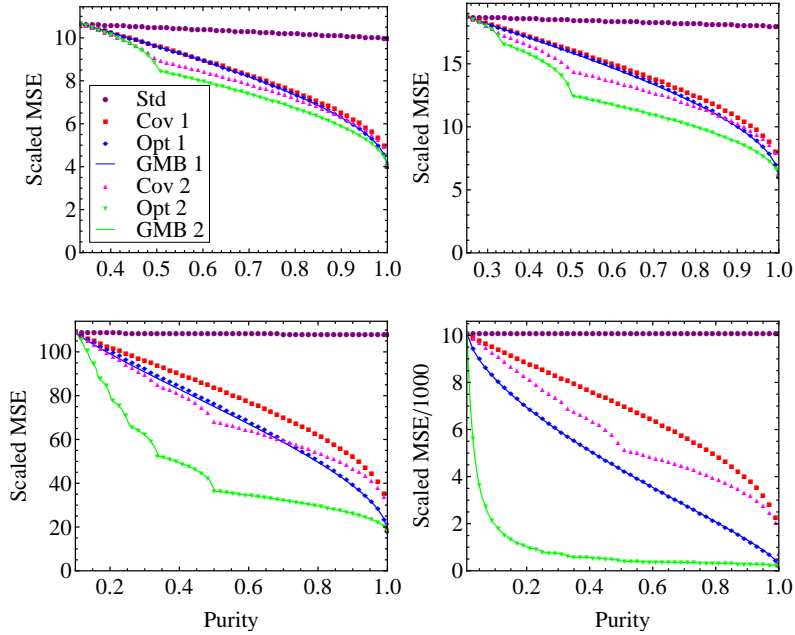


Figure 5.2: The minimal scaled MSHs in standard state estimation (Std), state estimation with covariant measurements (Cov), and state estimation with optimal adaptive measurements (Opt), respectively, for dimensions 3 (upper left), 4 (upper right), 10 (lower left), and 100 (lower right). The performances of the latter two strategies depend on not only the purity but also the spectrum: The curves Cov 1 and Opt 1 are applicable to the states in Eq. (5.38), while Cov 2 and Opt 2 are applicable to the states in Eq. (5.39). For comparison, the GM bounds (GMB) for the scaled MSHs are also plotted. The maximal gaps between the minimal scaled MSHs and the GM bounds are 0.7%, 1.2%, 1.9%, and 2.8% for dimensions 3, 4, 10, and 100, respectively.

assume the validity of Conjecture 5.2 when determining the performance of the optimal adaptive strategies, but our conclusion is independent of this assumption since the optimal performance under this assumption is quite close to the GM bound, as we shall see shortly.

When the scaled MSH is chosen as the figure of merit, our study in Sec. 5.3.3 suggests that, for given purity, the family of states in Eq. (5.38) are most difficult to estimate, whereas those in Eq. (5.39) are most easy to estimate. This observation is supported by numerical calculation based on the method described in Secs. 5.4.1 and 5.4.3. Therefore, the scaled MSHs associated with the two family of states can serve as

## 5.4. Optimal quantum state estimation with adaptive measurements

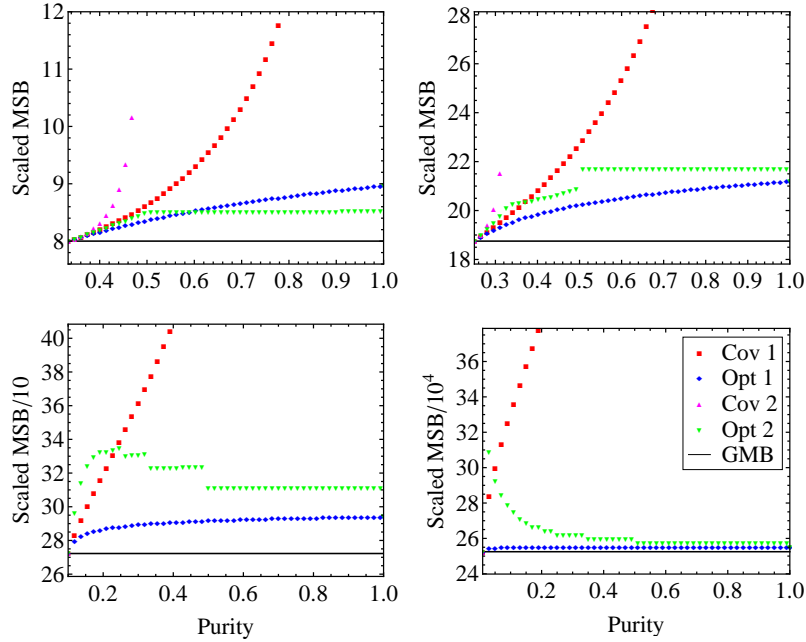


Figure 5.3: The minimal scaled MSBs for the covariant strategy and the optimal adaptive strategy with the same setting as in Fig. 5.2. When the true state becomes rank deficient, the scaled MSB for the covariant strategy diverges, while that for the optimal adaptive strategy is discontinuous (see Appendix D.3).

a benchmark for comparing resource requirements. Figure 5.2 shows the minimal scaled MSHs achievable with the standard strategy [244] (see Sec. 3.2.2), covariant strategy (see Sec. 4.4), and the optimal adaptive strategy, respectively. With the standard strategy, the scaled MSH has only a weak dependence on the purity of the true state and is independent of the spectrum for a given purity. By contrast, the scaled MSHs for the other two strategies, especially for the adaptive strategy, heavily depend on the purity and the spectrum. When the dimension is large, the adaptive strategy is much more efficient than the other two strategies. For states with low rank  $r \ll d$ , the minimal scaled MSH is approximately equal to  $2rd$ , which is roughly  $d/2r$  times smaller than the value in standard state estimation. Meanwhile, the minimal scaled MSH is very close to the GM bound. In other words, the tomographic efficiency of the optimal adaptive strategy is essentially characterized by the GM bound (see Sec. 5.3.3).

Figure 5.3 shows the performances of the covariant strategy and the optimal adap-

tive strategies with respect to the scaled MSB. We do not know the performance of the optimal standard strategies; suffice it to point out that it is no better than that of the covariant strategy. When the true state approaches the boundary of the state space, the scaled MSB for the covariant strategy diverges as pointed out in Sec. 4.4, and this problem gets more and more serious as the dimension of the Hilbert space increases. In contrast, the scaled MSB for the optimal adaptive strategies is finite, and it is generally quite close to the GM bound. Now, adaptive strategies are crucial to achieving high efficiency even if the dimension of the Hilbert space is small. However, special attention is necessary to ensure their robustness since the optimal measurement schemes and the minimal scaled MSB are strongly state dependent near the boundary of the state space, as explained in more detail in Appendix D.3.

## 5.5 Summary and open problems

We have studied the problem of optimal state estimation with adaptive measurements and proposed a general recipe for constructing optimal measurement schemes with respect to the WMSE based on a unitarily invariant distance. With this recipe, the optimization problem over POMs is reduced to that over Fisher information matrices, which greatly reduces the dimension of the parameter space and avoids the nuisance of nonconvexity in traditional optimization procedures. In addition, our approach provides a general framework for understanding the role of the complementarity principle in determining the tomographic efficiency. Furthermore, we showed that the GM bound for the MSB can be saturated approximately within a factor of two. Our numerical calculation indicates that the bound is nearly tight for a wide range of figures of merit, including the MSB and the MSH. In other words, the tomographic efficiencies of optimal adaptive strategies with respect to many figures of merit are essentially characterized by the GM bounds.

We further compared the tomographic efficiencies of adaptive schemes with that of nonadaptive ones and showed that the former can improve the tomographic efficiency significantly, especially when the dimension of the Hilbert space is large or the states of

## 5.5. Summary and open problems

---

interest have high purities. In many scenarios of experimental interest, even the scaling behavior of the efficiency with the dimension or the sample size can be improved by means of adaption. In that case, our study may help reduce resource consumption considerably.

There are a few open problems that we hope to address in the future.

1. Explore the connection between optimal state estimation and approximate simultaneous measurement of noncommuting observables.
2. Prove Conjecture 5.2 or characterize all extremal points of  $U_\rho$ -invariant Fisher information matrices.
3. Develop more efficient algorithms for minimizing the WMSE over the set of Fisher information matrices.
4. Construct an analytical proof that the GM bound for the scaled MSH can be saturated approximately.
5. Extend our approach to scenarios in which the figure of merit is not unitarily invariant, especially when the number of parameters of interest is much smaller than the dimension of the state space.
6. Investigate the optimal adaptive strategies in the case of limited sample size; compare two-step adaptive schemes with other alternatives.



# Quantum state estimation with collective measurements

---

## 6.1 Introduction

Collective measurements, which are often characterized by the use of quantum entanglement, are the most general measurements allowed by quantum mechanics. Their application to quantum state estimation is of paramount interest not only for reducing resource consumption in practice but also for understanding the distinctive features of quantum information processing as compared with classical information processing.

The problem of whether collective measurements can extract more information than individual measurements was first posed by Peres and Wootters [214] in the early 1990s. A positive answer was given by Massar and Popescu [191], who studied optimal estimation of qubit pure states based on the Bayesian approach. The same conclusion was later obtained for qubit mixed-state models [19, 20, 254, 259] and pure-state models in higher dimensions [49, 128]. In the large-sample limit, the CR approach is generally more suitable for investigating the efficiency gap. It turns out that separable measurements suffice to achieve the optimal performance for any pure-state model [193]. The same is true for several other models, such as estimation of a single parameter [45, 141] and of the spectrum of a qudit state [22]. In marked contrast, the efficiency advantage of collective measurements persists in the asymptotic limit in many mixed-state models, such as estimation of displaced thermal states [131] and of qubit mixed states [19, 20, 137, 252]. This observation reveals a radical departure of quantum information processing from its classical counterpart, in which the Fisher information is additive.

Recently, a major breakthrough in quantum estimation theory was made by Kahn and Guță et al. [125, 126, 160], who demonstrated local asymptotic normality for finite-dimensional quantum systems, which states that any quantum statistical model consisting of an ensemble of identically prepared systems can be approximated by a statistical model consisting of classical and quantum Gaussian variables in the asymptotic limit. This observation allows devising the optimal state-estimation strategies based on two-step adaptive schemes [28, 132, 136, 202]. Their work generalizes the earlier study of Hayashi [135] (see also Refs. [131] and [137]) on the applications of quantum central-limit theorem [111, 215] to quantum state estimation<sup>1</sup>.

Up to now, most studies on collective measurements presume the capability of joint measurements on arbitrary number of identically prepared quantum systems, which are hardly accessible in practice. An important problem left open is to determine the optimal estimation strategies and the corresponding tomographic efficiency in the case of limited access to collective measurements.

In this chapter, we study quantum state estimation in a more realistic scenario in which we are able to perform collective measurements but only on a limited number of systems. To circumvent the difficulty associated with traditional approaches, we introduce the concept of coherent measurements, which are composed of (generalized) coherent states [212, 276] as outcomes. As we shall see later, such measurements exhibit many nice features that make them an ideal starting point for investigating collective measurements. The GMT, which played a crucial role in studying individual measurements, will serve as a benchmark for comparing various measurement schemes.

We show that the GMT of any coherent measurement on  $\rho^{\otimes N}$  is a symmetric polynomial of the eigenvalues of  $\rho$ , which is independent of the specific coherent measurement. We believe that this polynomial is the maximum of the GMT over all possible measurements on  $\rho^{\otimes N}$  and prove our conjecture for several special yet important cases. These polynomials succinctly summarize the information trade-off among various parameters

---

<sup>1</sup>We are grateful to Masahito Hayashi for stimulating discussions on collective measurements and for several pertinent references.



## 6.2. Efficiency of asymptotic state estimation

---

in the case of collective measurements. They have profound implications for quantum estimation theory and, in particular, the open problem mentioned above, as we shall see later.

In the case of a two-level system, we propose a lower bound for the WMSE that is generally much tighter than any bound known previously. We then determine the set of Fisher information matrices of all coherent measurements and derive the maximal GMT over all measurements on  $\rho^{\otimes N}$ . As a byproduct, our study confirms a conjecture posed by Slater more than ten years ago [252]. Furthermore, we determine the tomographic efficiencies of the optimal coherent measurements in terms of the MSH and the MSB, and show that these measurements are almost optimal among all measurements. The distinctive features of collective measurements are also elaborated in comparison with individual measurements.

## 6.2 Efficiency of asymptotic state estimation

In this section, we briefly discuss the asymptotic tomographic efficiency based on the works of Hayashi [131, 135, 137], as well as Kahn and Guță [125, 126, 160]. In particular, we determine the maximal scaled GMT and the minimal scaled MSE and MSB, assuming that one can perform arbitrary collective measurements. Our study shows that the optimal measurements with respect to the three figures of merit are identical in the asymptotic limit, in marked contrast with state estimation using individual measurements. In addition, collective measurements can improve the scaling behavior of the tomographic efficiency with the dimension of the Hilbert space. The main tool in our study is another quantum CR bound based on the RLD [147, 274].

### 6.2.1 Quantum Cramér–Rao bound based on the right logarithmic derivative

Following the notation in Sec. 5.2, an operator  $\tilde{L}_j$  satisfying the equality

$$\rho_{,j} = \rho \tilde{L}_j \tag{6.1}$$

is called the RLD of  $\rho$  with respect to  $\theta_j$  [147, 274]. The *RLD Fisher information matrix*  $\tilde{J}$  is defined as

$$\tilde{J}_{jk} = \text{tr}(\rho \tilde{L}_k \tilde{L}_j^\dagger). \quad (6.2)$$

Like the SLD Fisher information matrix, it sets an upper bound for the Fisher information matrix  $I$ , and its inverse sets a lower bound for the MSE matrix  $C$  of any unbiased estimator, which is known as the RLD bound [147, 274].

To obtain an informative lower bound for the WMSE corresponding to a given weight matrix  $W$ , we need a lemma of Holevo [147] (Lemma 6.1 in Chapter VI).

**Lemma 6.1** (Holevo) *Let  $R$  be a complex Hermitian matrix; then*

$$\min_{Y \geq \pm R} \text{tr}(WY) = \text{tr} |\sqrt{W} R \sqrt{W}|, \quad (6.3)$$

and the minimum is achieved when  $Y = W^{-1/2} |\sqrt{W} R \sqrt{W}| W^{-1/2}$ .

When  $R$  is real or purely imaginary, the minimizing  $Y$  in Lemma 6.1 is real. When  $W$  and  $R$  commute, the minimum reduces to  $\text{tr}(W|R|)$  and is saturated at  $Y = |R|$ .

Since  $C$  is real, the bound  $C \geq \tilde{J}^{-1}$  implies that  $C - \Re(\tilde{J}^{-1}) \geq \pm i \Im(\tilde{J}^{-1})$ . According to Lemma 6.1,  $\text{tr}(WC)$  is lower bounded by

$$\mathcal{E}_W^{\text{RLD}} := \text{tr}\{W \Re(\tilde{J}^{-1})\} + \text{tr}(|\sqrt{W} \Im(\tilde{J}^{-1}) \sqrt{W}|), \quad (6.4)$$

and the bound is saturated if  $C$  is equal to

$$C_W^{\text{RLD}} = \Re(\tilde{J}^{-1}) + W^{-1/2} |\sqrt{W} \Im(\tilde{J}^{-1}) \sqrt{W}| W^{-1/2}. \quad (6.5)$$

When  $W$  and  $\tilde{J}$  commute, Eqs. (6.4) and (6.5) reduce to

$$\mathcal{E}_W^{\text{RLD}} = \text{tr}\{W[\Re(\tilde{J}^{-1}) + |\Im(\tilde{J}^{-1})|]\}, \quad C_W^{\text{RLD}} = \Re(\tilde{J}^{-1}) + |\Im(\tilde{J}^{-1})|. \quad (6.6)$$

Interestingly, the MSE matrix saturating the RLD bound is independent of the weight

## 6.2. Efficiency of asymptotic state estimation

---

matrix as long as it commutes with the RLD Fisher information matrix.

In the one-parameter setting, the RLD bound for the WMSE cannot be tighter than the SLD bound since the latter can be saturated. This is generally not the case in the multiparameter setting. To illustrate, we need to introduce the commutation superoperator  $\mathcal{D}$  first investigated by Holevo [147]. The superoperator is characterized by its action on an arbitrary linear operator  $A$ ,

$$\frac{1}{2}[\rho\mathcal{D}(A) + \mathcal{D}(A)\rho] = i(A\rho - \rho A); \quad (6.7)$$

it is linear and skew-Hermitian. A model is  $\mathcal{D}$ -invariant if the subspace spanned by the  $L_j$ s is invariant under the superoperator  $\mathcal{D}$ . In that case, the RLD bound is tighter than the SLD bound, as we shall see shortly. For a  $\mathcal{D}$ -invariant model, there is a simple relation between the RLD and the SLD Fisher information matrices [137, 147],

$$\tilde{J}^{-1} = J^{-1} + \frac{i}{2}J^{-1}DJ^{-1}, \quad (6.8)$$

where  $D$  is the real antisymmetric matrix defined by

$$D_{j,k} = \frac{1}{2}\text{tr}\{\rho[\mathcal{D}(L_k)L_j + L_j\mathcal{D}(L_k)]\} = i\text{tr}\{\rho(L_jL_k - L_kL_j)\}. \quad (6.9)$$

Now Eqs. (6.4) and (6.5) can be simplified by means of Eq. (6.8),

$$\begin{aligned} \mathcal{E}_W^{\text{RLD}} &= \text{tr}(WJ^{-1}) + \frac{1}{2}\text{tr}|W^{1/2}J^{-1}DJ^{-1}W^{1/2}|, \\ C_W^{\text{RLD}} &= J^{-1} + \frac{1}{2}W^{-1/2}|W^{1/2}J^{-1}DJ^{-1}W^{1/2}|W^{-1/2}. \end{aligned} \quad (6.10)$$

Compared with Eq. (5.15), the RLD bound for the scaled WMSE is tighter than the SLD bound as claimed.

Another important feature of a  $\mathcal{D}$ -invariant model is that the RLD bound is equal to the Holevo bound [137, 147].

When  $\rho$  is diagonal  $\rho = \sum_k \lambda_k |k\rangle\langle k|$ , the SLD and the SLD Fisher information

matrix can be computed as follows,

$$(L_j)_{kl} = (\rho, j)_{kl} \frac{2}{\lambda_k + \lambda_l}, \quad J_{jk} = \sum_{l,m=1}^d \frac{(\rho, j)_{lm}(\rho, k)_{ml} + (\rho, j)_{ml}(\rho, k)_{lm}}{\lambda_l + \lambda_m}. \quad (6.11)$$

As for the RLD and the RLD Fisher information matrix, we have

$$(\tilde{L}_j)_{kl} = \frac{(\rho, j)_{kl}}{\lambda_k}, \quad \tilde{J}_{jk} = \sum_{l,m=1}^d \frac{(\rho, j)_{ml}(\rho, k)_{lm}}{\lambda_l}. \quad (6.12)$$

### 6.2.2 Efficiency of the optimal state estimation in the asymptotic limit

In this section, we determine the maximal scaled GMT and the minimal scaled MSE and MSB in the asymptotic limit, assuming that one can perform arbitrary collective measurements. Our study is based on the fact that, for a  $\mathcal{D}$ -invariant model, the RLD bound is equal to the Holevo bound [137, 147], which can be saturated asymptotically according to Hayashi [131, 135, 137] (see also Refs. [125, 126, 160]).

To simplify the discussion, it is advisable to choose a suitable orthonormal basis of traceless Hermitian operators. Inspired by Ref. [107], we adopt a basis that comprises three types of elements,

$$\begin{aligned} \rho, jk+ &= E_{jk}^+, \quad \rho, jk- = E_{jk}^-, \quad 1 \leq j < k \leq d, \\ \rho, mm &= \sum_{k=1}^d a_{mk} E_{kk}, \quad m = 1, 2, \dots, d-1, \end{aligned} \quad (6.13)$$

where  $E_{kk}$  and  $E_{jk}^\pm$  are defined in Eqs. (2.20) and (2.21), and the real coefficients  $a_{mk}$  are chosen to ensure the orthonormality of the basis elements  $\rho, mm$ , whose specific values are not important. With this choice, the SLDs read

$$L_{jk\pm} = \frac{2}{\lambda_j + \lambda_k} E_{jk}^\pm, \quad L_{mm} = \sum_{k=1}^d \frac{a_{mk}}{\lambda_k} E_{kk}, \quad (6.14)$$

## 6.2. Efficiency of asymptotic state estimation

---

and they satisfy the equations

$$\mathcal{D}(L_{jk\pm}) = \pm \frac{2(\lambda_j - \lambda_k)}{\lambda_j + \lambda_k} L_{jk\mp}, \quad \mathcal{D}(L_{mm}) = 0. \quad (6.15)$$

Note that  $E_{jk}$  is an eigenvector of  $\mathcal{D}$  with eigenvalue  $-2i(\lambda_j - \lambda_k)/(\lambda_j + \lambda_k)$  [147]. In particular, the subspace spanned by the SLDs is invariant under the commutation superoperator; that is, our model is  $\mathcal{D}$ -invariant.

The SLD Fisher information matrix is diagonal with respect to the first two types of basis elements,

$$J_{jk\pm, jk\pm} = \frac{2}{\lambda_j + \lambda_k}. \quad (6.16)$$

The RLD Fisher information matrix is block diagonal with each block of size two, with nonzero entries given by

$$\tilde{J}_{jk\pm, jk\pm} = \frac{1}{2} \left( \frac{1}{\lambda_j} + \frac{1}{\lambda_k} \right), \quad \tilde{J}_{jk-, jk+} = -\tilde{J}_{jk+, jk-} = \frac{i}{2} \left( \frac{1}{\lambda_j} - \frac{1}{\lambda_k} \right). \quad (6.17)$$

Denote the submatrices of  $J$  and  $\tilde{J}$  with respect to the basis elements  $\rho_{,mm}$  by  $J_d$  and  $\tilde{J}_d$ , respectively (and define  $I_d$  and  $C_d$  in the same way); then we have  $J_d = \tilde{J}_d$  according to Eqs. (6.11) and (6.12). The matrix  $J_d$  satisfies the equality

$$\text{tr}(J_d^{-1}) = 1 - \text{tr}(\rho^2). \quad (6.18)$$

The proof follows from the discussion in Sec. 5.2.1,

$$\begin{aligned} \text{tr}(J_d^{-1}) &= \sum_{m=1}^{d-1} \langle\langle \rho_{,mm} | \bar{\mathcal{J}}^{-1}(\rho) | \rho_{,mm} \rangle\rangle = \sum_{m=1}^{d-1} \langle\langle \rho_{,mm} | [\mathcal{R}(\rho) - |\rho\rangle\rangle \langle\langle \rho | | \rho_{,mm} \rangle\rangle \\ &= \sum_{j=1}^d \langle\langle E_{jj} | [\mathcal{R}(\rho) - |\rho\rangle\rangle \langle\langle \rho | | E_{jj} \rangle\rangle = \text{tr}(\rho) - \text{tr}(\rho^2) = 1 - \text{tr}(\rho^2). \end{aligned} \quad (6.19)$$

According to the above analysis, it is straightforward to verify that  $\Re(\tilde{J}^{-1}) = J^{-1}$ , as required by Eq. (6.8) for a  $\mathcal{D}$ -invariant model.

Now we are ready to determine the minimal scaled MSE and MSB based on the inequality  $I \leq \tilde{J}$ . According to Eq. (6.6), the RLD bounds for both figures of merit are saturated at the same scaled MSE matrix  $C^{\text{RLD}} = \Re(\tilde{J}^{-1}) + |\Im(\tilde{J}^{-1})|$ , since the corresponding weight matrices 1 and  $J/4$  (see Sec. 5.2) commute with the RLD Fisher information matrix. Calculation shows that the scaled MSE matrix has the same block-diagonal structure as the SLD Fisher information matrix, with nonzero entries given by

$$C_d^{\text{RLD}} = J_d^{-1}, \quad C_{jk\pm, jk\pm}^{\text{RLD}} = \max(\lambda_j, \lambda_k). \quad (6.20)$$

The corresponding Fisher information matrix takes on the form

$$I_d^{\text{RLD}} = J_d, \quad I_{jk\pm, jk\pm}^{\text{RLD}} = \frac{1}{\max(\lambda_j, \lambda_k)}. \quad (6.21)$$

The RLD bound for the scaled MSE reads

$$\mathcal{E}^{\text{RLD}} = \text{tr}\{\Re(\tilde{J}^{-1}) + |\Im(\tilde{J}^{-1})|\} = d - \text{tr}(\rho^2) + \sum_{k>j=1}^d |\lambda_j - \lambda_k|, \quad (6.22)$$

where we have applied Eq. (6.18) in deriving the second equality. The minimum  $d-1/d$  of  $\mathcal{E}^{\text{RLD}}$  is attained when  $\rho$  is the completely mixed state, and the maximum  $2(d-1)$  when  $\rho$  is pure. By contrast, the RLD bound for the MSB reads

$$\mathcal{E}_{\text{SB}}^{\text{RLD}} = \frac{1}{4} \text{tr}\{J[\Re(\tilde{J}^{-1}) + |\Im(\tilde{J}^{-1})|]\} = \frac{d^2 - 1}{4} + \frac{1}{2} \sum_{k>j=1}^d \frac{|\lambda_j - \lambda_k|}{\lambda_j + \lambda_k}. \quad (6.23)$$

The minimum  $(d-1)(d+1)/4$  of  $\mathcal{E}_{\text{SB}}^{\text{RLD}}$  is attained at the completely mixed state, and the supremum  $(d-1)(2d+1)/4$  in the limit  $\lambda_j/\lambda_{j-1} \rightarrow 0$  for  $j = 2, 3, \dots, d$ . Except for the qubit, the bound is not well defined in the pure-state limit, and it can assume any value between  $(d-1)(d+3)/4$  and  $(d-1)(2d+1)/4$  depending on how the limit is taken. Similarly, the bound is not well defined when the rank of  $\rho$  is less than  $d-1$ .

Since our model is  $\mathcal{D}$ -invariant, the RLD bounds for the scaled MSE and MSB can be saturated asymptotically according to Refs. [131, 135, 137]. Figure 6.1 shows the contour plots of the asymptotic maximal scaled MSE and MSB in the eigenvalue

## 6.2. Efficiency of asymptotic state estimation

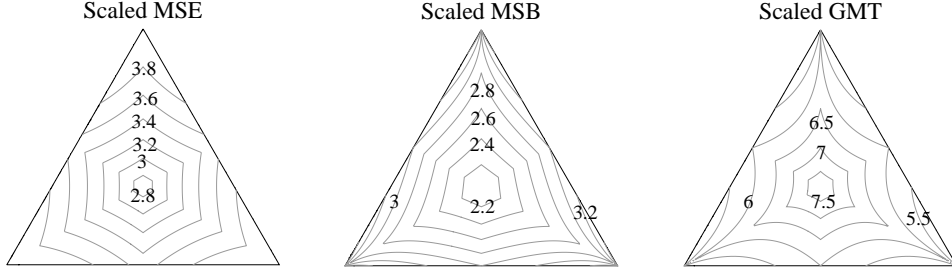


Figure 6.1: Contour plots of the asymptotic maximal scaled MSE, MSB, and minimal scaled GMT in the eigenvalue simplex for  $d = 3$ . The merging of contour lines at the extremal points in the middle and right plots indicates the singularity of the scaled MSB and GMT in the pure-state limit.

simplex for  $d = 3$ . The RLD bounds are generally much smaller than the GM bounds discussed in Sec. 5.3.3; for example, they are  $d + 1$  times smaller for the completely mixed state. Therefore, collective measurements can improve the scaling behaviors of tomographic efficiencies with the dimension of the Hilbert space.

The maximal scaled GMT  $t^{\text{RLD}}$  under the RLD bound  $I \leq \tilde{J}$  can be computed by means of Lemma 6.1,

$$\begin{aligned} t^{\text{RLD}} &= \text{tr}(J^{-1}\Re\tilde{J}) - \text{tr}|J^{-1/2}(\Im\tilde{J})J^{-1/2}| = \text{tr}\{J^{-1}(\Re\tilde{J} - |\Im\tilde{J}|)\} \\ &= d - 1 + \sum_{k>j=1}^d \frac{\lambda_j + \lambda_k}{\max(\lambda_j, \lambda_k)}. \end{aligned} \quad (6.24)$$

The maximum  $d^2 - 1$  is attained at the completely mixed state, and the infimum  $(d - 1)(d + 2)/2$  in the limit  $\lambda_j/\lambda_{j-1} \rightarrow 0$  for  $j = 2, 3, \dots, d$ . Except for the qubit, the bound is not well defined in the pure-state limit, and it may assume any value between  $(d - 1)(d + 2)/2$  and  $(d - 1)d$  depending on how the limit is taken. The Fisher information matrix saturating the upper bound is given by  $\Re\tilde{J} - |\Im\tilde{J}|$ , which is identical to the Fisher information matrix in Eq. (6.21). Therefore, minimizing the MSE or the MSB is equivalent to maximizing the GMT in the asymptotic limit. This observation further corroborates the significance of the GMT in the study of quantum state estimation.

### 6.3 Quantum state estimation with coherent measurements

In the previous section, we have studied the tomographic efficiency in the asymptotic limit, assuming one is capable of performing arbitrary collective measurements. There are two major problems left open: To what extent is this asymptotic analysis applicable in case of limited power in performing collective measurements? By how much can the efficiency be improved in that case compared with separable measurements. Although these problems are of paramount theoretical and practical interests, little is known about them because most traditional approaches are not effective in this scenario.

In this section, we study state estimation with collective measurements in a more realistic scenario, in which one is able to perform collective measurements but only on a limited number of systems each time. Nevertheless, the total sample available is still reasonably large so that the classical CR bound can be saturated. To circumvent the enormous difficulty associated with optimization, we use the GMT as a benchmark for comparing various measurement schemes and take a divide-and-conquer strategy. First, we exploit the underlying symmetry of the problem as characterized by the Schur–Weyl duality to reduce the problem on the whole Hilbert space  $\mathcal{H}^{\otimes N}$  to that on each irreducible subspace of the unitary group or the general linear group. Second, we introduce the concept of coherent measurements, measurements that are composed of (generalized) coherent states, and show that the GMT of any coherent measurement on  $\rho^{\otimes N}$  is a symmetric polynomial of the eigenvalues of  $\rho$ . Third, we prove that this polynomial is the maximum of the GMT over all possible measurements on  $\rho^{\otimes N}$  when either  $N = 2$  or  $d = 2$  and provide some evidence that this conclusion might hold in general. The implications of this polynomial for the two open problems mentioned above are also discussed in detail. Applications to state estimation on the two-level system is investigated in Sec. 6.4.



### 6.3. Quantum state estimation with coherent measurements

---

#### 6.3.1 Schur–Weyl duality and its implications

Let  $\mathcal{H}$  be a  $d$ -dimensional Hilbert space,  $\text{GL}(\mathcal{H})$  the general linear group, and  $S_N$  the symmetric group of  $N$  letters. There are two kinds of actions on the  $N$ -fold tensor space  $\mathcal{H}^{\otimes N}$ . Each operator  $X$  (in this chapter  $X$  denotes a generic operator instead of the cyclic-shift operator) in  $\text{GL}(\mathcal{H})$  acts on  $\mathcal{H}^{\otimes N}$  by simultaneous multiplication,

$$X^{\otimes N}(|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_N\rangle) = X|\psi_1\rangle \otimes X|\psi_2\rangle \otimes \cdots \otimes X|\psi_N\rangle; \quad (6.25)$$

each permutation  $\sigma$  in  $S_N$  acts by permuting the parties,

$$U_\sigma(|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_N\rangle) = |\psi_{\sigma^{-1}(1)}\rangle \otimes |\psi_{\sigma^{-1}(2)}\rangle \otimes \cdots \otimes |\psi_{\sigma^{-1}(N)}\rangle. \quad (6.26)$$

The two kinds of actions commute with each other. The *Schur–Weyl duality* states that the tensor space  $\mathcal{H}^{\otimes N}$  decomposes into a direct sum of tensor products under them [190, 222, 266],

$$\mathcal{H}^{\otimes N} = \bigoplus_{\mu} \mathcal{H}_{\mu} = \bigoplus_{\mu} (\mathcal{S}_{\mu} \otimes \mathcal{K}_{\mu}). \quad (6.27)$$

Here  $\mathcal{S}_{\mu}$  and  $\mathcal{K}_{\mu}$  are irreducible representations of the general linear group and the symmetric group, respectively, and the summation runs over all partitions  $\mu$  of  $N$  with no more than  $d$  parts. By convention, we take  $\mu$  as a vector of length  $d$  by adding zeros if necessary. The number of nonzero parts in  $\mu$  is called its height and is denoted by  $\text{ht}(\mu)$ . The dimension  $d_{\mu}$  of  $\mathcal{K}_{\mu}$  and the dimension  $D_{\mu}$  of  $\mathcal{S}_{\mu}$  are given by [222, 266],

$$d_{\mu} = \frac{N!}{h_{\mu}}, \quad D_{\mu} = \frac{y_{\mu}}{h_{\mu}}, \quad (6.28)$$

where

$$h_{\mu} = \frac{\prod_{j=1}^d (d + \mu_j - j)!}{\prod_{j < k} (\mu_j - \mu_k + k - j)}, \quad y_{\mu} = \prod_{j=1}^d \frac{(d + \mu_j - j)!}{(d - j)!}. \quad (6.29)$$

The representation space  $\mathcal{H}_{\mu}$  can be identified by its projector

$$H_{\mu} = \frac{d_{\mu}}{N!} \sum_{\sigma \in S_N} \chi_{\mu}(\sigma) U_{\sigma}, \quad (6.30)$$

where  $\chi_\mu(\sigma)$  is the character of  $\sigma$ . For example,  $H_{[2]} = (1+V)/2$  and  $H_{[1^2]} = (1-V)/2$ , where  $V$  is the *swap operator*. The space  $\mathcal{H}_\mu$  is composed of all symmetric states when  $\mu = [N]$  and of all antisymmetric states when  $\mu = [1^N]$ . In both cases,  $\mathcal{S}_\mu$  is identical with  $\mathcal{H}_\mu$ , and its projector  $S_\mu$  with  $H_\mu$ . In general, the multiplicity of the representation  $\mu$  of the general linear group is larger than 1, and the choice of  $\mathcal{S}_\mu$  or  $S_\mu$  is not unique; a convenient candidate will be introduced in Sec. 6.3.2.

The character of  $X$  in the representation  $\mu$  reads

$$s_\mu(X) = \text{tr}(S_\mu X^{\otimes N}) = \frac{1}{d_\mu} \text{tr}(H_\mu X^{\otimes N}); \quad (6.31)$$

note that it is independent of the choice of the subspace  $\mathcal{S}_\mu$ . The character  $s_\mu(X)$  is a symmetric polynomial of the eigenvalues  $x_1, x_2, \dots, x_d$  of  $X$ , which is known as the *Schur symmetric polynomial* [188, 222] and is denoted by  $s_\mu(x_1, \dots, x_d)$  or  $s_\mu(x)$  in short. According to Eq. (6.31),  $s_\mu(1, \dots, 1)$  is equal to the dimension of  $\mathcal{S}_\mu$ . In the following discussion,  $s_\mu(X)$  and  $s_\mu(x_1, \dots, x_d)$  are used interchangeably.

According to the Schur–Weyl duality,  $\rho^{\otimes N}$  is block diagonal with respect to the irreducible representations of the general linear group. As a consequence, the measurements that yield the maximal Fisher information can always be chosen such that all outcomes have the same block-diagonal structure and that equivalent representations yield the same Fisher information matrix. In that case, the total Fisher information matrix is a weighted sum,

$$I^N(\Pi, \theta) = \sum_{\mu} d_\mu I_\mu(\Pi_\mu, \theta), \quad (6.32)$$

where  $\Pi_\mu$  is a measurement on the subspace  $\mathcal{S}_\mu$ , and  $I_\mu$  the corresponding Fisher information matrix. The same is true for any figure of merit that is linear in the Fisher information matrix, such as the GMT. To optimize such a quantity, it suffices to optimize it on each irreducible component separately. We emphasize that the main merit of this approach is to simplify analysis and computation of the tomographic efficiency of collective measurements. It is not always necessary or practical to impose

### 6.3. Quantum state estimation with coherent measurements

---

such constraints on the measurements.

Let  $\{p_\sigma\}$  be a probability distribution on the symmetric group  $S_N$ . Given an optimal measurement  $\{\Pi_\xi\}$  on  $\rho^{\otimes N}$  with respect to a given figure of merit, then the measurement with outcomes

$$\Pi'_\xi = \sum_{\sigma \in S_N} p_\sigma U_\sigma \Pi_\xi U_\sigma^\dagger \quad (6.33)$$

is also optimal since it yields the same Fisher information matrix thanks to the permutation symmetry. Therefore, rank-one measurements are not crucial to achieving the optimal performance, unlike the scenario with individual measurements. This observation can help construct optimal measurements with fewer outcomes or simpler structure. However, to simplify the following discussion, we assume that all measurements are rank one in the rest of this chapter, except when stated otherwise.

#### 6.3.2 Highest-weight states and coherent states

Denote by  $\mathfrak{gl}(\mathcal{H})$  the Lie algebra of  $\text{GL}(\mathcal{H})$ . Then each operator  $O$  of  $\mathfrak{gl}(\mathcal{H})$  can be represented in  $\mathcal{H}^{\otimes N}$  as follows,

$$O^{(N)} = \sum_{k=1}^N 1^{\otimes(k-1)} \otimes O \otimes 1^{\otimes(N-k)}. \quad (6.34)$$

Let  $\mathfrak{u}^+$  be the subalgebra of  $\mathfrak{gl}(\mathcal{H})$  generated by  $|j\rangle\langle k|$  for  $1 \leq j < k \leq d$ . Then each operator in  $\mathfrak{u}^+$  has a strictly upper-triangular form in the standard basis.

A state in  $\mathcal{H}_\mu$  is a *highest-weight state* [222, 276] if it is annihilated by  $\mathfrak{u}^+$  or, equivalently, by  $|j\rangle\langle k|$  for  $1 \leq j < k \leq d$ . For example,  $|1\rangle^{\otimes N}$  is the unique highest-weight state in the symmetric subspace, and

$$|\Psi_{N-}\rangle := |1\rangle \wedge |2\rangle \wedge \cdots \wedge |N\rangle \quad (6.35)$$

is the unique highest-weight state in the antisymmetric subspace, where

$$|a_1\rangle \wedge |a_2\rangle \wedge \cdots \wedge |a_k\rangle = \frac{1}{\sqrt{k!}} \sum_{\sigma \in S_k} \text{sgn}(\sigma) |a_{\sigma(1)}, a_{\sigma(2)}, \dots, a_{\sigma(k)}\rangle \quad (6.36)$$

is known as a *Slater-determinant state*. For subspaces with mixed symmetry, highest-weight states are not unique. Since the actions of the unitary group and the symmetric group commute,  $|\Psi_\mu\rangle$  is a highest-weight state if and only if  $U_\sigma|\Psi_\mu\rangle$  is for any  $\sigma \in S_N$ . Any linear combination of highest-weight states in  $\mathcal{H}_\mu$  is also a highest-weight state by definition. The space spanned by all the highest-weight states in  $\mathcal{H}_\mu$  form an irreducible representation of the symmetric group.

Suppose  $|\Psi_\mu\rangle$  and  $|\Psi_\nu\rangle$  are highest-weight states of  $\mathcal{H}_\mu$  and  $\mathcal{H}_\nu$ , respectively. It is well known in representation theory that  $|\Psi_\mu\rangle \otimes |\Psi_\nu\rangle$  is a highest-weight state of  $\mathcal{H}_{\mu+\nu}$  [222], where  $\mu + \nu := [\mu_1 + \nu_1, \dots, \mu_d + \nu_d]$ . This observation points to a simple recipe for constructing highest-weight states in subspaces with mixed symmetry. In particular, it implies the existence of a highest-weight state in  $\mathcal{H}_\mu$  that is a tensor product of Slater-determinant states, a simple example being

$$\bigotimes_{j=1}^{\text{ht}(\tilde{\mu})} |\Psi_{\tilde{\mu}_j-}\rangle, \quad (6.37)$$

where  $\tilde{\mu}$  is the dual partition (also known as conjugate partition) of  $\mu$  [222].

A state in  $\mathcal{H}_\mu$  is a *coherent state* [5, 212, 276] if it can be generated from a highest-weight state by a unitary operator of the form  $U^{\otimes N}$ ; a coherent state is essentially a highest-weight state in a different local basis. For the symmetric subspace, each coherent state is a tensor power of a single-particle state and is thus separable. For the antisymmetric subspace, each coherent state is a Slater-determinant state, which is least entangled among all antisymmetric states [5, 220, 280]. In general, coherent states are characterized by maximal resemblance to classical states. They have been found useful in a variety of research areas, such as in the study of atomic systems and that of the classical and the thermodynamical limits of quantum mechanics [276].

The coherent states of  $\mathcal{H}_\mu$  form disjoint orbits under the action of  $U(\mathcal{H})$ , and the

### 6.3. Quantum state estimation with coherent measurements

---

span of all coherent states on one orbit is an irreducible subspace. There is a one-to-one correspondence among the trio: highest-weight states, orbits of coherent states, and irreducible subspaces in  $\mathcal{H}_\mu$ . The projector  $S_\mu$  onto the subspace generated from  $|\Psi_\mu\rangle$  can be constructed by twirling,

$$S_\mu = D_\mu \mathcal{P}_N(|\Psi_\mu\rangle\langle\Psi_\mu|) := D_\mu \int d\mu U^{\otimes N} |\Psi_\mu\rangle\langle\Psi_\mu| U^{\dagger\otimes N}, \quad (6.38)$$

where  $d\mu$  is the normalized Haar measure on the unitary group. The twirling can also be realized by a unitary  $N$ -design [122, 235], which may contain only a finite number of elements. The projector onto  $\mathcal{H}_\mu$  can be constructed in a similar way,

$$H_\mu = \frac{d_\mu D_\mu}{N!} \sum_{\sigma \in S_N} U_\sigma \mathcal{P}_N(|\Psi_\mu\rangle\langle\Psi_\mu|) U_\sigma^\dagger. \quad (6.39)$$

#### 6.3.3 Coherent measurements

A measurement on  $\mathcal{H}_\mu$  or  $\mathcal{S}_\mu$  is a *coherent measurement* if all outcomes are coherent states up to normalization. A measurement on  $\mathcal{H}^{\otimes N}$  is coherent if all outcomes are block diagonal with respect to the  $\mathcal{H}_\mu$ s, and its restriction on each  $\mathcal{H}_\mu$  is coherent. Coherent measurements are a very special class of collective measurements which, in a sense, are closest to separable measurements. Intuitively, one may not expect much relevance of such measurements to optimal state estimation, in view of the crucial role of entanglement in collective measurements. Quite surprisingly, they are actually optimal solutions to several special yet important state-estimation problems investigated in the literature [19, 20, 128, 254, 259], with or without being noticed. However, most previous studies chose Bayesian approaches and focused on two-level systems or pure-state systems. Here we shall study coherent measurements systematically from the point-wise perspective. To simplify the following discussion, we take  $\mathcal{S}_\mu$  to be the irreducible subspace generated from the highest-weight state defined by Eq. (6.37) in the rest of this chapter.

To illustrate the distinctive features of coherent measurements, let us first take the symmetric subspace as an example, assuming  $N \geq 2$ . Consider a coherent measurement

with outcomes  $\Pi_\xi = (|\psi_\xi\rangle\langle\psi_\xi|)^{\otimes N}$ , where the kets  $|\psi_\xi\rangle$  are not necessarily normalized.

Noticing that the probability  $p(\xi|\theta) = (\langle\psi_\xi|\rho|\psi_\xi\rangle)^N$  is factorized, we have

$$\begin{aligned} (I_{[N]})_{jk} &= \sum_{\xi} \frac{1}{p(\xi|\theta)} \frac{\partial p(\xi|\theta)}{\partial \theta_j} \frac{\partial p(\xi|\theta)}{\partial \theta_k} = N^2 \sum_{\xi} (\langle\psi_\xi|\rho|\psi_\xi\rangle)^{N-2} \langle\psi_\xi|\rho_{,j}|\psi_\xi\rangle \langle\psi_\xi|\rho_{,k}|\psi_\xi\rangle \\ &= N^2 \operatorname{tr}\{(\rho_{,j} \otimes \rho_{,k} \otimes \rho^{\otimes(N-2)}) S_{[N]}\} = \frac{N}{N-1} \frac{\partial^2 S_{[N]}(\rho)}{\partial \theta_j \partial \theta_k}. \end{aligned} \quad (6.40)$$

Interestingly, the Fisher information matrix is independent of the specific coherent measurement. In particular, it is invariant under the unitary transformation  $\Pi_\xi \rightarrow U^{\otimes N} \Pi_\xi U^{\dagger \otimes N}$  for any  $U \in U(\mathcal{H})$ , in sharp contrast with the scenario  $N = 1$ , in which this attribute pertains to only the covariant measurement.

It turns out that the invariance property of the Fisher information matrix is quite pervasive. According to Eq. (6.37), all coherent states in  $\mathcal{S}_\mu$  are unitarily equivalent to

$$\bigotimes_{r=1}^s |\Psi_{a_r-}\rangle^{\otimes b_r}, \quad (6.41)$$

where  $a_1 > a_2 \cdots > a_s$  are the sequence of distinct column lengths of the young diagram  $\mu$ , and  $b_1, b_2, \dots, b_s$  are the numbers of columns with the corresponding lengths. Therefore, each outcome  $\Pi_\xi$  of any coherent measurement on  $\mathcal{S}_\mu$  has a tensor-product form, and the probability of obtaining the outcome is factorized. The Fisher information matrix is independent of the specific coherent measurement whenever  $b_r \geq 2$  for  $r = 1, 2, \dots, s$ ; when  $a_1 = d$ , the same is true even if  $b_1 = 1$ .

When  $N$  is large, the conditions  $b_r \geq 2$  are satisfied for almost all partitions, so the total Fisher information matrix is almost independent of specific coherent measurements. In particular, all coherent measurements are equally efficient in the asymptotic limit.

### 6.3.4 Complementarity polynomials

Inspired by the previous analysis on the Fisher information matrix, in this section we show that the GMT of any coherent measurement on  $\rho^{\otimes N}$  is a symmetric polynomial

### 6.3. Quantum state estimation with coherent measurements

of the eigenvalues of  $\rho$ , which is independent of the specific coherent measurement. As we shall see shortly, this polynomial has profound implications for quantum estimation theory as well as for foundational issues, such as the complementarity principle.

To achieve our goal, we first prove a useful lemma concerning the SLD Fisher information matrix, following the notation in Sec. 5.2.

#### Lemma 6.2

$$\sum_{j,k=1}^{d^2-1} J^{-1}_{jk}(\rho_{,j} \otimes \rho_{,k}) = \frac{1}{2}V(\rho \otimes 1 + 1 \otimes \rho) - \rho^{\otimes 2}. \quad (6.42)$$

The proof follows from the observation that the equation in the lemma is equivalent to

$$\sum_{j,k=1}^{d^2-1} |\rho_{,j}\rangle\rangle J^{-1}_{jk} \langle\langle \rho_{,k}| = \frac{1}{2} \sum_{j,k=1}^d (|E_{jl}\rangle\rangle \rho_{jk} \langle\langle E_{kl}| + |E_{lk}\rangle\rangle \rho_{jk} \langle\langle E_{lj}|) - |\rho\rangle\rangle \langle\langle \rho|, \quad (6.43)$$

which is in turn equivalent to Eq. (5.9).

According to Lemma 6.2, the GMT of any measurement  $\Pi$  on  $\mathcal{S}_\mu$  takes on the form

$$t(\Pi, \rho) = \sum_{j,k=1}^{d^2-1} J^{-1}_{jk} \sum_{\xi} \frac{\text{tr}(\rho_{,j}^{\otimes N} \Pi_{\xi}) \text{tr}(\rho_{,k}^{\otimes N} \Pi_{\xi})}{\text{tr}(\rho^{\otimes N} \Pi_{\xi})} = \bar{t}(\Pi, \rho) - N^2 s_{\mu}(\rho), \quad (6.44)$$

where

$$\begin{aligned} \bar{t}(\Pi, \rho) &= \sum_{\xi} \frac{\text{tr}\{\Lambda_N(\rho) \Pi_{\xi}^{\otimes 2}\}}{\text{tr}(\rho^{\otimes N} \Pi_{\xi})}, \quad \Lambda_N(\rho) := \Lambda_{N,N}(\rho), \\ \Lambda_{M,N}(\rho) &:= \frac{1}{2} \sum_{\substack{1 \leq j \leq M \\ M < k \leq M+N}} V(j, k) (\rho^{\otimes(j-1)} \otimes 1 \otimes \rho^{\otimes(M+N-j)} \\ &\quad + \rho^{\otimes(k-1)} \otimes 1 \otimes \rho^{\otimes(M+N-k)}), \end{aligned} \quad (6.45)$$

and  $V(j, k)$  is the swap operator between party  $j$  and party  $k$ . Note that  $\Lambda_{M,N}(\rho)$  is a Hermitian operator. Now, it remains to show that  $\bar{t}(\Pi, \rho)$  is a symmetric polynomial of the eigenvalues of  $\rho$ . We shall demonstrate this point by constructing a Hermitian operator  $T_{\mu}(\rho)$  on  $\mathcal{H}^{\otimes N}$  such that the following equality holds for any coherent state  $|\Psi\rangle$  in  $\mathcal{S}_{\mu}$ :

$$\frac{\text{tr}\{\Lambda_N(\rho) (|\Psi\rangle\langle\Psi|)^{\otimes 2}\}}{\langle\Psi|\rho^{\otimes N}|\Psi\rangle} = \langle\Psi|T_{\mu}(\rho)|\Psi\rangle. \quad (6.46)$$

For the symmetric subspace, Eq. (6.46) is satisfied with

$$T_{[N]}(\rho) = N^2(1 \otimes \rho^{\otimes(N-1)}). \quad (6.47)$$

Recall that any coherent state in the symmetric subspace is a tensor power of a single-particle state. For the antisymmetric subspace, Eq. (6.46) is satisfied with

$$T_{[1N]}(\rho) = N(1 \otimes \rho^{\otimes(N-1)}), \quad (6.48)$$

which differs from  $T_{[N]}(\rho)$  by a factor of  $N$ . The proof of the equality

$$\frac{\text{tr}\{\Lambda_N(\rho)(|\Psi\rangle\langle\Psi|)^{\otimes 2}\}}{\langle\Psi|\rho^{\otimes N}|\Psi\rangle} = \langle\Psi|N(1 \otimes \rho^{\otimes(N-1)})|\Psi\rangle \quad (6.49)$$

for any Slater-determinant state  $|\Psi\rangle$  is relegated to Appendix E.1.

In general, the operator  $T_\mu(\rho)$  satisfying Eq. (6.46) can be constructed by induction. Let  $\mu$  and  $\nu$  be two partitions of  $M$  and  $N$ , respectively, that satisfy  $\tilde{\mu}_{\text{ht}(\tilde{\mu})} \geq \tilde{\nu}_1$ . Suppose  $T_\mu(\rho)$  and  $T_\nu(\rho)$  have been constructed such that Eq. (6.46) is satisfied for any coherent state in  $\mathcal{S}_\mu$  or in  $\mathcal{S}_\nu$ . Let

$$T_{\mu+\nu}(\rho) = T_\mu(\rho) \otimes \rho^N + \rho^{\otimes M} \otimes T_\nu(\rho) + 2\Lambda_{M,N}(\rho), \quad (6.50)$$

then Eq. (6.46) is satisfied for any coherent state in  $\mathcal{S}_{\mu+\nu}$ . To demonstrate this point, note that any coherent state  $|\Psi\rangle$  in  $\mathcal{S}_{\mu+\nu}$  can be written as a tensor product, namely,  $|\Psi\rangle = |\Psi_\mu\rangle \otimes |\Psi_\nu\rangle$ , where  $|\Psi_\mu\rangle$  and  $|\Psi_\nu\rangle$  are coherent states in  $\mathcal{S}_\mu$  and in  $\mathcal{S}_\nu$ , respectively. Accordingly, we have

$$\begin{aligned} \frac{\langle\Psi|^{\otimes 2}\Lambda_{M+N}(\rho)|\Psi\rangle^{\otimes 2}}{\langle\Psi|\rho^{\otimes M+N}|\Psi\rangle} &= \frac{\langle\Psi_\mu|^{\otimes 2}\Lambda_M(\rho)|\Psi_\mu\rangle^{\otimes 2}}{\langle\Psi_\mu|\rho^{\otimes M}|\Psi_\mu\rangle} \langle\Psi_\nu|\rho^{\otimes N}|\Psi_\nu\rangle \\ &+ \langle\Psi_\mu|\rho^{\otimes M}|\Psi_\mu\rangle \frac{\langle\Psi_\nu|^{\otimes 2}\Lambda_N(\rho)|\Psi_\nu\rangle^{\otimes 2}}{\langle\Psi_\nu|\rho^{\otimes N}|\Psi_\nu\rangle} + 2\langle\Psi|\Lambda_{M,N}(\rho)|\Psi\rangle \\ &= \langle\Psi|[T_\mu(\rho) \otimes \rho^{\otimes N} + \rho^{\otimes M} \otimes T_\nu(\rho) + 2\Lambda_{M,N}(\rho)]|\Psi\rangle. \end{aligned} \quad (6.51)$$



### 6.3. Quantum state estimation with coherent measurements

---

Now the operator  $T_\mu(\rho)$  can be constructed explicitly,

$$T_\mu(\rho) = \sum_{k=1}^{\text{ht}(\tilde{\mu})} (\tilde{\mu}_k \rho^{\otimes a_k - 1} \otimes 1 \otimes \rho^{\otimes (N - a_k)}) + \sum_{k > j=1}^{\text{ht}(\tilde{\mu})} \left[ \tilde{\mu}_j \tilde{\mu}_k V(a_j, a_k) \times (\rho^{\otimes (a_j - 1)} \otimes 1 \otimes \rho^{\otimes (N - a_j)} + \rho^{\otimes (a_k - 1)} \otimes 1 \otimes \rho^{\otimes (N - a_k)}) \right], \quad (6.52)$$

where  $a_k = 1 + \sum_{j=1}^{k-1} \tilde{\mu}_j$  and  $\tilde{\mu}$  is the dual partition of  $\mu$ . To get an operator with a simple expression, we have taken into account the permutation symmetry of coherent states in  $\mathcal{S}_\mu$ , and Eq. (6.50) is not guaranteed with the above choice.

Define

$$\bar{t}_\mu(\rho) = \text{tr}\{T_\mu(\rho)S_\mu\}, \quad t_\mu(\rho) = \bar{t}_\mu(\rho) - N^2 s_\mu(\rho); \quad (6.53)$$

then the GMT of any coherent measurement on  $\mathcal{S}_\mu$  is equal to  $t_\mu(\rho)$ , which is a symmetric polynomial of the eigenvalues of  $\rho$ . Although the operator  $T_\mu(\rho)$  is generally not unique, the polynomial  $t_\mu(\rho)$  is. For the symmetric and the antisymmetric subspaces, these polynomials are given by

$$\begin{aligned} t_{[N]}(\rho) &= N(d + N - 1)s_{[N-1]}(\rho) - N^2 s_{[N]}(\rho), \\ t_{[1^N]}(\rho) &= (d - N + 1)e_{N-1}(\rho) - N^2 e_N(\rho). \end{aligned} \quad (6.54)$$

where  $e_k(\rho) := s_{[1^k]}(\rho)$  is the  $k$ th elementary symmetric polynomial [188, 222]. Define

$$\bar{t}^N(\rho) = \sum_{\mu} d_{\mu} \bar{t}_{\mu}(\rho), \quad t^N(\rho) = \sum_{\mu} d_{\mu} t_{\mu}(\rho) = \bar{t}^N(\rho) - N^2 [\text{tr}(\rho)]^N. \quad (6.55)$$

Then the GMT of any coherent measurement on  $\rho^{\otimes N}$  is equal to  $t^N(\rho)$ . The polynomials  $t_\mu(\rho)$  and  $t^N(\rho)$  are called *complementarity polynomials* for reasons that will become clear shortly. The polynomials  $\bar{t}_\mu(\rho)$  and  $\bar{t}^N(\rho)$  are called *homogeneous complementarity polynomials* because of their close connection with complementarity polynomials and their homogeneity.

Complementarity polynomials succinctly summarize the information trade-off among various parameters in state estimation with coherent measurements. There

is some evidence that such trade-off also applies to arbitrary measurements on a given number of identically prepared systems, as to be explained shortly. The implications of these polynomials can be elaborated in three aspects: First, to a large extent, these polynomials determine how much efficiency can be improved with coherent measurements as compared with separable measurements. Second, they can serve as an indicator about when the analysis on asymptotic tomographic efficiency (see Sec. 6.2.2) is approximately applicable. Third, they reflect the importance of adaption and are thus crucial to understanding the differences between collective measurements and individual measurements in quantum state estimation (see Sec. 6.4 for the case of a qubit).

**Conjecture 6.3** *The GMT is upper bounded by  $t_\mu(\rho)$  for any measurement on the subspace  $\mathcal{S}_\mu$  and by  $t^N(\rho)$  on the tensor space  $\mathcal{H}^{\otimes N}$ . In each case, the bound is saturated if and only if the measurement is coherent.*

According to Eq. (6.44), Conjecture 6.3 is a consequence of the following conjecture, which is mathematically much more amenable.

**Conjecture 6.4** *Any state  $|\Psi\rangle$  in  $\mathcal{S}_\mu$  satisfies the inequality*

$$\frac{\langle \Psi |^{\otimes 2} \Lambda_{|\mu|}(\rho) | \Psi \rangle^{\otimes 2}}{\langle \Psi | \rho^{\otimes |\mu|} | \Psi \rangle} \leq \langle \Psi | T_\mu(\rho) | \Psi \rangle, \quad (6.56)$$

where  $|\mu| := \sum_{j=1}^d \mu_j$ , and the inequality is saturated if and only if  $|\Psi\rangle$  is a coherent state.

In addition to offering a promising approach for investigating Conjecture 6.3, this conjecture also furnishes an alternative characterization of coherent states.

For the  $N$ -partite symmetric subspace and the bipartite antisymmetric subspace, Conjecture 6.4 is proved in Appendix E.2. In conjunction with Eq. (6.54), we have

**Lemma 6.5** *The GMT is upper bounded by  $N(d + N - 1)s_{[N-1]}(\rho) - N^2 s_{[N]}(\rho)$  for any measurement on the  $N$ -partite symmetric subspace and by  $d - 3 + 2\text{tr}(\rho^2)$  on the bipartite antisymmetric subspace. In each case, the upper bound is saturated if and only if the measurement is coherent.*

### 6.3. Quantum state estimation with coherent measurements

---

Lemma 6.5 confirms Conjecture 6.3 when  $N = 2$  or  $d = 2$ . The first case is obvious since the symmetric subspace and the antisymmetric subspace are the only two irreducible subspaces in  $\mathcal{H}^{\otimes N}$ . To tackle the second case, note that any state in  $\mathcal{S}_{[k^d] + \mu}$  is a tensor product of a  $d$ -partite Slater-determinant state and a state in  $\mathcal{S}_\mu$ . If Conjecture 6.3 holds for  $\mathcal{S}_\mu$ , then it also holds for  $\mathcal{S}_{[k^d] + \mu}$ . Now our claim follows from the observation that each partition in the case  $d = 2$  has the form  $[k^2] + [j]$  for some integers  $j$  and  $k$ .

**Theorem 6.6** *The GMT of any measurement on  $\mathcal{H}^{\otimes 2}$  is upper bounded by  $3(d - 1)$ . When  $d = 2$ , the GMT of any measurement on the subspace  $\mathcal{S}_\mu$  is upper bounded by  $t_\mu(\rho)$ , and that of any measurement on  $\mathcal{H}^{\otimes N}$  is upper bounded by  $t^N(\rho)$ . In each case, the upper bound is saturated if and only if the measurement is coherent.*

To illustrate the improvement of collective measurements over separable measurements, the polynomials  $t^N(\rho)$  for  $N = 1, 2, 3, 4$  are listed below (more details will be presented elsewhere [277]),

$$\begin{aligned} t^1(\rho) &= (d - 1), & t^2(\rho) &= 3(d - 1), \\ t^3(\rho) &= \frac{16d - 17}{3} + \frac{2 - d}{3} \operatorname{tr} \rho^2, \\ t^4(\rho) &= \frac{1}{12} [98d - 111 - (9d - 27) \operatorname{tr} \rho^2 - 5d \operatorname{tr} \rho^3]. \end{aligned} \tag{6.57}$$

When  $N = 1, 2$  or when  $N = 3$  and  $d = 2$ , the polynomial  $t^N(\rho)$  is independent of the eigenvalues of  $\rho$ ; for example,  $t^2(\rho)$  is three times as large as  $t^1(\rho)$ . In general,  $t^N(\rho)$  is larger for states with low purities, in which case collective measurements are more effective.

#### 6.3.5 Estimation of highly mixed states with collective measurements

In this section, we determine the maximal GMT at the point  $\rho = 1/d$  over all possible measurements on  $\mathcal{H}^{\otimes N}$  and confirm Conjecture 6.3 in this special case. This result allows us to compute the minimal MSE of any unbiased estimator. Our study reveals a formal connection between the Pauli-exclusion principle [5] and the optimal state estimation, whose implications are yet to be explored.

Given the affine parametrization specified in Eq. (2.19), the SLD Fisher information matrix is  $d$  times the identity matrix when  $\theta = 0$ , according to Sec. 5.2 or 6.2. Consider a rank-one measurement on  $\mathcal{S}_\mu$  composed of the outcomes  $\Pi_{\mu,\xi} = w_\xi |\Psi_{\mu,\xi}\rangle\langle\Psi_{\mu,\xi}|$ , where  $\sum_\xi w_\xi = D_\mu$ . The Fisher information matrix and the GMT are respectively given by

$$\begin{aligned} I_{\mu,jk} &= d^{2-N} \sum_\xi w_\xi \operatorname{tr}\{E_j Q(\Psi_{\mu,\xi})\} \operatorname{tr}\{E_k Q(\Psi_{\mu,\xi})\}, \\ \frac{1}{d} \operatorname{tr}(I_\mu) &= d^{1-N} \sum_\xi w_\xi \left( \operatorname{tr}\{Q(\Psi_{\mu,\xi})^2\} - \frac{N^2}{d} \right). \end{aligned} \quad (6.58)$$

where  $Q(\Psi) := \sum_{m=1}^N \operatorname{tr}_{\hat{m}}(|\Psi\rangle\langle\Psi|)$ , and “ $\operatorname{tr}_{\hat{m}}$ ” means taking the trace of all the parties except  $m$ .

For the antisymmetric subspace, the (nonzero) eigenvalues of  $Q(\Psi)$  may be interpreted as the occupation numbers of certain single-particle states. According to the Pauli-exclusion principle, they are no larger than one, and they are all equal to one if and only if  $|\Psi\rangle$  is a Slater-determinant state. Therefore, the maximum of the GMT for the antisymmetric subspace is formally determined by the Pauli-exclusion Principle. Remarkably, a similar connection pertains to subspaces with mixed symmetry. According to Theorem 3 of Altunbulak and Klyachko [5], given any normalized state  $|\Psi\rangle$  in  $\mathcal{H}_\mu$ , the eigenvalues of  $Q(\Psi)$  (arranged in decreasing order) is majorized by  $\mu$ , and the equality is realized if and only if  $|\Psi\rangle$  is a coherent state. Noticing that  $\operatorname{tr}\{Q(\Psi)^2\}$  is a Schur-convex function of the eigenvalues, we have

$$\frac{1}{d} \operatorname{tr}(I_\mu) \leq d^{1-N} D_\mu \left( \max_{|\Psi\rangle \in \mathcal{S}_\mu} \operatorname{tr}\{Q(\Psi)^2\} - \frac{N^2}{d} \right) = d^{1-N} D_\mu \left( \mu^2 - \frac{N^2}{d} \right), \quad (6.59)$$

where  $\mu^2 := \sum_j \mu_j^2$ , and the maximum is attained if and only if each outcome is a coherent state up to normalization.

According to the above analysis, the values of the complementarity polynomials  $t_\mu(\rho)$  and  $t^N(\rho)$  at the point  $\rho = 1/d$  are given by

$$t_\mu\left(\frac{1}{d}\right) = d^{1-N} D_\mu \left( \mu^2 - \frac{N^2}{d} \right), \quad t^N\left(\frac{1}{d}\right) = \sum_\mu d^{1-N} d_\mu D_\mu \left( \mu^2 - \frac{N^2}{d} \right). \quad (6.60)$$

### 6.3. Quantum state estimation with coherent measurements

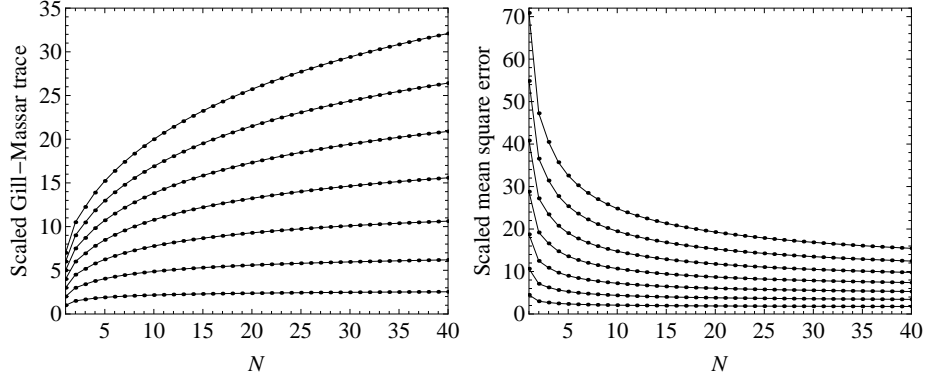


Figure 6.2: The maximal scaled GMT (left) and the minimal scaled MSE (right) at  $\rho = 1/d$  over all collective measurements on  $\rho^{\otimes N}$  for  $d = 2, 3, \dots, 8$  (from bottom to top).

And they are the maxima of the GMTs over all measurements on  $\mathcal{S}_\mu$  and on  $\mathcal{H}^{\otimes N}$ , respectively, in agreement with Conjecture 6.3 in the case  $\rho = 1/d$ . To illustrate the improvement of collective measurements over separable measurements, the values of the maximal scaled GMT for  $N = 1, 2, 3, 4$  are listed below,

$$\frac{t^N\left(\frac{1}{d}\right)}{N} = \begin{cases} d-1, & N=1, \\ \frac{3}{2}(d-1), & N=2, \\ \frac{2(8d-1)(d-1)}{9d}, & N=3, \\ \frac{(49d-11)(d-1)}{24d}, & N=4. \end{cases} \quad (6.61)$$

The left plot of Fig. 6.2 shows the maximal scaled GMT for  $d = 2, 3, \dots, 8$  and  $N = 1, 2, \dots, 40$ . The maximum increases monotonically with  $N$  although with a decreasing slope. Compared with the maximal GMT over separable measurements, there is 50% improvement for collective measurements on two identically prepared systems. When  $N \geq 3$ , the improvement of collective measurements is more significant for large dimensions.

The asymptotic maximal scaled GMT is determined by the following theorem, which is proved in Appendix E.3.

#### Theorem 6.7

$$\lim_{N \rightarrow \infty} \frac{t^N\left(\frac{1}{d}\right)}{N} = d^2 - 1. \quad (6.62)$$

The asymptotic GMT saturates the RLD bound (see Sec. 6.2.2) and is  $d + 1$  times the maximal value over separable measurements. Surprisingly, the maximal information about each parameter can be extracted almost simultaneously in the asymptotic limit when the states of interest are nearly completely mixed.

The maximal scaled GMT sets a lower bound  $N(d^2 - 1)^2/[dt^N(1/d)]$  for the minimal scaled MSE or, equivalently, for the minimal scaled MSH. The bound can be saturated by the covariant coherent measurement, whose Fisher information matrix is proportional to the identity matrix. In other words, the minimal scaled MSE is inversely proportional to the maximal scaled GMT, as depicted in the right plot of Fig. 6.2. In the large- $N$  limit, the minimal scaled MSE  $d - 1/d$  saturates the RLD bound (see Sec. 6.2.2) and is  $d + 1$  times smaller than the corresponding value for separable measurements. Therefore, collective measurements can improve the scaling behavior of the tomographic efficiency with the dimension of the Hilbert space.

## 6.4 Collective measurements in qubit state estimation

Qubit state estimation with collective measurements has received intensive attention in the past two decades. Most studies in the literature were based on the Bayesian approach, which allows deriving optimal solutions in certain scenarios without any assumption on the sample size. For example, coherent measurements are known to maximize the mean fidelity in both the pure-state model [48, 128, 191] and the mixed-state model [18, 19, 20, 254, 259] given an isotropic prior. The disadvantage of this approach lies in the difficulty in determining how the tomographic efficiency improves with the increasing power in performing collective measurements since the optimal estimation strategy usually entails a one-shot measurement on all the samples available. Owing to technical reasons, most studies based on the CR approach focused on the asymptotic regime, assuming the capability of performing arbitrary collective measurements [124, 126, 137]. One exception was the work of Slater [252] that built on the earlier work of Vidal et al. [259].

In this section, we apply the theory developed in Sec. 6.3 to studying qubit state

## 6.4. Collective measurements in qubit state estimation

---

estimation with collective measurements. Our main goal is to quantify the improvement in the tomographic efficiency resulting from the increasing power in performing collective measurements, so as to bridge the gap between asymptotic state estimation and state estimation based on individual measurements. The distinctive features of collective measurements are also discussed in detail.

### 6.4.1 A lower bound for the weighted mean square error

In this section, we introduce a lower bound for the WMSE based on the generalized GM bound and the RLD bound. For a qubit system, the bound is almost tight for a wide range of WMSEs including the MSH and the MSB, as we shall see later.

Consider state estimation by repeated measurements on  $\rho^{\otimes N}$ . We have known three general bounds for the scaled WMSE for any unbiased estimator: the SLD bound [139, 141], the RLD bound [147, 274], and the GM bound [107]. For a  $\mathcal{D}$ -invariant model, the RLD bound is tighter than the SLD bound [137, 147] (see Sec. 6.2). The original GM bound does not apply to entangled measurements, but a straightforward generalization does. Let  $t$  be the maximal scaled GMT (here the dependence on  $N$  and  $\rho$  is suppressed for simplicity); then the scaled WMSE is lower bounded by

$$\mathcal{E}_W = \frac{(\text{tr} \sqrt{W^{1/2} J^{-1} W^{1/2}})^2}{t} = \frac{(\text{tr} \sqrt{J^{-1/2} W J^{-1/2}})^2}{t}, \quad (6.63)$$

according to similar reasoning in Sec. 5.3. If the lower bound is saturated, the scaled MSE matrix and the Fisher information matrix are given by

$$\frac{1}{C_W} = I_W = t J^{1/2} \frac{\sqrt{J^{-1/2} W J^{-1/2}}}{\text{tr} \sqrt{J^{-1/2} W J^{-1/2}}} J^{1/2}. \quad (6.64)$$

The parameter  $t$  in the above equations can be replaced by the scaled complementarity polynomial  $t^N(\rho)/N$  if Conjecture 6.3 holds, as is the case when  $N = 2$  or  $d = 2$  (see Theorem 6.6).

The generalized GM bound is usually more informative for states with low purity, in sharp contrast with the RLD bound, which exhibits the opposite behavior. Therefore,

it is advisable to combine the two bounds. A naive combination is the maximum of the two bounds; a better alternative is the *joint bound* defined as follows,

$$\tilde{\mathcal{E}}_W := \min_I \{ \text{tr}(WI^{-1}) \mid \text{tr}(J^{-1}I) = t, I \leq \tilde{J} \}. \quad (6.65)$$

For a qubit system, the three components of the Bloch vector provide a convenient parametrization of the true state,  $\rho = (1 + \mathbf{r} \cdot \boldsymbol{\sigma})/2$ . Here we assume that the Bloch vector of the true state is aligned with  $\sigma_z$ ; that is,  $\mathbf{r} = (0, 0, r)$ . The SLD and RLD Fisher information matrices follow from Eqs. (6.11) and (6.12), with the result [107, 137]

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{1-r^2} \end{pmatrix}, \quad \tilde{J} = \frac{1}{1-r^2} \begin{pmatrix} 1 & ir & 0 \\ -ir & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (6.66)$$

For any WMSE based on a unitarily invariant distance, like the MSH and the MSB, the weight matrix has a diagonal form  $W = \text{diag}(w_1, w_1, w_3)$ , so the generalized GM bound reduces to

$$\mathcal{E}_W = \frac{(2\sqrt{w_1} + \sqrt{(1-r^2)w_3})^2}{t}. \quad (6.67)$$

It is saturated if the Fisher information matrix is equal to

$$I_W = \frac{t \text{diag}(\sqrt{w_1}, \sqrt{w_1}, \sqrt{(1-r^2)^{-1}w_3})}{2\sqrt{w_1} + \sqrt{(1-r^2)w_3}}. \quad (6.68)$$

Simple analysis shows that the Fisher information matrix saturating the joint bound can be chosen to be diagonal with  $I_{22} = I_{11}$ , in which case Eq. (6.65) reduces to

$$\tilde{\mathcal{E}}_W = \min_{I_{11}, I_{33}} \left\{ \frac{2w_1}{I_{11}} + \frac{w_3}{I_{33}} \mid 2I_{11} + (1-r^2)I_{33} = t, I_{11} \leq \frac{1}{1+r}, I_{33} \leq \frac{1}{1-r^2} \right\}. \quad (6.69)$$

The solution is

$$\tilde{\mathcal{E}}_W = \begin{cases} \mathcal{E}_W & \text{if } I_{W,11} \leq \frac{1}{1+r} \text{ and } I_{W,33} \leq \frac{1}{1-r^2}, \\ 2(1+r)w_1 + \frac{(1+r)(1-r^2)w_3}{(1+r)t-2} & \text{if } I_{W,11} > \frac{1}{1+r}, \\ (1-r^2)w_3 + \frac{4w_1}{t-1} & \text{if } I_{W,33} > \frac{1}{1-r^2}. \end{cases} \quad (6.70)$$



#### 6.4. Collective measurements in qubit state estimation

---

The joint bound reduces to the GM bound when  $N = 1$  and converges to the RLD bound in the limit  $N \rightarrow \infty$ . The bound can be saturated in both cases. In general, any measurement (if there is one) saturating the generalized GM bound or the joint bound has GMT equal to the complementarity polynomial  $t^N(\rho)$  and is thus necessarily coherent according to Theorem 6.6.

For the MSH, the weight matrix is equal to  $\frac{1}{2}$ , and the generalized GM bound is

$$\mathcal{E}_{\text{SH}}^{\text{GM}} = \frac{(2 + \sqrt{1 - r^2})^2}{2t}. \quad (6.71)$$

For a coherent measurement, the bound is saturated if and only if  $I$  is diagonal and

$$I_{11} = I_{22} = \sqrt{1 - r^2} I_{33}. \quad (6.72)$$

The joint bound is

$$\tilde{\mathcal{E}}_{\text{SH}} = \begin{cases} \frac{(2 + \sqrt{1 - r^2})^2}{2t} & \text{if } \frac{t}{2 + \sqrt{1 - r^2}} \leq \frac{1}{1 + r}, \\ (1 + r) + \frac{(1 + r)(1 - r^2)}{2[(1 + r)t - 2]} & \text{otherwise,} \end{cases} \quad (6.73)$$

where in deriving the equation we have exploited the fact that the third case in Eq. (6.70) never occurs. For the MSB, the weight matrix is equal to  $J/4$  (see Sec. 5.2), and the generalized GM bound is

$$\mathcal{E}_{\text{SB}}^{\text{GM}} = \frac{9}{4t}. \quad (6.74)$$

It is saturated if and only if  $I$  is diagonal and

$$I_{11} = I_{22} = (1 - r^2) I_{33}. \quad (6.75)$$

The joint bound is

$$\tilde{\mathcal{E}}_{\text{SB}} = \begin{cases} \frac{9}{4t} & \text{if } \frac{t}{3} \leq \frac{1}{1 + r}, \\ \frac{1}{2}(1 + r) + \frac{(1 + r)}{4[(1 + r)t - 2]} & \text{otherwise.} \end{cases} \quad (6.76)$$

In the pure-state limit, the joint bound for the MSH coincides with the RLD bound, which can be saturated by covariant measurements on individual systems (see Sec. 4.4). When  $r = 0$ , the joint bounds for both figures of merit coincide with the generalized GM bounds, which can be saturated by suitable coherent measurements (see Sec. 6.3.5). In general, the bounds may not be saturated because  $I_{11}$  of any coherent measurement is smaller than  $I_{W,11}$  and  $1/(1+r)$ . Nevertheless, they can be saturated approximately with a high precision, as we shall see in Sec. 6.4.4.

### 6.4.2 Fisher information matrices for coherent measurements

When  $d = 2$ , the irreducible components of  $\mathcal{H}^{\otimes N}$  are specified by two nonnegative integers  $\mu_1$  and  $\mu_2$  that satisfy  $\mu_1 + \mu_2 = N$  and  $\mu_2 \leq \mu_1$ . The dimensions of  $\mathcal{S}_\mu$  and  $\mathcal{K}_\mu$  are given by

$$D_\mu = \mu_1 - \mu_2 + 1, \quad d_\mu = \frac{N!(\mu_1 - \mu_2 + 1)}{(\mu_1 + 1)!\mu_2!}. \quad (6.77)$$

The Schur symmetric polynomial  $s_\mu(r)$  reduces to

$$s_\mu(r) = \frac{(\lambda_1 \lambda_2)^{\mu_2} (\lambda_1^{q+1} - \lambda_2^{q+1})}{\lambda_1 - \lambda_2} = \frac{|\rho|^{\mu_2} (\lambda_1^{q+1} - \lambda_2^{q+1})}{\lambda_1 - \lambda_2}, \quad (6.78)$$

where  $q = \mu_1 - \mu_2$ ,  $\lambda_{1,2} = (1 \pm r)/2$ , and  $|\rho| := \det(\rho) = \lambda_1 \lambda_2 = (1 - r^2)/4$ . In the pure-state limit,  $s_\mu(r)$  converges to 1 if  $\mu_2 = 0$  and to 0 otherwise.

Any coherent measurement on  $\mathcal{S}_\mu$  has outcomes of the form  $\Pi_\xi = a_\xi |\Psi_\xi\rangle\langle\Psi_\xi|$  with  $|\Psi_\xi\rangle = |\Psi_-\rangle^{\otimes \mu_2} \otimes |\psi_\xi\rangle^{\otimes q}$ , where  $|\Psi_-\rangle$  is the singlet and  $|\psi_\xi\rangle$  is a single-particle state such that  $\{|\psi_\xi\rangle, a_\xi\}$  forms a weighted  $q$ -design. For example, any minimal coherent measurement on the bipartite symmetric subspace is composed of four outcomes  $\frac{3}{4}(|\psi_\xi\rangle\langle\psi_\xi|)^{\otimes 2}$ , such that the states  $|\psi_\xi\rangle\langle\psi_\xi|$  form a SIC POM; accordingly, any minimal coherent measurement on  $\mathcal{H}^{\otimes 2}$  is composed of five outcomes, which include the singlet in addition to the above four outcomes.

Given the coherent measurement mentioned above, the probability of obtaining the outcome  $\Pi_\xi$  has the form  $p_\xi = |\rho|^{\mu_2} (\langle\psi_\xi|\rho|\psi_\xi\rangle)^q$ . The Fisher information matrix is thus

#### 6.4. Collective measurements in qubit state estimation

---

given by

$$\begin{aligned}
(I_\mu)_{jk} &= \mu_2^2 |\rho|^{\mu_2-2} s_{[q]}(r) \frac{\partial |\rho|}{\partial r_j} \frac{\partial |\rho|}{\partial r_k} + |\rho|^{\mu_2} (I_{[q]})_{jk} \\
&\quad + \mu_2 |\rho|^{\mu_2-1} \left( \frac{\partial |\rho|}{\partial r_j} \frac{\partial s_{[q]}(r)}{\partial r_k} + \frac{\partial |\rho|}{\partial r_k} \frac{\partial s_{[q]}(r)}{\partial r_j} \right) \\
&= \frac{\mu_2 r}{4} |\rho|^{\mu_2-2} \delta_{j3} \delta_{k3} [\mu_2 r s_{[q]}(r) - 4 |\rho| s'_{[q]}(r)] + |\rho|^{\mu_2} (I_{[q]})_{jk}, \tag{6.79}
\end{aligned}$$

where  $I_{[q]}$  is the Fisher information matrix associated with the measurement on  $\mathcal{S}_{[q]}$  that is composed of the outcomes  $a_\xi(|\psi_\xi\rangle\langle\psi_\xi|)^{\otimes q}$ . In deriving the last equality we have taken into account the assumption that the Bloch vector of  $\rho$  is aligned with  $\sigma_z$ . The derivative of  $s_{[q]}(r)$  with respect to  $r$  can be computed by means of Eq. (6.78), with the result

$$s'_{[q]}(r) = \frac{(q+1)(\lambda_1^q + \lambda_2^q)}{2(\lambda_1 - \lambda_2)} - \frac{\lambda_1^{q+1} - \lambda_2^{q+1}}{(\lambda_1 - \lambda_2)^2} = \frac{(q-1)s_{[q]}(r) - (q+1)|\rho|s_{[q-2]}(r)}{2(\lambda_1 - \lambda_2)}. \tag{6.80}$$

Strictly speaking, the second equality in the above equation is valid only when  $q \geq 2$ ; nevertheless, this restriction can be lifted by extending the definition of  $s_{[q]}(r)$  according to Eq. (6.78).

To discuss the properties of the Fisher information matrix in more detail, we need to distinguish several cases depending on the value of  $q$ . When  $q \neq 1$ , Eq. (6.40) implies that

$$(I_{[q]})_{jk} = \frac{q}{q-1} \frac{\partial^2 s_{[q]}(r)}{\partial r_j \partial r_k} = \frac{q}{q-1} \frac{\partial^2 s_{[q]}(r)}{\partial r_j^2} \delta_{jk}; \tag{6.81}$$

accordingly, Eq. (6.79) reduces to

$$(I_\mu)_{jk} = \frac{\mu_2 r}{4} |\rho|^{\mu_2-2} \delta_{j3} \delta_{k3} [\mu_2 r s_{[q]}(r) - 4 |\rho| s'_{[q]}(r)] + \frac{q}{q-1} |\rho|^{\mu_2} \frac{\partial^2 s_{[q]}(r)}{\partial r_j^2} \delta_{jk}. \tag{6.82}$$

The Fisher information matrix  $I_\mu$  is diagonal with diagonal entries

$$I_{\mu,11}(r) = I_{\mu,22}(r) = f_{\mu,1}(r) + f_{\mu,1}(-r), \quad I_{\mu,33}(r) = f_{\mu,3}(r) + f_{\mu,3}(-r), \tag{6.83}$$

where

$$\begin{aligned}
 f_{\mu,1}(r) &= \frac{q|\rho|^{\mu_2}}{2r^3} \left[ -\frac{q(q+1)|\rho|}{(q-1)} \left(\frac{1+r}{2}\right)^{q-1} - q \left(\frac{1+r}{2}\right)^{q+1} + (q+1) \left(\frac{1+r}{2}\right)^q \right], \\
 f_{\mu,3}(r) &= \frac{|\rho|^{\mu_2-2}}{4r^3} \left[ (\mu_2 r^2 - 2q|\rho|)^2 \left(\frac{1+r}{2}\right)^{q+1} + (q+1)|\rho|(Nr^2 - q) \left(\frac{1+r}{2}\right)^q \right. \\
 &\quad \left. + \frac{q^2(q+1)|\rho|^2}{(q-1)} \left(\frac{1+r}{2}\right)^{q-1} \right]. \tag{6.84}
 \end{aligned}$$

In the special case  $q = 0$ , that is,  $\mu_1 = \mu_2 = N/2$ , we have

$$I_{\mu,jk}(r) = \frac{N^2}{16} r^2 |\rho|^{\mu_2-2} \delta_{j3} \delta_{k3}. \tag{6.85}$$

When  $q = 1$ , because  $s_{[q]}(r) = 1$  and  $s'_{[q]}(r) = 0$ , Eq. (6.79) reduces to

$$(I_{\mu})_{jk} = \frac{\mu_2^2 r^2}{4} |\rho|^{\mu_2-2} \delta_{j3} \delta_{k3} + |\rho|^{\mu_2} (I_{[1]})_{jk}. \tag{6.86}$$

Note that  $I_{\mu}$  is determined by the Fisher information matrix of a measurement on a single copy of  $\rho$ .

For concreteness, let us take the covariant coherent measurement as an example. Let  $\Pi_{\mu,\mathbf{v}} = \rho_-^{\otimes \mu_2} \otimes \rho_{\mathbf{v}}^{\otimes q}$ , where  $\rho_- = |\Psi_- \rangle \langle \Psi_-|$  and  $\rho_{\mathbf{v}}$  is the single-particle pure state defined by the Bloch vector  $\mathbf{v} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ . Then the outcome of the covariant coherent measurement on  $\mathcal{S}_{\mu}$  can be written as  $\Pi_{\mu,\mathbf{v}} D_{\mu} \sin \theta d\theta d\phi / 4\pi$ , and the probability density of obtaining the outcome is  $p(\mathbf{v}|\mathbf{r}) = |\rho|^{\mu_2} [(1 + \mathbf{v} \cdot \mathbf{r})/2]^q$ . The Fisher information matrix is given by

$$\begin{aligned}
 I_{\mu,jk}(r) &= D_{\mu} \int \frac{d\theta d\phi \sin \theta}{4\pi} \frac{1}{p(\mathbf{v}|\mathbf{r})} \frac{\partial p(\mathbf{v}|\mathbf{r})}{\partial r_j} \frac{\partial p(\mathbf{v}|\mathbf{r})}{\partial r_k} \\
 &= D_{\mu} \int \frac{d\theta d\phi \sin \theta}{4\pi} |\rho|^{\mu_2-2} \left(\frac{1+r \cos \theta}{2}\right)^{q-2} \left( -\frac{\mu_2 r}{2} \frac{1+r \cos \theta}{2} \delta_{j3} + q|\rho| \frac{v_j}{2} \right) \\
 &\quad \times \left( -\frac{\mu_2 r}{2} \frac{1+r \cos \theta}{2} \delta_{k3} + q|\rho| \frac{v_k}{2} \right). \tag{6.87}
 \end{aligned}$$

When  $q = 0$  or  $q \geq 2$ , the Fisher information matrix is identical with that in Eq. (6.83) as expected. Otherwise, it takes on the form in Eq. (6.86), where  $I_{[1]}$  is diagonal with

#### 6.4. Collective measurements in qubit state estimation

---

diagonal entries given by

$$I_{[1],22} = I_{[1],11} = \frac{1}{4} \left( \frac{2}{r^2} + \frac{r^2 - 1}{r^3} \ln \frac{1+r}{1-r} \right), \quad I_{[1],33} = \frac{1}{2} \left( -\frac{2}{r^2} + \frac{1}{r^3} \ln \frac{1+r}{1-r} \right). \quad (6.88)$$

When  $N$  is even, the total Fisher information matrix  $I^N(r) = \sum_{\mu} d_{\mu} I_{\mu}(r)$  is identical for all coherent measurements as is  $I_{\mu}$ . Otherwise, the same holds for  $I_{\mu}$  except when  $\mu = [(N+1)/2, (N-1)/2]$ . In the large- $N$  limit, the contribution from this irreducible component is negligible, so the Fisher information matrix is almost independent of specific coherent measurements even if  $N$  is odd.

When  $N \geq 2$ , according to Eq. (6.79) or (6.83), in the pure-state limit,

$$I_{\mu}(r) \rightarrow \begin{cases} \text{diag}\left(\frac{N}{2}, \frac{N}{2}, \frac{N(N^2-3N+4)}{4(N-1)}\right) & \text{if } \mu = [N], \\ \text{diag}\left(0, 0, \frac{1}{1-r^2}\right) & \text{if } \mu = [N-1, 1], \\ \text{diag}(0, 0, 1) & \text{if } \mu = [N-2, 2], \\ 0 & \text{otherwise.} \end{cases} \quad (6.89)$$

Noting that the major contribution to the Fisher information matrix stems from only two kinds of irreducible components, we get

$$I^N(r) \rightarrow \text{diag}\left(\frac{N}{2}, \frac{N}{2}, \frac{N-1}{1-r^2}\right). \quad (6.90)$$

The asymptotic scaled Fisher information matrix of any coherent measurement is determined by the following theorem, whose proof is relegated to Appendix E.4.

**Theorem 6.8** *In the large- $N$  limit, the scaled Fisher information matrix of any coherent measurement on  $N$  identically prepared qubit systems is given by*

$$\lim_{N \rightarrow \infty} \frac{I^N(r)}{N} = \text{diag}\left(\frac{1}{1+r}, \frac{1}{1+r}, \frac{1}{1-r^2}\right). \quad (6.91)$$

The limit is independent of the specific coherent measurement and is maximal among all Fisher information matrices that commute with and are upper bounded by the RLD Fisher information matrix. These crucial properties guarantee that all coherent measurements are optimal globally in the asymptotic limit with respect to any WMSE

based on a unitarily invariant distance (see also Secs. 6.2 and 6.4.4).

### 6.4.3 Complementarity polynomials

The complementarity polynomial for the subspace  $\mathcal{S}_\mu$  can be derived based on Eqs. (6.66) and (6.79), with the result

$$t_\mu(r) = \mu_2 r |\rho|^{\mu_2-1} [\mu_2 r s_{[q]}(r) - 4|\rho| s'_{[q]}(r)] + |\rho|^{\mu_2} t_{[q]}(r) = f_\mu(r) + f_\mu(-r), \quad (6.92)$$

where  $t_{[q]}(r) = q(q+1)s_{[q-1]}(r) - q^2 s_{[q]}(r)$  [see Eq. (6.54)] and

$$f_\mu(r) = \frac{1}{r} \left( \frac{1+r}{2} \right)^{\mu_1} \left( \frac{1-r}{2} \right)^{\mu_2-1} \left[ \frac{(Nr-q)^2}{4} + \frac{N(1-r)}{2} \right]. \quad (6.93)$$

When  $q \geq 2$ , the polynomial  $t_\mu(r)$  has a simple decomposition in terms of Schur symmetric polynomials,

$$\begin{aligned} t_\mu(r) &= \mu_2^2 r^2 |\rho|^{\mu_2-1} s_{[q]}(r) + 2\mu_2(q+1) |\rho|^{\mu_2+1} s_{[q-2]}(r) \\ &\quad + |\rho|^{\mu_2} [q(q+1)s_{[q-1]}(r) - (qN - 2\mu_2)s_{[q]}(r)]. \end{aligned} \quad (6.94)$$

This equation is also applicable when  $q = 0, 1$  if the definition of  $s_{[q]}(r)$  is extended according to Eq. (6.78).

According to Theorem 6.6, the complementarity polynomial  $t^N(r)$  is the maximum of the GMT over all measurements on  $\rho^{\otimes N}$ . Calculation shows that it equals the constants 1, 3, 5 when  $N = 1, 2, 3$ , but decreases monotonically with  $r$  when  $N \geq 4$  (see Fig. 6.3). In the large- $N$  limit, the scaled GMT can be determined based on Theorem 6.8,

$$\lim_{N \rightarrow \infty} \frac{t^N(r)}{N} = \frac{2}{1+r} + 1. \quad (6.95)$$

Alternatively, it can be derived directly from Eq. (6.92). The asymptotic scaled GMT of any coherent measurement saturates the RLD bound in Eq. (6.24).

More than a decade ago, Vidal et al. [259] investigated optimal estimation of qubit mixed states from the Bayesian perspective and constructed a family of measurements

#### 6.4. Collective measurements in qubit state estimation

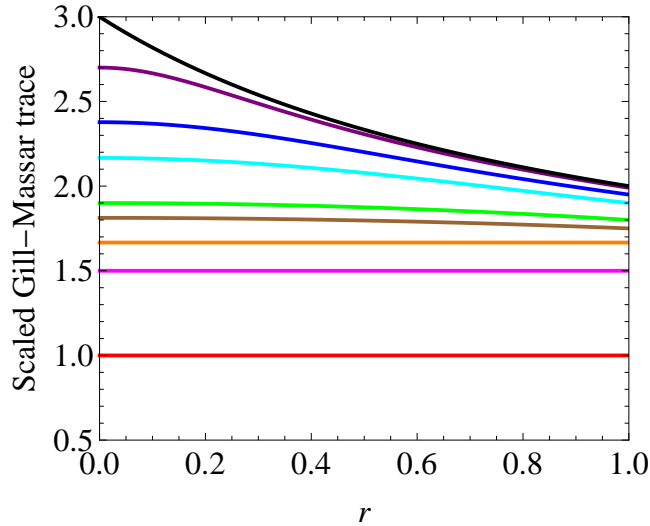


Figure 6.3: The maximal scaled GMT over all measurements on  $N$  copies of a qubit state for  $N = 1, 2, 3, 4, 5, 10, 20, 100, \infty$  (from bottom to top). It does not depend on  $r$  when  $N = 1, 2, 3$ , but decreases monotonically with  $r$  otherwise.

that maximize the mean fidelity over all measurements on  $\rho^{\otimes N}$ . Later, Slater [252] inspected these measurements for  $2 \leq N \leq 7$  from the point-wise perspective and found that the GMTs are polynomials of  $r$ . In addition, he conjectured that the maximal GMT over all measurements on  $\rho^{\otimes N}$  is equal to  $2N - 1$  in the pure-state limit.

The measurements constructed in Ref. [259] are actually coherent in our terminology except that they are not necessarily rank one. Consequently, the polynomials found by Slater are special examples of complementarity polynomials. Based on the previous analysis, we can now confirm Slater’s conjecture. In the pure-state limit, according to Eq. (6.92) or (6.89),  $t_\mu(r)$  converges to

$$t_\mu(1) = \begin{cases} N & \text{if } \mu = [N], \\ 1 & \text{if } \mu = [N - 1, 1], \\ 0 & \text{otherwise.} \end{cases} \quad (6.96)$$

Therefore, the maximal GMT  $t^N(r)$  is equal to  $2N - 1$  in the pure-state limit.

#### 6.4.4 Mean square error and mean square Bures distance

In this section we determine the performances of coherent measurements with respect to the MSH and the MSB. Surprisingly, our study reveals that any coherent measurement is nearly optimal whenever  $N \geq 2$ .

When  $N$  is even, both the scaled MSH and MSB are independent of specific coherent measurements, as is the Fisher information matrix. In the special case  $N = 2$ , the Fisher information matrix  $I^2(r)$  is equal to the SLD Fisher information matrix, while the scaled MSH and MSB are equal to  $3 - r^2$  and  $3/2$ , respectively. The scaled MSB saturates the generalized GM bound in Eq. (6.74), so any coherent measurement on  $\rho^{\otimes 2}$  is optimal in minimizing the MSB globally. Compared with separable measurements (see Secs. 5.3.2 and 5.3.3), coherent measurements can improve the efficiency by 50% without any adaption.

When  $N$  is odd, to minimize the scaled MSB over coherent measurements, it suffices to consider Fisher information matrices that are diagonal and satisfy  $I_{11}^N(r) = I_{22}^N(r)$ . When  $N = 1$ , the GM bound can always be saturated [107] (see Sec. 5.3.2). When  $N \geq 3$  and  $r$  is smaller than a certain threshold  $r_{\text{th}}$ , it is possible to choose a suitable coherent measurement on  $\mathcal{S}_\mu$  for  $\mu = [(N + 1)/2, (N - 1)/2]$  such that Eq. (6.75) is satisfied and that the generalized GM bound is saturated. Otherwise,  $I_{11}^N(r)$  is too small to satisfy the equation, and the generalized GM bound cannot be saturated. A measurement is optimal among coherent measurements if and only if it yields the largest value of  $I_{11}^N(r) = I_{22}^N(r)$ , that is, if  $I_{[1]}$  in Eq. (6.86) is equal to  $\text{diag}(1/2, 1/2, 0)$ ; unfortunately, we are not sure whether such a measurement is optimal among all measurements on  $\rho^{\otimes N}$ . The threshold is determined by solving Eq. (6.75) after setting  $I_{[1]} = \text{diag}(1/2, 1/2, 0)$  in Eq. (6.86). When  $N = 3$ , for example, we have

$$\begin{aligned} I_{[3]}(r) &= \frac{3}{2} \text{diag}(1, 1, 1), & I_{[2,1]}(r) &= \text{diag}\left(0, 0, \frac{r^2}{1-r^2}\right) + \frac{1-r^2}{4} I_{[1]}, \\ I^3(r) &= I_{[3]}(r) + 2I_{[2,1]}(r) = \text{diag}\left(\frac{3}{2}, \frac{3}{2}, \frac{3+r^2}{2-2r^2}\right) + \frac{1-r^2}{2} I_{[1]}, \end{aligned} \quad (6.97)$$

so the threshold is  $r_{\text{th}} = 1/\sqrt{3}$ . The same analysis also applies to minimizing the scaled



#### 6.4. Collective measurements in qubit state estimation

---

MSH, except that Eq. (6.75) should be replaced by Eq. (6.72) and, as a consequence, the threshold is smaller. For example, the threshold is approximately equal to 0.39 when  $N = 3$ .

In the pure-state limit, comparison of Eqs. (6.73), (6.76), and (6.89) shows that any coherent measurement saturates the joint bounds for the MSH and the MSB (see Sec. 6.4.1) whenever  $N \geq 2$ . For example, the minimal scaled MSB is given by

$$\frac{\mathcal{E}_{\text{SB}}^N(r \rightarrow 1)}{N} = 1 + \frac{N}{4(N-1)}, \quad (6.98)$$

which decreases monotonically with  $N$ . Therefore, increasing power in performing collective measurements implies increasing tomographic efficiency. The minimal scaled MSH is independent of  $N$  since the joint bound converges, in the pure-state limit, to the RLD bound, which is independent of  $N$ . The situation is different once the purity of the true state deviates from the unit (see Fig. 6.4).

In the large- $N$  limit, the scaled MSH and MSB of any coherent measurement are determined by Theorem 6.8,

$$\lim_{N \rightarrow \infty} \frac{\mathcal{E}_{\text{SH}}^N(r)}{N} = \frac{1}{2}(3 + 2r - r^2), \quad \lim_{N \rightarrow \infty} \frac{\mathcal{E}_{\text{SB}}^N(r)}{N} = \frac{1}{4}(3 + 2r). \quad (6.99)$$

The first formula agrees with the analysis of Hayashi and Matsumoto [137] based on a different approach. Both figures of merit saturate the RLD bounds, so all coherent measurements are optimal with respect to them in the asymptotic limit. The same is true for any other figure of merit as long as the weight matrix commutes with the RLD Fisher information matrix. Recall that the Fisher information matrix saturating the RLD bound is independent of the weight matrix under the commutativity assumption (see Sec. 6.2.1).

Figure 6.4 shows the scaled MSH and the scaled MSB of the covariant coherent measurement and the optimal coherent measurement on  $N$  identically prepared qubit states for  $N = 1, 2, 3, 4, 5, 10, 20, 100, \infty$ . The scaled MSH of the optimal coherent measurement decreases monotonically with  $r$  when  $N$  is small but increases monotonically

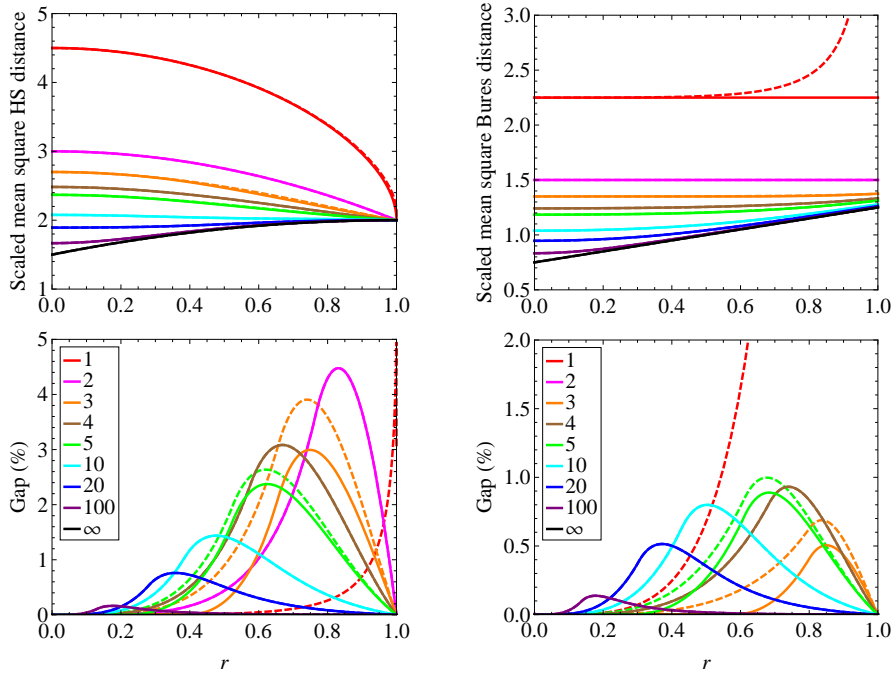


Figure 6.4: The scaled MSHs (upper left) and the scaled MSBs (upper right) of the covariant coherent measurements (dashed) and the optimal coherent measurements (solid) on  $N$  copies of a qubit state for  $N = 1, 2, 3, 4, 5, 10, 20, 100, \infty$  (from top to bottom). For comparison, the lower plots illustrate the gaps between them and the joint bounds (see Sec. 6.4.1). The performances of the two kinds of measurements are identical when  $N$  is even.

when  $N$  is large; calculation shows that it is not monotonic for certain values of  $N$ , such as 13 and 14. By contrast, the scaled MSB is independent of  $r$  when  $N = 1, 2$  but increases monotonically otherwise. When  $N = 1$ , the scaled MSB of the covariant measurement diverges in the pure-state limit. In marked contrast, the scaled MSH saturates the RLD bound and the GM bound in this limit; however, the maximal gap between the covariant measurement and the optimal measurement occurs when the true state is nearly pure. When  $N \geq 2$ , both types of measurements are nearly optimal with respect to both figures of merit, as witnessed by the small gaps from the joint bounds. Further analysis shows that all coherent measurements are nearly optimal globally with respect to both figures of merit, as well as many others, in sharp contrast to state estimation with individual measurements, for which the optimal measurements heavily depend on the true state and the figure of merit (see Chapter 5).

### 6.5 Summary and open problems

We have addressed the major open problem: By how much can the efficiency in quantum state estimation be increased with collective measurements in comparison with individual measurements? The distinctive features of collective measurements and their implications have also been discussed in detail.

First, we investigated asymptotic quantum state estimation and determined the minimal scaled MSE, MSB, and the maximal scaled GMT. Our study showed that collective measurements can improve the scaling behavior of the tomographic efficiency with the dimension of the Hilbert space and that the optimal measurements are universal in the asymptotic limit with respect to a wide range of figures of merit including the three choices mentioned above.

Second, we introduced the concept of coherent measurements as a primitive for understanding quantum state estimation in the case of limited access to collective measurements. We proved that the GMT of any coherent measurement on  $\rho^{\otimes N}$  is a symmetric polynomial—the complementarity polynomial—of the eigenvalues of  $\rho$ . Furthermore, this polynomial is the maximum of the GMT over all possible measurements on  $\rho^{\otimes N}$  when either  $N = 2$  or  $d = 2$ . We believe that this conclusion may hold in general. As the name suggests, complementarity polynomials concisely summarize the information trade-off among complementary aspects of a quantum system. They are useful not only in explicating the efficiency advantage of collective measurements but also in determining the conditions under which the analysis on asymptotic state estimation is applicable. In addition, our study provides a simple framework for understanding the emergence of universality in optimal state estimation as  $N$  increases and the importance of adaption decreases.

In the case of a qubit system, we determined the set of Fisher information matrices of all coherent measurements and calculated all complementarity polynomials. We also analyzed the tomographic efficiencies of the optimal and covariant coherent measurements in terms of the MSE and the MSB, thereby revealing that all coherent measurements are nearly optimal with respect to both figures merit whenever  $N \geq 2$ .

## Chapter 6. Quantum state estimation with collective measurements

---

Interestingly, to achieve nearly optimal performances, it suffices to perform collective measurements that are least entangled as long as they comply with certain symmetry requirement.

There are a few problems left open, which may serve as future research topics.

1. Prove that the complementarity polynomial  $t^N(\rho)$  is the maximum of the GMT over all measurements on  $\rho^{\otimes N}$  and that the maximum is saturated if and only if the measurement is coherent (see Conjectures 6.3 and 6.4; the claim holds when either  $N = 2$  or  $d = 2$  according to Theorem 6.6).
2. Develop more efficient algorithms for computing complementarity polynomials.
3. Design experiments for measuring complementarity polynomials.
4. Explore the implications of complementarity polynomials for foundational issues, such as wave-particle duality.
5. Generalize the analysis on qubit state estimation to more general settings.

# Symmetric informationally complete POMs

---

## 7.1 Introduction

In a  $d$ -dimensional Hilbert space, a *symmetric informationally complete probability operator measurement* (SIC POM) [62, 230, 232, 275] is composed of  $d^2$  subnormalized projectors onto pure states with equal pairwise fidelity. SIC POMs exhibit plenty of elegant properties, which are rooted in their very definition. For example, they are simultaneously minimal 2-designs and maximal sets of equiangular lines [76, 144, 230, 232, 263, 275]; they are optimal in linear quantum state tomography [217, 218, 226, 230, 244, 281] (see Chapter 3) and measurement-based quantum cloning [244]. They are useful in quantum cryptography [84, 91, 100, 230, 231], quantum fingerprinting [246], and signal processing [151]. They also play a crucial role in studying foundational issues [98, 99, 101, 230] and in understanding the geometry of quantum states [13, 29, 30]. Besides, their connections with mutually unbiased bases (MUB) [2, 8, 12, 161, 162, 271] and discrete Wigner functions [67, 230] are hot topics. Recently, they have also attracted the attention of many experimentalists; for example, qubit SIC POMs [80, 84, 183] and qutrit SIC POMs [195] were implemented in experiments. In addition, a novel scheme for realizing SIC POMs by successive measurements was proposed [161, 162].

Every known SIC POM is *group covariant* in the sense that it can be generated from a *fiducial state* by a group composed of unitary operators. Moreover, almost all group covariant SIC POMs known so far are covariant with respect to the *Heisenberg–Weyl* (HW) group [232, 245, 266, 275]. Up to now, analytical solutions of HW covariant

SIC POMs have been constructed in dimensions 2, 3 [76, 275]; 4, 5 [275]; 6 [115]; 7 [6]; 8 [116, 245]; 9–15 [116, 117, 118, 119, 245]; 16 [10]; 19 [6]; and 24, 28, 31, 35, 37, 43, 48 [9, 120, 245]. Numerical solutions with high precision have been found up to dimension 67 [232, 245]. All the evidence supports the belief that HW covariant SIC POMs exist in any Hilbert space of finite dimension. However, there is neither a universal recipe for constructing SIC POMs nor a general proof of their existence, despite the efforts of many researchers in the past decade. Most known solutions are derived from solving systems of nonlinear equations [118, 245] or optimizing certain target functions numerically [232, 245], both of which are computationally very demanding, and it is increasingly more difficult to obtain new solutions without introducing new ideas. Furthermore, understanding the properties of SIC POMs has remained one of the most challenging tasks in this field. Many fundamental questions are awaiting for answers, as explained as follows.

In the mathematical community, SIC POMs have been studied under the name of *equiangular lines* for more than half a century [76, 112, 127, 164, 175, 184, 233]; see Ref. [164] for a historical survey. When the lines are represented by unit kets, the equiangular condition means that the pairwise fidelities among the kets are the same. A cursory inspection of the Gram matrix of the lines reveals that there are at most  $d^2$  equiangular lines in a (complex) Hilbert space [76]. SIC POMs stand out as sets of equiangular lines that saturate the absolute upper bound. When the pairwise fidelity  $\mu$  is smaller than  $1/(d+1)$ , there is a tighter bound for the number  $n$  of lines,

$$n \leq \frac{d - \mu d}{1 - \mu d}, \quad (7.1)$$

which is known as the *Welch bound* [263]. A set of equiangular lines is *tight* if it saturates the Welch bound. Besides, the study of SIC POMs has drawn much inspiration from the study of spherical codes and designs [26, 77, 144, 145, 165, 230, 232, 244], as well as frame theory [61, 81, 230, 232, 244].

In the 1990s, Zauner [275] started to investigate SIC POMs systematically under the general theme of establishing a quantum design theory by generalizing classical combi-

## 7.1. Introduction

---

natorial design theory to the noncommutative setting. He found that many SIC POMs can be generated from fiducial states by certain groups composed of unitary operators, a prominent example being the HW group. For example, he constructed analytical fiducial states of the HW group in dimensions from 2 to 5 and numerical ones in dimensions 6 and 7. It turned out that each of these fiducial states is stabilized by an order-3 unitary transformation, which is now known as *Zauner unitary transformation*. This observation inspired his conjecture that a fiducial state with this additional symmetry exists in every finite dimension. Up to now, Zauner's conjecture has remained one of the most intriguing problems concerning SIC POMs. Unfortunately, Zauner's work was not fully recognized for a long time, partially because the German-written thesis was not accessible to many researchers<sup>1</sup>. A few years later, the study of SIC POMs was revitalized by Renes, Blume-Kohout, Scott, and Caves [232] (see also Refs. [62, 230]), who performed extensive numerical calculations and found fiducial states of the HW group for dimensions up to 45. Their work considerably strengthened the belief that HW covariant SIC POMs exist in all finite dimensions.

An important tool in the study of HW covariant SIC POMs is the *Clifford group*, the normalizer of the HW group within the group of all unitary operators; the extended Clifford group also contains antiunitary operators. The two groups have played an important role in quantum information science, such as in quantum error correction and quantum computation [113, 114]. Their relevance to the study of SIC POMs was first noticed by Grassl [115] and was investigated in detail by Appleby [6, 7] (see also Refs. [10, 11, 95, 278, 279]). The extended Clifford group classify HW covariant SIC POMs into disjoint orbits, such that those on the same orbit are *equivalent* in the sense that they can be turned into each other by unitary or antiunitary transformations. However, it does not help understand the relations among SIC POMs on different orbits. In addition, Appleby [6] introduced the concept of *canonical order-3 Clifford unitary transformations* (see also Ref. [95]), which include Zauner unitary transformation as a special example, and pointed out that every fiducial state found by Renes et al. [232]

---

<sup>1</sup>Recently, a English translation of Zauner's thesis was published [275].

is stabilized by a canonical order-3 Clifford unitary transformation. Inspired by this observation, he formulated a few variants of Zauner's conjecture, some of which have remained elusive up to now. By virtue of symmetry consideration, he also derived fiducial states in dimensions 7 and 19 with a purely analytical approach. Unfortunately, it seems very difficult to generalize his construction to other dimensions [163].

Recently, Scott and Grassl [245] compiled a comprehensive list of numerical fiducial states of the HW group in dimensions up to 67, which is putatively complete up to dimension 50; they also constructed several new analytical fiducial states based on the approach developed earlier by Grassl [115, 116, 117, 118, 119]. Their study not only confirmed Zauner's conjecture [275] in dimensions up to 67 but also revealed the existence of fiducial states with other symmetry, which disproved a variant of Zauner's conjecture formulated by Appleby [6]<sup>2</sup> (see also Ref. [116]). In addition, their work revealed the important role played by Galois field theory [75] in constructing SIC POMs and in understanding their properties, but the potential of this line of thinking is still not clear.

Because of technical reasons, most studies on the symmetry of HW covariant SIC POMs have focused on Clifford operations. Although there is a widespread speculation that all symmetry operations of an HW covariant SIC POM belong to the Clifford group, there is no rigorous proof. In a recent work [279] (see Chapter 8), we proved that any group covariant SIC POM in a prime dimension is covariant with respect to the HW group and that the above speculation holds whenever the prime is not equal to 3. After detailed analysis on the peculiarity in dimension 3, we also established complete equivalence relations among group covariant SIC POMs in dimension 3 and classified inequivalent ones according to the geometric phases associated with fiducial states (see Ref. [64] for an alternative classification method). Our study also clarified the relations among SIC POMs on different orbits of the Clifford group in every prime dimension. In collaboration with our colleagues, we explored the structure of two-qubit HW covariant

---

<sup>2</sup>Zauner's conjecture for dimension 66 was not confirmed in Ref. [245], but by Andrew Scott later, according to private communication with Markus Grassl.



## 7.1. Introduction

---

SIC POMs and showed that the symmetry group of each SIC POM is contained in the Clifford group. On the other hand, our investigation revealed two kinds of additional symmetry for which Clifford operations cannot account [278, 283] (see Chapter 9). The situation in other dimensions is still not well understood and deserves further study.

The quest for SIC POMs with other group symmetry has had a convoluted journey. More than a decade ago, Hoggar [146] constructed a set of 64 equiangular lines in dimension 8, which is known as the *Hoggar lines* and is covariant with respect to the three-qubit Pauli group [275], an alternative version of the HW group. Later, Renes et al. [230, 232] showed with numerics that some other nice error bases [166, 168, 169] can also generate SIC POMs, but they did not give the detail. Using nice error bases with non-Abelian index groups, Grassl constructed an analytical SIC POM in dimensions 6 and 8, respectively [116]. Unfortunately, there is still no systematic study about groups that can generate SIC POMs, and there seem to be more confusions than conclusions about the state of affairs. Even many rudimentary questions are still open. For example, does there exist any SIC POM that is not covariant with respect to the HW group? The main obstacle lies in the difficulty in determining the full symmetry group of a SIC POM and the equivalence relation between two SIC POMs. What is worse, confusions often result from overlooking the fact that some SIC POMs are covariant with respect to more than one group [279].

Most studies on the properties of SIC POMs have presumed group covariance, partially because group covariant SIC POMs are much easier to construct and to analyze. Except in dimension 2, however, there is still no conclusive answer to the basic question: Does there exist a SIC POM that is not group covariant? Besides, in many applications, such as quantum state tomography [226, 230, 244, 281] (see Chapter 3), quantum cryptography [84, 91, 100, 230, 231], and foundational studies [13, 98, 99, 101, 230], the group structure is not essential. Therefore, it is instructive to drop the covariance assumption at a certain stage if we want to get a complete picture about SIC POMs. Recently, a major step along this direction was made by Appleby, Flammia, and Fuchs [14], who demonstrated that the traces of the triple products among states in a SIC POM

characterize the SIC POM up to unitary equivalence. Based on this observation, they established a simple connection between the existence of SIC POMs and the existence of certain structures related to Lie algebras. However, neither existence problem is easy to solve. The full potential of their approach is yet to be explored.

Motivated by the persistent open problems and confusions mentioned above, we study the properties of SIC POMs systematically in the rest of this thesis. Our main concern are the symmetry problem: What symmetry can a SIC POM possess? and the equivalence problem: How can we determine whether two SIC POMs are equivalent or not. In this way, we hope to establish a clear picture about SIC POMs already known and shed some light on those SIC POMs yet to be discovered.

The rest of this chapter sets the stage for later discussions. Several results presented here are also interesting in their own right. We start by introducing the concepts of symmetry group and group covariance. We then derive a necessary condition on the groups that can generate SIC POMs based on the works of Zauner [275] and Grassl [119], thereby revealing the crucial role of nice error bases in the study of SIC POMs. Next, we establish a simple criterion for determining the equivalence relations among SIC POMs that are covariant with respect to the same group. Finally, we review the basic properties of the HW group and the Clifford group following the work of Appleby [6]. For the convenience of later discussions, some supplementary materials concerning the Clifford group are presented in Appendix H.

### 7.2 Symmetry and group covariance

In this section we investigate several fundamental questions pertinent to group covariant SIC POMs: What are the necessary requirements on a group that can generate a SIC POM? What relations exist among SIC POMs that are covariant with respect to the same group.

Let  $G$  be any group composed of unitary operators. The *collineation group*  $\overline{G}$  of  $G$  is derived from  $G$  by identifying operators that are proportional to each other [37]. A group and its collineation group are referred to with the same name when there is

## 7.2. Symmetry and group covariance

---

no confusion. By convention, when the symbol is not specified, the attributes of the collineation group, such as its order, are taken for granted as the attributes of the group. The main advantage of working with collineation groups is the convenience of discussion; for example, almost all groups relevant to our study are finite groups in this way. Given any finite collineation group  $\overline{G}$ , there exists a unimodular unitary group  $G$  (a unitary group whose elements have determinant 1) such that their orders satisfy  $|G| \leq d|\overline{G}|$  [37]. This observation is useful when we need to consider the group of unitary operators itself. For the convenience of later discussions, a group of unitary operators is also perceived as a representation of itself.

The *symmetry group*  $G_{\text{sym}}$  of a SIC POM is composed of all unitary operators that leave the SIC POM invariant. The extended symmetry group  $EG_{\text{sym}}$  is the larger group that also contains antiunitary operators. The corresponding collineation groups  $\overline{G}_{\text{sym}}$  and  $\overline{EG}_{\text{sym}}$  are isomorphic to subgroups of the permutation group of  $d^2$  letters and are thus finite groups [279]. A SIC POM is group covariant if it can be generated from a single state—the fiducial state—by a group composed of unitary operations, in which case the states in the SIC POM form a single orbit under the action of its symmetry group. It is *strong group covariant* if, in addition, the generating group has order  $d^2$ . It should be noted that many researchers take strong group covariance as the definition of group covariance [6, 232, 245, 275]. By our definition, strong group covariance implies group covariance; however, it is not known whether the converse also holds, although all SIC POMs known so far are strong group covariant.

### 7.2.1 Groups that can generate SIC POMs

In this section we derive a necessary condition on every group of unitary operators that can generate a SIC POM. Our study starts from inspecting the action on a SIC POM by its symmetry group.

The stepping stone of our analysis is Theorem 2.34 in Zauner’s thesis [275]<sup>3</sup>, as reformulated as follows.

---

<sup>3</sup>We are grateful to Markus Grassl for bringing Zauner’s theorem and Ref. [119] to our attention.

**Theorem 7.1** (Zauner) *Suppose  $\{\Pi_0, \Pi_1, \dots, \Pi_{d^2-1}\}$  is a set of  $d \times d$  linearly independent orthogonal projection matrices and  $U$  is an element in the symmetry group of the set of matrices. Then the number of fixed points  $f(U)$  of  $U$  is equal to the absolute square of the trace of  $U$ ; that is,  $f(U) = |\text{tr}(U)|^2$ .*

**Proof.** Represent  $\Pi_j$  with the column vector  $((\Pi_j)_{00}, (\Pi_j)_{01}, \dots)^T$  and let  $P$  be the matrix formed by juxtaposing the column representations of the  $\Pi_j$ s; then the conjugation by  $U$  can be cast as multiplication by  $U \otimes U^*$ ,

$$(U \otimes U^*)P = PM_U, \quad (7.2)$$

where  $M_U$  is a permutation matrix determined by  $U$ . Since the  $\Pi_j$ s span the matrix space, the matrix  $P$  is invertible and, as a consequence,

$$\text{tr}(M_U) = \text{tr}(U \otimes U^*) = |\text{tr}(U)|^2. \quad (7.3)$$

Now the theorem follows from the observation that the trace of  $M_U$  is equal to  $f(U)$ . □

According to the above proof, the  $\Pi_j$ s are not necessarily restricted to orthogonal projection matrices; they may be any set of positive semidefinite matrices that span the matrix space.

**Lemma 7.2** *Suppose  $G$  is a subgroup of the symmetry group of a SIC POM. Then the number of orbits of states in the SIC POM under the action of  $G$  is equal to the sum of the squared multiplicities of all the inequivalent irreducible components of  $G$ .*

**Proof.** Suppose the states form  $k$  orbits under the action of  $G$ ; let  $f_j(U)$  denote the number of fixed points of  $U \in G$  on the orbit  $j$ . For each orbit,  $|\overline{G}|$  is equal to the order of the stabilizer of each state multiplied by the length of the orbit, which is equal to the sum of the numbers of fixed points of the elements in  $\overline{G}$ ; therefore,

$$|\overline{G}| = \sum_{U \in \overline{G}} f_j(U) \quad \text{for } j = 1, 2, \dots, k. \quad (7.4)$$

## 7.2. Symmetry and group covariance

---

Upon summing over all the orbits and applying Zauner's theorem, we have

$$k|\overline{G}| = \sum_{U \in \overline{G}} f(U) = \sum_{U \in \overline{G}} |\text{tr}(U)|^2. \quad (7.5)$$

According to representation theory [69],  $k$  is equal to the sum of the squared multiplicities of all the inequivalent irreducible components of  $G$ .  $\square$

**Theorem 7.3** *Let  $G$  be a subgroup of the symmetry group of a SIC POM. Then the SIC POM is covariant with respect to  $G$  if and only if  $G$  is irreducible; it is strong covariant with respect to  $G$  if and only if  $G$  is a nice error basis.*

The first part of the theorem is an immediate consequence of Lemma 7.2. Previously, unaware of Zauner's theorem, we proved that  $G$  is necessarily non-Abelian under the same assumption as in the theorem [279]. The second part of the theorem was first shown by Grassl [119] based on Zauner's theorem. It follows from the first part and the properties of nice error bases [166, 168, 169] (see Appendix F). This theorem reveals the importance of nice error bases in the study of SIC POMs.

### 7.2.2 Orbits and equivalence of SIC POMs

Let  $N(G)$  be the normalizer of  $G$ ; then  $U|\psi\rangle$  is a fiducial ket for any  $U \in N(G)$  whenever  $|\psi\rangle$  is. Under the action of  $N(G)$ , all the fiducial states of  $G$  form disjoint orbits, and so do all  $G$ -covariant SIC POMs. SIC POMs on the same orbit are unitarily equivalent. The equivalence relations among SIC POMs on different orbits are determined by the following theorem, assuming we know the symmetry group of the SIC POM.

**Theorem 7.4** *Suppose a SIC POM is covariant with respect to the group  $G$ , and its symmetry group  $G_{\text{sym}}$  contains  $m$  subgroups  $G_1 = G, G_2, \dots, G_m$  that are unitarily equivalent to  $G$ ; let  $U_j$  be a unitary transformation from  $G_j$  to  $G$ ; that is,  $U_j G_j U_j^\dagger = G$ . Then any other  $G$ -covariant SIC POM is unitarily equivalent to the given SIC POM if and only if it is on the same orbit as the image of the given SIC POM under the action of  $U_j$  for some  $j$ . The number of orbits of  $G$ -covariant SIC POMs that are*

*unitarily equivalent to the given SIC POM is equal to the number of conjugacy classes of subgroups of  $G_{\text{sym}}$  that are unitarily equivalent to  $G$ .*

Theorem 7.4 is also applicable when  $G$  and  $G_{\text{sym}}$  are replaced by their collineation groups. In that case, however, both the number of orbits and that of conjugacy classes may change. This theorem is quite useful to investigating the structure of group covariant SIC POMs, especially those in dimension 3, as we shall see later.

**Proof.** According to Theorem 7.3, the given SIC POM is covariant with respect to  $G_j$  for all  $j$ . By assumption, the image of the given SIC POM under the action of  $U_j$ , henceforth denoted by SIC POM  $j$ , is covariant with respect to  $G$ . On the other hand, if a unitary transformation maps the given SIC POM to another  $G$ -covariant SIC POM, then it must map  $G_j$  for some  $j$  to  $G$ . Moreover, the orbit  $\text{orb}(U_j)$  of the image SIC POM does not depend on the choice of  $U_j$ . To see this, let  $U'_j$  be another unitary transformation from  $G_j$  to  $G$ ; then  $U'_j U_j^\dagger \in N(G)$  since it stabilizes  $G$ . So the images of the given SIC POM under the actions of  $U_j$  and  $U'_j$ , respectively, are on the same orbit; note that they can be transformed from one to another by  $U'_j U_j^\dagger$ .

To prove the second part of the theorem, we shall establish a one-to-one correspondence between the conjugacy classes of subgroups of  $G_{\text{sym}}$  that are equivalent to  $G$  and the orbits of  $G$ -covariant SIC POMs that are equivalent to the given SIC POM. More precisely, we shall show that  $\text{orb}(U_j) = \text{orb}(U_k)$  if and only if  $G_j$  and  $G_k$  are conjugated to each other under  $G_{\text{sym}}$ . Suppose  $\text{orb}(U_j) = \text{orb}(U_k)$  and  $V \in N(G)$  transforms SIC POM  $j$  to SIC POM  $k$ ; then  $U_k^\dagger V U_j \in G_{\text{sym}}$  and it maps  $G_j$  to  $G_k$ . On the other hand, suppose  $G_j$  is mapped to  $G_k$  under the conjugation of  $U \in G_{\text{sym}}$ ; then  $G_j$  is mapped to  $G$  under the conjugation of  $U_k U$ , which implies that  $\text{orb}(U_j) = \text{orb}(U_k U) = \text{orb}(U_k)$ . □

### 7.3 Heisenberg–Weyl group and Clifford group

The HW group (also known as the generalized Pauli group) was first introduced by Weyl [266] in the study of quantum kinematics. The HW group and its normalizer—the Clifford group—have played an important role in quantum information science,

### 7.3. Heisenberg–Weyl group and Clifford group

---

such as quantum error correction and quantum computation [113, 114]. During the past decade, they have also found extensive applications in the study of SIC POMs [6, 7, 10, 12, 95, 115, 232, 245, 275, 278, 279] and MUB [7, 83, 155, 272]. It should be noted that there are different versions of the HW group and, accordingly, different versions of the Clifford group [6, 7, 113, 114]. We are mainly concerned with a particular version to be defined shortly that is most relevant to the study of SIC POMs.

In this section we review the basic properties of the HW group and the Clifford group following the work of Appleby [6]. For the convenience of later discussion, in Appendix H we provide some supplementary materials about the Clifford group, such as the trace of a Clifford operator, the normalizer of the Clifford group, and HW groups in the Clifford group.

#### 7.3.1 Heisenberg–Weyl group

The HW group  $D$  is generated by the phase operator  $Z$  and the cyclic-shift operator  $X$  defined by their action on the kets  $|e_r\rangle$  of the “computational basis”,

$$Z|e_r\rangle = \omega^r|e_r\rangle, \quad X|e_r\rangle = |e_{r+1}\rangle, \quad (7.6)$$

where  $r \in \mathbb{Z}_d$ ,  $\omega = e^{2\pi i/d}$  is a primitive  $d$ th root of unity, and  $\mathbb{Z}_d$  is the ring of integers modulo  $d$ . The HW group is a nice error basis [166, 168, 169] (see Appendix F) by definition. The two generators obey the canonical commutation relation

$$XZX^{-1}Z^{-1} = \omega^{-1}, \quad (7.7)$$

which determines the HW group up to unitary equivalence and overall phase factors [266]. All elements of the HW group take on the form

$$D_{k_1, k_2} = \tau^{k_1 k_2} X^{k_1} Z^{k_2}, \quad (7.8)$$

where  $\tau = -e^{\pi i/d}$ , and  $k_1, k_2 \in \mathbb{Z}_d$ ; here the phase factor has been chosen following Appleby [6] to simplify the discussion. Note that  $\tau$  is a primitive  $d$ th root of unity

when  $d$  is odd but a  $2d$ th root of unity otherwise. These elements satisfy the following relations [6]:

$$\begin{aligned}
 D_{\mathbf{k}}^\dagger &= D_{-\mathbf{k}}, \\
 D_{\mathbf{k}}D_{\mathbf{q}} &= \tau^{\langle \mathbf{k}, \mathbf{q} \rangle} D_{\mathbf{k}+\mathbf{q}}, \\
 D_{\mathbf{k}+d\mathbf{q}} &= \begin{cases} D_{\mathbf{k}} & \text{if } d \text{ is odd,} \\ (-1)^{\langle \mathbf{k}, \mathbf{q} \rangle} D_{\mathbf{k}} & \text{if } d \text{ is even,} \end{cases}
 \end{aligned} \tag{7.9}$$

where  $\langle \mathbf{k}, \mathbf{q} \rangle := k_2q_1 - k_1q_2$  is the symplectic form. Note that  $D_{\mathbf{k}+d\mathbf{q}}$  may differ from  $D_{\mathbf{k}}$  by a sign factor when  $d$  is even. Both  $D$  and  $\bar{D}$  are referred to as HW groups, but they possess quite different properties. For example, the former is a non-Abelian group, whereas the latter is an Abelian group that is isomorphic to  $\mathbb{Z}_d^2$ .

According to the definition of a SIC POM [see Eq. (1.1)], a fiducial ket  $|\psi\rangle$  of the HW group obeys

$$|\langle \psi | D_{k_1, k_2} | \psi \rangle| = \frac{1}{\sqrt{d+1}} \tag{7.10}$$

for all  $(k_1, k_2) \neq (0, 0)$ . Up to now, analytical fiducial kets of the HW group have been constructed for  $d \leq 16$  and  $d = 19, 24, 28, 31, 35, 37, 43, 48$  [6, 9, 120, 232, 245, 275]; numerical fiducial kets with high precision have been found up to  $d = 67$  [232, 245]. Moreover, almost all known SIC POMs are covariant with respect to the HW group, partially because the HW group is most familiar to many researchers and is universal in the sense that it exists in every dimension.

When the dimension is a prime power  $p^k$  with  $k \geq 2$ , there is another version of the HW group that is the  $k$ -fold tensor product of the usual HW group in prime dimension  $p$ . This HW group is usually called  $k$ -qubit Pauli group when  $p = 2$ . In dimension 8, the three-qubit Pauli group can generate the set of Hoggar lines [146, 275]. However, no other multiqubit Pauli group can generate any SIC POM according to Godsil and Roy [112]. The situation is still not clear in the case of odd prime-power dimensions.



### 7.3. Heisenberg–Weyl group and Clifford group

---

#### 7.3.2 Special linear group

To understand the structure of the Clifford group, we need to take a detour reviewing the concept of the *special linear group*  $\mathrm{SL}(2, \mathbb{Z}_d)$  [6, 95, 279], which is composed  $2 \times 2$  matrices

$$F = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad (7.11)$$

with entries in  $\mathbb{Z}_d$  and determinant  $1 \pmod{d}$ . Likewise, the extended special linear group  $\mathrm{ESL}(2, \mathbb{Z}_d)$  is the larger group that also contains matrices with determinant  $-1$ . It can be written as a union  $\mathrm{ESL}(2, \mathbb{Z}_d) = \mathrm{SL}(2, \mathbb{Z}_d) \cup \mathrm{JSL}(2, \mathbb{Z}_d)$ , where

$$J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (7.12)$$

The group  $\mathrm{SL}(2, \mathbb{Z}_d) \ltimes (\mathbb{Z}_d)^2$  is the semidirect product defined by the product rule

$$(F_1, \chi_1) \circ (F_2, \chi_2) = (F_1 F_2, \chi_1 + F_1 \chi_2), \quad (7.13)$$

where  $F_1, F_2 \in \mathrm{SL}(2, \mathbb{Z}_d)$  and  $\chi_1, \chi_2 \in (\mathbb{Z}_d)^2$ . The group  $\mathrm{ESL}(2, \mathbb{Z}_d) \ltimes (\mathbb{Z}_d)^2$  is defined in the same manner.

Denote

$$\bar{d} = \begin{cases} d & \text{if } d \text{ is odd,} \\ 2d & \text{if } d \text{ is even.} \end{cases} \quad (7.14)$$

Then the group  $\mathrm{SL}(2, \mathbb{Z}_{\bar{d}}) \ltimes (\mathbb{Z}_d)^2$  can be defined in the same way as  $\mathrm{SL}(2, \mathbb{Z}_d) \ltimes (\mathbb{Z}_d)^2$ , except that, when computing the product  $F_1 \chi_2$  in Eq. (7.13), we implicitly take the following natural homomorphism from  $\mathrm{SL}(2, \mathbb{Z}_{\bar{d}})$  to  $\mathrm{SL}(2, \mathbb{Z}_d)$ :  $F_1 \rightarrow F_1 \pmod{d}$ . The same recipe applies to  $\mathrm{ESL}(2, \mathbb{Z}_{\bar{d}}) \ltimes (\mathbb{Z}_d)^2$ .

#### 7.3.3 Understanding the Clifford group from a homomorphism

The Clifford group  $\overline{\mathrm{C}}(d)$  is the normalizer (within the group of all unitary operations) of the HW group, that is, the group of unitary operations that map the HW group to itself. Likewise, the extended Clifford group  $\overline{\mathrm{EC}}(d)$  is the larger group that also

contains anti-unitary operations [6].

The structure of the extended Clifford group can be characterized succinctly by the following surjective homomorphism given by Appleby [6]:

$$\begin{aligned} f_E : \quad \text{ESL}(2, \mathbb{Z}_{\bar{d}}) \times (\mathbb{Z}_d)^2 &\rightarrow \overline{\text{EC}}(d), \\ UD_{\mathbf{k}}U^\dagger = \omega^{\langle \chi, F\mathbf{k} \rangle} D_{F\mathbf{k}} \quad \text{for} \quad U &= f_E(F, \chi). \end{aligned} \quad (7.15)$$

When  $d$  is odd,  $f_E$  is an isomorphism; that is, the kernel  $K(d)$  is trivial. Otherwise, the kernel  $K(d)$  is composed of the eight elements

$$\left( \left( \begin{array}{cc} 1+rd & sd \\ td & 1+rd \end{array} \right), \left( \begin{array}{c} sd/2 \\ td/2 \end{array} \right) \right) \quad \text{for} \quad r, s, t = 0, 1. \quad (7.16)$$

Consider  $F$  as given in Eq. (7.11) with  $\det(F) = 1 \pmod{\bar{d}}$ . If  $\beta$  is invertible in  $\mathbb{Z}_{\bar{d}}$ , then the homomorphism image of  $(F, \chi)$  reduces to  $D_\chi V_F$  [6], where

$$V_F = \frac{1}{\sqrt{d}} \sum_{r,s=0}^{d-1} |e_r\rangle \tau^{\beta^{-1}(\alpha s^2 - 2rs + \delta r^2)} \langle e_s|. \quad (7.17)$$

If  $\beta$  is not invertible, then  $F$  can be written as a product  $F = F_1 F_2$  with

$$F_1 = \begin{pmatrix} 0 & -1 \\ 1 & x \end{pmatrix}, \quad F_2 = \begin{pmatrix} \gamma + x\alpha & \delta + x\beta \\ -\alpha & -\beta \end{pmatrix}, \quad (7.18)$$

and accordingly,  $V_F = V_{F_1} V_{F_2}$ . One can ensure that  $\delta + x\beta$  be invertible with a suitable choice of  $x$ , in which case  $V_{F_1}$  and  $V_{F_2}$  can be computed by means of Eq. (7.17) [6].

When  $\beta = 0$ , for example,  $x$  can be set to 0, and we have

$$V_F = \sum_{s=0}^{d-1} |e_{\alpha s}\rangle \tau^{\alpha \gamma s^2} \langle e_s|. \quad (7.19)$$

If  $\det(F) = -1$ , then  $\det(FJ) = 1$  and  $(FJ, \chi) \in \text{SL}(2, \mathbb{Z}_{\bar{d}}) \times (\mathbb{Z}_d)^2$ . So the images of the elements in  $\text{ESL}(2, \mathbb{Z}_{\bar{d}}) \times (\mathbb{Z}_d)^2$  can be determined once the images of the elements in  $\text{SL}(2, \mathbb{Z}_{\bar{d}}) \times (\mathbb{Z}_d)^2$  and that of  $(J, \mathbf{0})$  are known, where  $\mathbf{0}$  is a shorthand for  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . The

### 7.3. Heisenberg–Weyl group and Clifford group

---

image of  $(J, \mathbf{0})$  is the complex-conjugation operator [6]

$$\hat{J} : \sum_{r=0}^{d-1} |e_r\rangle a_r \mapsto \sum_{r=0}^{d-1} |e_r\rangle a_r^*, \quad (7.20)$$

which is clearly basis dependent (here defined with respect to the computational basis) and has no physical meaning.

Following Appleby, we denote by  $[F, \chi]$  (This is not a commutator!) the homomorphism image of  $(F, \chi)$  in the rest of the thesis. For the convenience of later discussions,  $[F, \chi]$  is often represented by a specific element in the equivalent class; for example,  $[F, \mathbf{0}]$  by  $V_F$ . By convention, elements in the HW group are called *displacement operations*, whereas elements of the form  $[F, \mathbf{0}]$  in the Clifford group are called *symplectic operations*.

The following two elements in the special linear group deserve special attention,

$$F_Z = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad F_L = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \quad (7.21)$$

Their corresponding Clifford operators are given by

$$V_Z = \sum_{r,s=0}^{d-1} \tau^{2rs+r^2} |e_r\rangle \langle e_s|, \quad V_L = \sum_{s=0}^{d-1} \tau^{s^2} |e_s\rangle \langle e_s|. \quad (7.22)$$

The matrix  $F_Z$  and the operator  $V_Z$  are known as Zauner matrix and Zauner unitary transformation [6, 275].



# SIC POMs in prime dimensions

---

## 8.1 Introduction

Almost all known SIC POMs are covariant with respect to the HW group [6, 232, 245, 275]. Is it due to a deep reason or simply because such SIC POMs are easy to construct? In either case, a better understanding of HW covariant SIC POMs is crucial to unravel the mystery. An important tool for this endeavor is the Clifford group, which divides HW covariant SIC POMs into disjoint orbits, such that SIC POMs on the same orbit are unitarily equivalent. The relations among SIC POMs on different orbits have remained elusive despite the efforts of many researchers in the past decade [6, 7, 95, 115, 245, 279]. The main difficulty lies in determining the relation between the symmetry group of an HW covariant SIC POM and the Clifford group, as manifested by the long-standing open question: Is the symmetry group necessarily a subgroup of the Clifford group? Although an affirmative answer has been speculated for a long time, no rigorous proof is known. Actually, this speculation has caused some confusions about the structure of SIC POMs, especially those in dimension 3, which exhibit a plethora of peculiar properties, including the existence of a continuous family of orbits [6, 232, 275, 279].

In this chapter<sup>1</sup> we settle the open problems mentioned above for any prime dimension  $p$  and clarify several subtle points about SIC POMs in dimension 3. The situation beyond prime dimensions is also discussed briefly. All unitary groups considered here are assumed to be unimodular for convenience. So the order of any unitary group containing the HW group is  $p$  times the order of the corresponding collineation group.

---

<sup>1</sup>This chapter is based on the following paper: H. Zhu, *SIC POVMs and Clifford groups in prime dimensions*, J. Phys. A: Math. Theor. **43**, 305305 (IOP Publishing, 2010).

## 8.2 Group covariant SIC POMs are HW covariant

**Theorem 8.1** *Any group covariant SIC POM in a prime dimension is covariant with respect to the HW group.*

In any prime dimension, a group covariant SIC POM is necessarily covariant with respect to each Sylow  $p$ -subgroup, say  $P$ , of its symmetry group  $G_{\text{sym}}$  (see Ref. [172] or Appendix G for a brief introduction of Sylow  $p$ -subgroups). The center of  $P$  has order at least  $p$  since any  $p$ -group has a nontrivial center. Meanwhile,  $P$  must be irreducible according to Theorem 7.3. So any element in its center is proportional to the identity, which, together with the unimodular condition, implies that the center is the cyclic group  $\langle \omega \rangle$  generated by the constant  $\omega$ . Since the  $p$ -group  $P/\langle \omega \rangle$  also has a nontrivial center, there exists an element  $X' \in P$  such that  $X'$  does not belong to the center of  $P$  but  $X'\langle \omega \rangle$  belongs to the center of  $P/\langle \omega \rangle$ . Hence, there exists another element  $Z' \in P$  such that  $Z'X'Z'^{-1}X'^{-1} = \omega^k$  with  $1 \leq k < p$ . The integer  $k$  may be set to 1 by replacing  $Z'$  with its suitable power if necessary. Now the group generated by  $X'$  and  $Z'$  must be an HW group according to their commutation relation [266] (see Sec. 7.3.1), and the theorem follows.

## 8.3 Qubit SIC POMs

In dimension 2, the Clifford group is generated by the Hadamard operator  $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$  and the phase operator  $\begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$  [206]; the extended Clifford group is generated by the complex-conjugation operator  $\hat{J}$  besides the two operators. The orders of the two groups are 24 and 48, respectively. There is only one orbit of eight fiducial states, one of which reads [6, 232, 275]

$$|\psi\rangle \hat{=} \begin{pmatrix} \sqrt{\frac{3+\sqrt{3}}{6}} \\ e^{i\pi/4} \sqrt{\frac{3-\sqrt{3}}{6}} \end{pmatrix}. \quad (8.1)$$

The orders of the stabilizers of each fiducial state within the Clifford group and the extended Clifford group are 3 and 6, respectively.

#### 8.4. SIC POMs in prime dimensions not equal to 3

---

When represented on the Bloch sphere, the eight fiducial states constitute the corners of a cube, and the two SIC POMs of two regular tetrahedrons, which are related to each other by inversion. The Clifford group corresponds to the (rotational) symmetry group of the cube. The symmetry group of each SIC POM corresponds to the symmetry group of the tetrahedron, which is a subgroup of the symmetry group of the cube. It is thus a subgroup of the Clifford group and contains only one HW group. By the same token, the extended symmetry group is a subgroup of the extended Clifford group. Incidentally, all SIC POMs in dimension 2 are covariant with respect to the HW group and are unitarily equivalent to each other.

#### 8.4 SIC POMs in prime dimensions not equal to 3

In this section, we prove that, in any prime dimension not equal to 3, any HW covariant SIC POM is covariant with respect to a unique HW group, and its (extended) symmetry group is a subgroup of the (extended) Clifford group. As a consequence, two HW covariant SIC POMs are equivalent if and only if they are on the same orbit. To achieve this goal, we first prove that the HW group is a Sylow  $p$ -subgroup of the symmetry group and then prove that it is a normal subgroup of the symmetry group. Since the case  $p = 2$  has been handled in Sec. 8.3, it remains to consider the case  $p \geq 5$ .

**Lemma 8.2** *In any prime dimension  $p \neq 3$ , the HW group is a Sylow  $p$ -subgroup of the symmetry group of any HW covariant SIC POM.*

Suppose the HW group is not a Sylow  $p$ -subgroup of the symmetry group; then it is a proper subgroup of one of the Sylow  $p$ -subgroups. The normalizer of the HW group within this Sylow  $p$ -subgroup is strictly larger than the HW group, so that each fiducial state is stabilized by an order- $p$  Clifford operation  $[F, \chi]$ , which is not traceless according to Theorem 7.1. Thanks to Theorem H.1, we may assume that  $\chi = \mathbf{0}$  without loss of generality. Therefore,  $F$  is an order- $p$  element in  $\text{SL}(2, \mathbb{Z}_p)$  and is thus conjugated to either  $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$  or  $\begin{pmatrix} 1 & 0 \\ \nu & 1 \end{pmatrix}$ , where  $\nu$  is a primitive element in the Galois field  $\mathbb{Z}_p$  [154, 279]. In either case,  $|\text{tr}([F, \chi])|^2 = p$  and each eigenspace of  $[F, \chi]$  has dimension at most

two. According to Theorem 7.1,  $[F, \chi]$  stabilizes  $p$  fiducial states simultaneously, which necessarily belong to a same eigenspace. However, a two-dimensional eigenspace cannot accommodate more than two fiducial states when  $p \geq 5$  according to the Welch bound [see Eq. (7.1)]. This contradiction completes the proof of the lemma.

To prove that the HW group is a normal subgroup of the symmetry group, we shall distinguish two cases depending on whether the symmetry group admits a monomial representation. One of the cases can be handled based on a theorem of Sibley [251]:

**Theorem 8.3** (Sibley) *Suppose  $G$  is a finite group with a faithful, irreducible, unimodular, and quasiprimitive representation of prime degree  $p \geq 5$ . If a Sylow  $p$ -subgroup  $P$  of  $G$  has order  $p^3$ , then  $P$  is normal in  $G$ , and  $G/P$  is isomorphic to a subgroup of  $\text{SL}(2, \mathbb{Z}_p)$ .*

A quasiprimitive representation is one whose restriction to every subgroup is homogeneous, that is, a multiple of one irreducible representation of the subgroup. An irreducible representation of prime degree that is not quasiprimitive is monomial [251].

If  $G_{\text{sym}}$  is quasiprimitive, then the HW group  $D$  is a normal subgroup of  $G_{\text{sym}}$  according to Sibley's theorem; in other words,  $G_{\text{sym}}$  is a subgroup of the Clifford group. In addition,  $G_{\text{sym}}$  contains only one HW group according to Sylow's theorem [172] since the HW group is also a Sylow  $p$ -subgroup of  $G_{\text{sym}}$ .

Now suppose  $G_{\text{sym}}$  is in monomial form; that is, all elements in  $G_{\text{sym}}$  are monomial after a suitable choice of basis if necessary. Let  $T$  be the normal subgroup of  $G_{\text{sym}}$  that is composed of its diagonal elements; then  $T$  contains the subgroup generated by  $\omega$  and a nontrivial displacement operator, say  $Z$ . It turns out that  $T$  contains no other elements. To see this, note that the order of  $T$  is not divisible by  $p^3$  since, otherwise, it would contain an HW group. In addition,  $T$  is a direct sum of  $p$  inequivalent one-dimensional representations when taken as a representation of itself. According to Lemma 7.2, the fiducial states of the SIC POM form  $p$  orbits of equal length  $p$  under the action of  $T$ ; in other words, two fiducial states are on the same orbit generated by  $T$  if and only if they are on the same orbit generated by  $Z$ . Suppose  $|T| > p^2$ ; then  $T$  contains an element  $U$  whose order is not divisible by  $p$ , and there exists at



### 8.5. SIC POMs in dimension 3

---

least one fiducial state, say  $|\psi\rangle$ , that is not stabilized by  $U$ . Since  $U|\psi\rangle$  and  $|\psi\rangle$  are on the same orbit of  $Z$ , there exist an integer  $1 \leq k \leq p-1$  and a phase factor  $e^{i\phi}$  such that  $U|\psi\rangle = e^{i\phi} Z^k |\psi\rangle$ . Note that  $|\psi\rangle$  has at least two nonzero entries. Let  $e^{i\phi_1}$  and  $e^{i\phi_2}$  be the two diagonal entries of  $U$  corresponding to any two nonzero entries of  $|\psi\rangle$ , respectively; then  $e^{i(\phi_1-\phi_2)}$  is a primitive  $p$ th root of unity, contradicting the fact that the order of  $U$  is not a multiple of  $p$ . This contradiction confirms that  $T$  is generated by  $\omega$  and  $Z$ . As a consequence,  $\overline{D}$  is the centralizer of  $\overline{T}$  within  $\overline{G}_{\text{sym}}$  and is thus a normal subgroup of  $\overline{G}_{\text{sym}}$ . In other words,  $\overline{G}_{\text{sym}}$  is a subgroup of the Clifford group and contains only one HW group.

In summary, for any prime dimension larger than 3, each group covariant SIC POM is covariant with respect to a unique HW group, and its symmetry group is a subgroup of the Clifford group. The conclusion can also be extended to cover antiunitary operations since any antiunitary operation in the extended symmetry group of the SIC POM must stabilize the HW group and thus belong to the extended Clifford group. In addition, the same conclusion holds for dimension 2 according to Sec. 8.3.

**Theorem 8.4** *In any prime dimension not equal to 3, each group covariant SIC POM is covariant with respect to a unique HW group. Furthermore, its (extended) symmetry group is a subgroup of the (extended) Clifford group and contains the HW group as a normal Sylow  $p$ -subgroup.*

In conjunction with Theorem 7.4, we can settle the equivalence problem of HW covariant SIC POMs in prime dimensions.

**Corollary 8.5** *In any prime dimension not equal to 3, two SIC POMs covariant with respect to the same HW group are unitarily (unitarily or antiunitarily) equivalent if and only if they are on the same orbit of the Clifford group (extended Clifford group).*

### 8.5 SIC POMs in dimension 3

The existence of a SIC POM in dimension 3 was already noted in the 1970s by Del-sarte, Goethals, and Seidel [76]. More than a decade ago, Zauner [275] constructed

a continuous family of SIC POMs by virtue of the HW group and determined their symmetry groups without reference to the Clifford group. However, he did not specify his approach, and there were some mistakes in his conclusions about the symmetry groups. Later, these SIC POMs attracted the attention of many other researchers [6, 64, 162, 195, 232, 240, 279], but their peculiarity is still a source of confusion to the SIC community.

In this section, we shall clarify several subtle points concerning the symmetry and the equivalence problems about group covariant SIC POMs in dimension 3. In sharp contrast with group covariant SIC POMs in other prime dimensions, each one in dimension 3 is covariant with respect to either three or nine HW groups. Its symmetry group is a subgroup of at least one of the Clifford groups of these HW groups, respectively, but not necessarily a subgroup of the Clifford of the HW group defined in the standard basis. As a consequence, SIC POMs on different orbits of the Clifford group can be equivalent even if their respective stabilizers have different orders, assuming that we considers symmetry operations only within the Clifford group, as is the case in almost all studies in the past decade because of the technical difficulty in determining the full symmetry groups. For the convenience of the following discussion, all HW groups in the Clifford group for any prime dimension are figured out in Appendix H.3.

### 8.5.1 Symmetry of SIC POMs

If  $G_{\text{sym}}$  is not monomial, then the order of each Sylow  $p$ -subgroup of  $G_{\text{sym}}$  is at most  $p^4$  according to the classification of finite linear groups of degree 3 by Blichfeldt [37]. Moreover,  $G_{\text{sym}}$  is isomorphic to a subgroup of the Clifford group if the order is  $p^4$ . Otherwise, there is a counterexample to Sibley's theorem (Theorem 8.3)—a unimodular unitary group of order 1080 whose collineation group (of order 360) is isomorphic to the alternating group of six letters [37]. However, this group cannot be the symmetry group of any SIC POM. To demonstrate this point, suppose otherwise. Let  $U$  be an order-5 element in the group; then the nine fiducial states of the SIC POM form disjoint orbits of length either 1 or 5 under the action of  $U$ . It follows that there are

### 8.5. SIC POMs in dimension 3

---

four orbits of length 1, that is, four fiducial states stabilized by  $U$ . The four fiducial states must belong to a same eigenspace of  $U$ , whose dimension is at most two since  $U$  is not proportional to the identity. However, a two-dimensional subspace cannot admit four fiducial states [see Eq. (7.1)]. This contradiction dictates that  $G_{\text{sym}}$  must be a subgroup of some Clifford group when it is not monomial.

Now suppose that  $G_{\text{sym}}$  is in monomial form and that one of the HW groups contained in  $G_{\text{sym}}$  is in the standard form, as in the case  $p \geq 5$ . Let  $T$  be the normal subgroup of  $G_{\text{sym}}$  consisting of its diagonal elements.

If  $T = \langle \omega, Z \rangle$ , then  $G_{\text{sym}}$  contains only one HW group and it is a subgroup of the Clifford group, according to similar reasoning in Sec. 8.4. However, it turns out that this scenario does not occur for the special case  $p = 3$  [6] (see also Sec. 8.5.2), in sharp contrast with the general case  $p \geq 5$ .

Otherwise, each fiducial state, say  $|\psi\rangle$ , is stabilized by some element in  $T$  that is not proportional to the identity, say  $U$ . Simple analysis shows that  $U$  has two identical diagonal entries and that  $|\psi\rangle$  has two nonzero entries with equal modulus  $\frac{1}{\sqrt{2}}$  and a zero entry. Without loss of generality, we may assume that  $U \cong e^{i\phi'} \text{diag}(1, 1, e^{i\phi})$  and  $|\psi\rangle \cong \frac{1}{\sqrt{2}}(1, e^{it}, 0)$ ; indeed, all kets of this form are fiducial kets [6] (see also Sec. 8.5.2). To ensure that  $UX|\psi\rangle$  be a fiducial ket in the SIC POM,  $\phi$  may take on only two possible values  $\pm \frac{2\pi}{3}$ . We may choose  $U = e^{-i2\pi/9} \text{diag}(1, 1, \omega)$  for definiteness, where  $\phi'$  has been chosen such that  $U$  is unimodular. Now  $T$  cannot contain any element other than those generated by  $\omega$ ,  $Z$ , and  $U$ , and therefore  $|T| = 27$ .

Define  $R = G_{\text{sym}}/T$ ; then  $R$  is isomorphic to a subgroup of the symmetry group of three letters that contains an order-3 cycle and thus has order either 3 or 6. Accordingly, the order of  $G_{\text{sym}}$  is either 81 or 162. If  $|R| = 3$ , then  $G_{\text{sym}}$  is a  $p$ -group of order  $3^4$ , and  $D$  is a normal subgroup of  $G_{\text{sym}}$  since the normalizer of a proper subgroup of a  $p$ -group is strictly larger than the subgroup [172] (see Appendix G). If  $|R| = 6$ , then  $G_{\text{sym}}$  contains a Sylow  $p$ -subgroup  $P$  of order 81 and with index 2, such that  $D$  is a normal subgroup of  $P$ . The group  $P$  is also a Sylow  $p$ -subgroup of the Clifford group and thus contains two other normal subgroups that are also HW groups but in different

bases (see Appendix H.3). One of the three HW groups is a normal subgroup of  $G_{\text{sym}}$ , whereas the other two are conjugated to each other. Therefore,  $G_{\text{sym}}$  is also a subgroup of some Clifford group when  $G_{\text{sym}}$  is monomial.

In summary, the symmetry group of any group covariant SIC POM in dimension 3 is a subgroup of some Clifford group. It should be emphasized that a SIC POM can be covariant with respect to more than one HW group. Its symmetry group can be a subgroup of the Clifford group of one of the HW groups but not of other Clifford groups, say, the standard Clifford group, the Clifford group of the HW group in the standard basis.

Suppose we have chosen the HW group  $\overline{D}$  such that it is a normal subgroup of  $\overline{G}_{\text{sym}}$ . Let  $\overline{D}'$  be any other HW group in  $\overline{G}_{\text{sym}}$  and  $\overline{C}'(3)$  its Clifford group. Then  $\overline{C}'(3) \cap C(3)$  is a Sylow 3-subgroup of  $C(3)$ , which contains three HW groups including  $\overline{D}$  (see Appendix H.3). In other words, the Clifford group of any HW group in  $\overline{G}_{\text{sym}}$  contains at least one HW group whose Clifford group contains  $\overline{G}_{\text{sym}}$ . This observation is very useful for determining the symmetry group of a group covariant SIC POM in dimension 3, no matter whether it is a subgroup of the standard Clifford group.

Since the stabilizer  $\overline{S}$  of each fiducial state is isomorphic to  $\overline{G}_{\text{sym}}/\overline{D}$ , which may be identified as a subgroup of  $\text{SL}(2, \mathbb{Z}_3)$ , its order may assume at most six possible values: 2, 3, 4, 6, 8, and 24 (see Ref. [187]). Analysis shows that, if a fiducial state is stabilized by an order-2 Clifford operation, then it is automatically stabilized by an order-3 Clifford operation. So the order of  $\overline{S}$  may take on only the three values 3, 6, or 24; accordingly, the order of  $\overline{G}_{\text{sym}}$  may take on only the three values 27, 54, or 216. When  $|\overline{S}| = 3$ , the stabilizer  $\overline{S}$  is conjugated to the group generated by the Zauner operation  $[F_Z, \mathbf{0}]$ , and  $\overline{G}_{\text{sym}}$  contains three HW groups, all of which are normal. When  $|\overline{S}| = 6$ , the stabilizer is conjugated to the subgroup generated by  $[-F_Z, \mathbf{0}]$ , and  $\overline{G}_{\text{sym}}$  also contains three HW groups, but only one of which is normal. When  $|\overline{S}| = 24$ , the stabilizer is conjugated to the group composed of all symplectic operations, and  $\overline{G}_{\text{sym}}$  coincides with the Clifford group and contains nine HW groups, only one of which is normal, whereas the other eight are conjugated to each other. In either of the latter

### 8.5. SIC POMs in dimension 3

---

two cases, starting from different HW groups, we can “see” different symmetry groups if considering symmetry operations only within the Clifford group of the given HW group.

It is straightforward to extend the above analysis to show that the extended symmetry group  $\overline{EG}_{\text{sym}}$  of a group covariant SIC POM in dimension 3 is a subgroup of some extended Clifford group. In the case  $|\overline{G}_{\text{sym}}| = 27$ , at least one of the three HW groups in  $\overline{G}_{\text{sym}}$  is stabilized by  $\overline{EG}_{\text{sym}}$ . In the case  $|\overline{G}_{\text{sym}}| = 54$  or 216, the HW group that is stabilized by  $\overline{G}_{\text{sym}}$  is also stabilized by  $\overline{EG}_{\text{sym}}$ .

Any HW covariant SIC POM in dimension 3 possesses antiunitary symmetry. To see this, we may assume without loss of generality that one fiducial state of the SIC POM is stabilized by  $[F_L, \mathbf{0}]$ , which is conjugated to  $[F_Z, \mathbf{0}]$  [see Eq. (7.21)]. Then each fiducial state of the SIC POM has the form  $\frac{1}{\sqrt{2}}(1, e^{it}, 0)^T$  up to permutations of the three entries. Now our claim follows from the observation that the extended symmetry group of any SIC POM in this family contains antiunitary operations. As a consequence, the orbits of the Clifford group coincide with that of the extended Clifford group [6] (see Sec. 8.5.2).

**Theorem 8.6** *In dimension 3, the symmetry group of each HW covariant SIC POM can have three possible orders 27, 54, and 216. Accordingly, it contains three, three, and nine HW groups, respectively, all of which are normal subgroups in the first case, whereas only one is normal in the latter two cases. In all three cases, its intersection with the standard Clifford group contains at least three HW groups. The (extended) symmetry group is a subgroup of at least one of the (extended) Clifford groups associated with these HW groups, respectively.*

As a consequence of Theorems 7.4 and 8.6, SIC POMs on different orbits of the HW group can be equivalent, in sharp contrast with the situation in other prime dimensions.

**Corollary 8.7** *In dimension 3, for each HW covariant SIC POM, there exist three orbits of equivalent SIC POMs if its symmetry group has order 27, and two if the symmetry group has order 54 or 216. In either case, the orbits of equivalent SIC POMs are*

connected to each other by unitary transformations that map additional HW groups in the intersection of the symmetry group and the standard Clifford group to the standard HW group.

### 8.5.2 Infinitely many inequivalent SIC POMs

There is a one-parameter family of fiducial kets in dimension 3,

$$|\psi_f(t)\rangle \hat{=} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -e^{it} \end{pmatrix}. \quad (8.2)$$

For each distinct orbit, there is a unique value of  $t \in [0, \frac{\pi}{3}]$  such that  $|\psi_f(t)\rangle$  is on the orbit. There are three kinds of orbits: two exceptional orbits corresponding to the end points  $t = 0$  and  $t = \frac{\pi}{3}$ , and infinitely many generic orbits corresponding to  $0 < t < \frac{\pi}{3}$  [6, 232, 275].

According to Appleby [6], the order of the stabilizer within the Clifford group (extended Clifford group) of each fiducial state is 24, 6, 3 (48, 12, 6) for the three kinds of orbits, respectively, and the number of SIC POMs on each orbit is 1, 4, 8. The stabilizer (within the extended Clifford group) of the exceptional fiducial ket  $|\psi_f(0)\rangle$  consists of all operations of the form  $[F, \mathbf{0}]$  with  $F \in \text{ESL}(2, \mathbb{Z}_3)$ . The stabilizer of the exceptional fiducial ket  $|\psi_f(\frac{\pi}{3})\rangle$  is generated by the unitary operation

$$[F, \chi] = \left[ \begin{pmatrix} -1 & 0 \\ -1 & -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] \quad (8.3)$$

and the antiunitary operation

$$[A, \chi] = \left[ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]. \quad (8.4)$$

For a generic fiducial ket  $|\psi_f(t)\rangle$  with  $0 < t < \frac{\pi}{3}$ , the stabilizer is generated by the

## 8.5. SIC POMs in dimension 3

---

following two operations,

$$[F, \chi]^2 = \left[ \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right], \quad [F, \chi] \circ [A, \chi] = \left[ \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right]. \quad (8.5)$$

Note that it is independent of  $t$ .

### 8.5.2.1 Equivalence relations among SIC POMs on different orbits

By virtue of Theorem 8.6 and Corollary 8.7, we can now establish complete equivalence relations among group covariant SIC POMs in dimension 3. Note that the symmetry group of the SIC POM generated from the fiducial ket  $|\psi_f(t)\rangle$  contains as a subgroup the Sylow 3-subgroup of the Clifford group that is generated by  $X$  and  $V_L$  [see Eq. (7.22)]. This Sylow-3 subgroup contains three HW groups, which are generated by  $Z, X$ ;  $Z, V_L X$ ; and  $Z, V_L^2 X$ , respectively, and are conjugated to each other under the unitary transformation  $U \cong \text{diag}(1, e^{-2i\pi/9}, e^{-4i\pi/9})$  (see Appendix H.3). According to Corollary 8.7, the SIC POMs on the three orbits generated from  $|\psi_f(t)\rangle$ ,  $U^\dagger|\psi_f(t)\rangle$ , and  $U^{\dagger 2}|\psi_f(t)\rangle$ , respectively, are unitarily equivalent. That is, the SIC POMs on the three orbits corresponding to  $t$ ,  $\frac{2\pi}{9} + t$ , and  $\frac{2\pi}{9} - t$  for each  $0 \leq t \leq \frac{\pi}{9}$  (two of the three orbits can merge when  $t = 0$  or  $t = \frac{\pi}{9}$ ) are unitarily equivalent. Moreover, SIC POMs on any two different orbits are not equivalent when restricted to the orbits with  $0 \leq t \leq \frac{\pi}{9}$ . There are two orbits of equivalent SIC POMs for each exceptional orbit, but three for each generic orbit with  $t \neq \frac{\pi}{9}, \frac{2\pi}{9}$ .

The equivalence between the SIC POM on the exceptional orbit with  $t = 0$  and those on the generic orbit with  $t = \frac{2\pi}{9}$  is particularly surprising at first glance since their stabilizers within the (extended) Clifford group have different orders. Equally surprising is the equivalence between the SIC POMs on the exceptional orbit with  $t = \frac{\pi}{3}$  and those on the generic orbit with  $t = \frac{\pi}{9}$ .

Although the SIC POMs on the three orbits with  $t$ ,  $\frac{2\pi}{9} - t$ , and  $\frac{2\pi}{9} + t$  are equivalent, the orbits themselves are not equivalent in the sense that there is no unitary or antiunitary transformation that can map all SIC POMs on one of the three orbits to

that on another one. For example, under the transformation induced by  $U^\dagger$ , only six out of the 24 SIC POMs on the three orbits are permuted among each other, whereas the other 18 are no longer on the three orbits. To see this point more clearly, we can look into the additional SIC POMs constructed by regrouping the fiducial states on each orbit; see Ref. [279] for more details.

Above analysis also implies that the (extended) symmetry group of each SIC POM is a subgroup of the standard (extended) Clifford group except for those on the orbit with  $t = \frac{\pi}{9}$  or  $t = \frac{2\pi}{9}$ . For each SIC POM on the two special orbits, its (extended) symmetry group is a subgroup of the (extended) Clifford group associated with another HW group that is contained in the intersection of the symmetry group and the standard Clifford group. Incidentally, according to Theorem 8.4, the two orbits are the only cases in prime dimensions in which the (extended) symmetry group of each SIC POM is not a subgroup of the standard (extended) Clifford group.

### 8.5.2.2 Classification of SIC POMs based on geometric phases

To better characterize those inequivalent SIC POMs, we need to find some invariants that can distinguish them. The simplest invariant is the triple product  $\text{tr}(\rho_1\rho_2\rho_3)$  associated with three different fiducial states in a SIC POM. According to the definition of a SIC POM,  $|\text{tr}(\rho_1\rho_2\rho_3)| = \frac{1}{8}$  for  $d = 3$ , so the relevant invariant is the phase of the triple product, which takes on a value between  $-\pi$  and  $\pi$ . Since the odd permutations or the complex conjugation of the three states reverses the sign of the phase, it is advisable to focus on the absolute value of the phase  $\phi = |\arg[\text{tr}(\rho_1\rho_2\rho_3)]|$ , which is independent of the permutations and the complex conjugation. The phase  $\phi$  is known as the discrete geometric phase [1, 35] or as the Bargmann invariant [27] associated with the three states  $\rho_1$ ,  $\rho_2$ , and  $\rho_3$ .

Given a SIC POM generated from the fiducial state in Eq. (8.2), thanks to the group covariance, we may assume that  $\rho_1 = |\psi_f(t)\rangle\langle\psi_f(t)|$  without loss of generality. There are  $\binom{8}{2} = 28$  different choices for the remaining two fiducial states. Analysis of the symmetry group of the SIC POM reveals that  $\phi$  may assume at most five distinct



### 8.5. SIC POMs in dimension 3

Table 8.1: Geometric phases  $\phi = |\arg[\text{tr}(\rho_1\rho_2\rho_3)]|$  associated with five different triple products among states in the SIC POM generated from the fiducial state in Eq. (8.2) for  $t \in [0, \frac{\pi}{9}]$ . To simplify the notation, the fiducial states are represented by displacement operators; for example, the fiducial state  $Z|\psi_f(t)\rangle\langle\psi_f(t)|Z^\dagger$  is represented by  $Z$ . Thanks to group covariance, one fiducial state may be chosen a priori, and the remaining two may take 28 different choices. The second column shows the numbers of choices that give rise to the specific geometric phases presented in the third column.

$\rho_1, \rho_2, \rho_3$	multiplicity	geometric phase
$1, Z, Z^2$	1	$\phi_1 = \pi$
$1, X, Z$	18	$\phi_2 = \frac{\pi}{3}$
$1, X, X^2$	3	$\phi_3 = \pi - 3t$
$1, X, X^2Z$	3	$\phi_4 = \frac{\pi}{3} - 3t$
$1, X, X^2Z^2$	3	$\phi_5 = \frac{\pi}{3} + 3t$

values. Table 8.1 lists the five distinct geometric phases associated with five triple products for  $0 \leq t \leq \frac{\pi}{9}$ . Figure 8.1 shows the variations of the five phases with  $t$  in a wider range. The two phases  $\phi_1$  and  $\phi_2$  are independent of the parameter  $t$ , whereas the other three phases  $\phi_3, \phi_4$ , and  $\phi_5$  are periodic functions of  $t$  with the same shape and period  $\frac{2\pi}{3}$ , but shifted from each other by  $\pm\frac{2\pi}{9}$ . If the three phases  $\phi_3, \phi_4$ , and  $\phi_5$  are not distinguished, then the pattern displays a period of  $\frac{2\pi}{9}$ , with additional mirror symmetry with respect to  $t = \frac{k\pi}{9}$  for  $k = 0, \pm 1, \pm 2, \dots$ . Figure 8.1 demonstrates that two SIC POMs on any two different orbits cannot be equivalent if the corresponding values of  $t$  belong to an open interval of length  $\frac{\pi}{9}$ . By contrast, the equivalence among SIC POMs on the three orbits corresponding to  $t, \frac{2\pi}{9} - t$ , and  $\frac{2\pi}{9} + t$  is underpinned.

Let  $\phi_{\min}$  be the minimum of the five phases listed in Table 8.1. Then  $0 \leq \phi_{\min} \leq \frac{\pi}{3}$ , and there is a one-to-one correspondence between  $\phi_{\min}$  and  $t$  when  $0 \leq t \leq \frac{\pi}{9}$ :

$$\phi_{\min} = \frac{\pi}{3} - 3t. \quad (8.6)$$

Therefore,  $\phi_{\min}$  uniquely specifies the equivalence class of a group covariant SIC POM in dimension 3. Unlike the parameter  $t$ , the phases  $\phi_j$ s and  $\phi_{\min}$  are intrinsic quantities of the SIC POM, which are independent of the parametrization. They are useful even if the SIC POM is not constructed from a fiducial state or the information about the

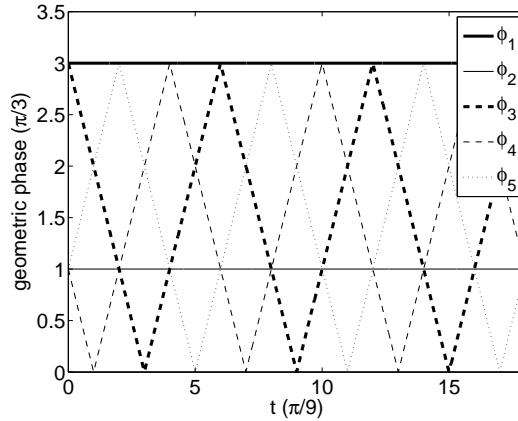


Figure 8.1: Geometric phases associated with five different triple products among states in the SIC POM generated from the fiducial state in Eq. (8.2) for  $t \in [0, 2\pi]$ . See the main text and Table 8.1 for the definitions of the five phases.

symmetry group is missing.

## 8.6 Beyond prime dimensions

In this section we briefly discuss the situation beyond prime dimensions followed by several conjectures; more details will be presented elsewhere [277].

**Theorem 8.8** *Suppose  $\overline{G}_{\text{sym}}$  is the symmetry group of an HW covariant SIC POM in any dimension not equal to 3, and  $\overline{G}$  is its intersection with the Clifford group. Then  $\overline{G}$  contains only one nice error basis with an Abelian index group. If the dimension is a prime power  $p^n$ , then each Sylow  $p$ -subgroup of  $\overline{G}$  is also a Sylow  $p$ -subgroup of  $\overline{G}_{\text{sym}}$ , and all HW groups in  $\overline{G}_{\text{sym}}$  are conjugated to each other. If  $p \geq 5$  or  $p^n = 2, 4$ , then  $\overline{G}$  contains only one nice error basis, and all nice error bases in  $\overline{G}_{\text{sym}}$  are Sylow  $p$ -subgroups and are thus conjugated to each other.*

According to Theorem 8.8, in any prime-power dimension, an HW covariant SIC POM cannot be covariant with respect to the tensor-product version of the HW group, and vice versa. In particular, the set of Hoggar lines [146, 275] is not covariant with respect to the usual HW group. The latter conclusion has been speculated for a long time, but we are not aware of any rigorous proof in the literature. As far as we

## 8.7. Summary

---

know, the set of Hoggar lines is the only SIC POM that is known to be not covariant with respect to the usual HW group (see Chapter 10 for more details).

In conjunction with Theorem 7.4, we can now settle the equivalence problem of HW covariant SIC POMs in any prime-power dimension.

**Corollary 8.9** *In any prime-power dimension not equal to 3, two HW covariant SIC POMs are equivalent if and only if they are on the same orbit.*

A few conjectures deserve further study.

**Conjecture 8.10** *In any dimension not equal to 3, the symmetry group of any HW covariant SIC POM is a subgroup of the Clifford group.*

**Conjecture 8.11** *In any dimension not equal to 3, any HW covariant SIC POM is covariant with respect to only one HW group.*

**Conjecture 8.12** *In any dimension not equal to 3, any HW covariant SIC POM is covariant with respect to only one nice error basis with an Abelian index group.*

According to Theorem 8.8, the above three conjectures are actually equivalent. If any one of them holds, then Corollary 8.9 may be generalized to any dimension not equal to 3.

## 8.7 Summary

The equivalence relations among SIC POMs on different orbits of the (extended) Clifford group have been an elusive question in the SIC community. So has been the closely related question: Is the (extended) symmetry group of an HW covariant SIC POM a subgroup of the (extended) Clifford group? In this chapter we clarified these open questions for all prime dimensions. More specifically, we proved that, in any prime dimension not equal to 3, each group covariant SIC POM is covariant with respect to a unique HW group; its (extended) symmetry group is a subgroup of the (extended) Clifford group. Therefore, two HW covariant SIC POMs are equivalent if and only if

they are on the same orbit. In dimension 3, each group covariant SIC POM can be covariant with respect to three or nine HW groups; its symmetry group is a subgroup of at least one of the Clifford groups associated with these HW groups, respectively. There can exist two or three orbits of equivalent SIC POMs depending on the order of the symmetry group. In addition, we established complete equivalence relations among all group covariant SIC POMs in dimension 3 and classified inequivalent ones by means of the geometric phases associated with fiducial states.

Finally, we briefly discussed some generalizations of the previous results to dimensions that are not necessarily prime. In particular, we proved that, in any prime-power dimension not equal to 3, two HW covariant SIC POMs are equivalent if and only if they are on the same orbit. We also showed that the set of Hoggar lines is not covariant with respect to the usual HW group, thereby confirming a long-standing speculation. A major problem left open is to determine the relation between the symmetry group of an HW covariant SIC POM and the Clifford group (see Conjecture 8.10).

# Two-qubit SIC POMs

---

## 9.1 Introduction

When the dimension is not a prime, SIC POMs can exhibit many exotic features not present in those in prime dimensions. To illustrate, in this chapter we take HW covariant SIC POMs in dimension 4 as an example, which exhibit remarkable additional symmetry beyond what is reflected in the name<sup>1</sup>. The situation in other dimensions is also discussed briefly when appropriate.

According to the analysis of Appleby et al. [10] (see also Refs. [232, 245]), there exists a single orbit of 256 fiducial states in dimension 4, constituting 16 SIC POMs. We shall characterize these fiducial states and SIC POMs by examining the symmetry transformations within a given SIC POM and among different SIC POMs. The symmetry group of each SIC POM is shown to be a subgroup of the Clifford group, thereby extending previous results on prime dimensions [279] (see Chapter 8). In addition, we find 16 additional SIC POMs by a suitable regrouping of the 256 fiducial states and demonstrate that they are unitarily equivalent to the 16 original SIC POMs by establishing an explicit unitary transformation. These additional SIC POMs were also noticed by Grassl [119]. Furthermore, we show that similar regrouping phenomena also appear on the orbits 8b and 12b among all the orbits cataloged by Scott and Grassl [245] and propose a unified explanation of all these regrouping phenomena by virtue of the structure of the Clifford group and its normalizer explicated in Appendix H.2.

---

<sup>1</sup>This chapter is based on the following paper: H. Zhu, Y. S. Teo, and B.-G. Englert, *Two-qubit symmetric informationally complete positive-operator-valued measures*, Phys. Rev. A **82**, 042308 (APS, 2010). Part result was presented at APS March meeting 2011 [278].

We then reveal the additional structure of these SIC POMs when the four-dimensional Hilbert space is taken as the tensor product of two qubit Hilbert spaces. A concise representation of the fiducial states is introduced in terms of generalized Bloch vectors, which allows us to explore the intriguing symmetry of the two-qubit SIC POMs. In particular, when either the standard product basis or the Bell basis is chosen as the defining basis of the HW group, in eight of the 16 HW covariant SIC POMs, all the fiducial states have the same concurrence of  $\sqrt{2/5}$ . These fiducial states can be turned into each other just by local unitary transformations. SIC POMs with this attribute are particularly appealing for an experimental implementation, because local unitary transformations are much easier to implement than global ones.

## 9.2 Structure of SIC POMs in the four-dimensional Hilbert space

For  $d = 4$ , the order of the Clifford group is 768, and that of the extended Clifford group is 1536. The analysis of Appleby et al. [10] (see also Refs. [232, 245]) shows that there is only one orbit of fiducial states (under either the Clifford group or the extended Clifford group). One of the fiducial states is  $\rho_f = |\psi_f\rangle\langle\psi_f|$  with

$$|\psi_f\rangle \hat{=} \frac{1}{2\sqrt{3+\Gamma}} \begin{pmatrix} 1 + e^{-i\pi/4} \\ e^{i\pi/4} + i\Gamma^{-3/2} \\ 1 - e^{-i\pi/4} \\ e^{i\pi/4} - i\Gamma^{-3/2} \end{pmatrix}, \quad (9.1)$$

where  $\Gamma = (\sqrt{5} - 1)/2$  is the golden ratio [6, 255]. The stabilizer (within the extended Clifford group) of this fiducial state is the order-6 cyclic group generated by the antiunitary operation

$$[A_4, \chi_4] = \left[ \begin{pmatrix} -1 & 1 \\ -1 & 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix} \right] = V\hat{J}, \quad (9.2)$$

## 9.2. Structure of SIC POMs in the four-dimensional Hilbert space

---

where

$$V \cong \frac{1}{2} \begin{pmatrix} 1 & e^{i\pi/4} & -1 & e^{i\pi/4} \\ i & e^{-3i\pi/4} & i & e^{i\pi/4} \\ 1 & e^{-3i\pi/4} & -1 & e^{-3i\pi/4} \\ i & e^{i\pi/4} & i & e^{-3i\pi/4} \end{pmatrix}, \quad (9.3)$$

and  $\hat{J}$  is the complex-conjugation operator [see Eq. (7.20)]. Within the Clifford group, the stabilizer is generated by  $[A_4, \chi_4]^2$ . So there are 256 fiducial states, constituting 16 SIC POMs on the orbit [6, 232].

### 9.2.1 Symmetry transformations within an HW covariant SIC POM

In this section, we focus on the symmetry property of a single HW covariant SIC POM for  $d = 4$ . In particular, we show that the symmetry group is a subgroup of the Clifford group and that the SIC POM is covariant with respect to a unique HW group.

Since all SIC POMs form a single orbit, it is enough to focus on the SIC POM generated from the fiducial state  $\rho_f$  [see Eq. (9.1)] under the action of the HW group. To demonstrate that the symmetry group  $\overline{G}_{\text{sym}}$  (extended symmetry group  $\overline{EG}_{\text{sym}}$ ) of this SIC POM is a subgroup of the Clifford group (extended Clifford group), it suffices to show that the stabilizer of the fiducial state  $\rho_f$  within the symmetry group is the same as that within the Clifford group, which is generated by  $[A_4, \chi_4]^2$ .

To simplify the notation in the following discussion, we use the ordered pair  $(k_1, k_2)$  to represent the fiducial state defined by the ket  $D_{k_1, k_2} |\psi_f\rangle$ . Under the action of  $[A_4, \chi_4]^2$ , the 15 fiducial states other than  $\rho_f \hat{=} (0, 0)$  in the SIC POM form five orbits:

$$\begin{aligned} O_1 &= \{(1, 0), (0, 3), (3, 1)\}, & O_2 &= \{(3, 3), (3, 2), (2, 3)\}, & O_3 &= \{(0, 1), (1, 3), (3, 0)\}, \\ O_4 &= \{(1, 2), (2, 1), (1, 1)\}, & O_5 &= \{(2, 0), (0, 2), (2, 2)\}. \end{aligned} \quad (9.4)$$

Any unitary transformation in the stabilizer (in the symmetry group of the SIC POM) of  $\rho_f$  must preserve triple products of the form  $\text{tr}(\rho_{j_1} \rho_{j_2} \rho_{j_3})$ , where  $\rho_{j_1}$ ,  $\rho_{j_2}$  and  $\rho_{j_3}$  are any triple of distinct fiducial states in the SIC POM. However, at least one of these triple products would be violated if there exists any unitary transformation in

the stabilizer other than those generated by  $[A_4, \chi_4]^2$ . This contradiction demonstrates that the symmetry group of each HW covariant SIC POM for  $d = 4$  is a subgroup of the Clifford group.

According to the previous discussion, the order of the symmetry group  $\overline{G}_{\text{sym}}$  (extended symmetry group  $\overline{EG}_{\text{sym}}$ ) of each SIC POM is 48 (96), which is much smaller than that of the symmetry group of a 15-dimensional regular simplex. It is not always possible to transform a pair of fiducial states to another pair with either a unitary or an antiunitary operation within the extended symmetry group. Since the HW group is a normal Sylow 2-subgroup of  $\overline{G}_{\text{sym}}$ , it follows from Sylow's theorem that  $\overline{G}_{\text{sym}}$  contains only one nice error basis, namely, the HW group (see Appendix G for a brief introduction to Sylow subgroups and Sylow's theorem). In other words, each HW covariant SIC POM in dimension 4 is covariant with respect to a unique nice error basis. This observation extends the previous result on prime dimensions not equal to 3 [279] and confirms Conjectures 8.10, 8.11, and 8.12 in the case of dimension 4 (see Chapter 8).

### 9.2.2 Symmetry transformations among HW covariant SIC POMs

In this section we investigate the symmetry transformations among the 16 HW covariant SIC POMs. To describe such operations, we need to label each SIC POM with a unique number for later reference. Let  $V_n \in [F_n, \mathbf{0}]$  for  $n = 1, 2, \dots, 16$ , where the  $F_n$ s are defined by

$$\begin{aligned}
 & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 3 \\ 5 & 7 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 6 & 7 \\ 3 & 5 \end{pmatrix}, \\
 & \begin{pmatrix} 0 & 3 \\ 5 & 5 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 7 & 1 \end{pmatrix}, \quad \begin{pmatrix} 6 & 7 \\ 7 & 7 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 \\ 1 & 6 \end{pmatrix}, \\
 & \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 6 & 7 \\ 1 & 4 \end{pmatrix}, \quad \begin{pmatrix} 0 & 3 \\ 5 & 6 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 7 & 0 \end{pmatrix}, \\
 & \begin{pmatrix} 6 & 7 \\ 5 & 6 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 7 & 2 \end{pmatrix}, \quad \begin{pmatrix} 0 & 3 \\ 5 & 0 \end{pmatrix}.
 \end{aligned} \tag{9.5}$$



## 9.2. Structure of SIC POMs in the four-dimensional Hilbert space

---

Table 9.1: Arrangement of the 16 HW covariant SIC POMs. Each number  $n$ , with  $1 \leq n \leq 16$ , represents the HW covariant SIC POM obtained from transforming the SIC POM containing the fiducial state  $\rho_f$  with the unitary transformation  $[F_n, \mathbf{0}]$ , where the  $F_n$ s are specified in Eq. (9.5).

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

They have been chosen with foresight to simplify the following discussion. Denote by SIC POM No.  $n$  the image of the SIC POM containing the fiducial state  $\rho_f$  under the transformation  $V_n$ ; then this correspondence between the 16 HW covariant SIC POMs and the 16 numbers  $n = 1, 2, \dots, 16$  is one to one.

We are now concerned with the transformations only among different SIC POMs, so the two groups  $G_{\text{SYM}} = \overline{\text{C}}(4)/\overline{\text{D}}$  and  $EG_{\text{SYM}} = \overline{\text{EC}}(4)/\overline{\text{D}}$  properly describe the symmetry operations of interest. As an abstract group,  $G_{\text{SYM}}$  is isomorphic to the special linear group  $\text{SL}(2, \mathbb{Z}_4)$  (see Sec. 7.3.3); likewise,  $EG_{\text{SYM}}$  is isomorphic to the extended special linear group  $\text{ESL}(2, \mathbb{Z}_4)$ . Coincidentally, the order of  $G_{\text{SYM}}$  is the same as the order of the symmetry group  $\overline{G}_{\text{SYM}}$  of a single SIC POM, namely, 48; however, the two groups are not isomorphic. The group  $G_{\text{SYM}}$  consists of the identity, seven order-2 elements, eight order-3 elements, 24 order-4 elements, and eight order-6 elements. Order-2 elements form three conjugacy classes, with one, three, and three elements, respectively. Order-4 elements constitute four conjugacy classes, each with six elements; elements in two of the classes are the inverses of those in the other two classes. Order-3 elements make up a single conjugacy class, and so do order-6 elements. The center of  $G_{\text{SYM}}$  is generated by the order-2 element that has only one conjugate.

If the 16 HW covariant SIC POMs are arranged in a  $4 \times 4$  square as in Table 9.1, then the effect of the symmetry transformations of the group  $G_{\text{SYM}}$  can be delineated in a pictorial way as shown in Fig. 9.1. The effect of only one element in each conjugacy class is shown; the effect of other group elements within the same conjugacy class can be derived simply by permuting the columns that represent the SIC POMs.

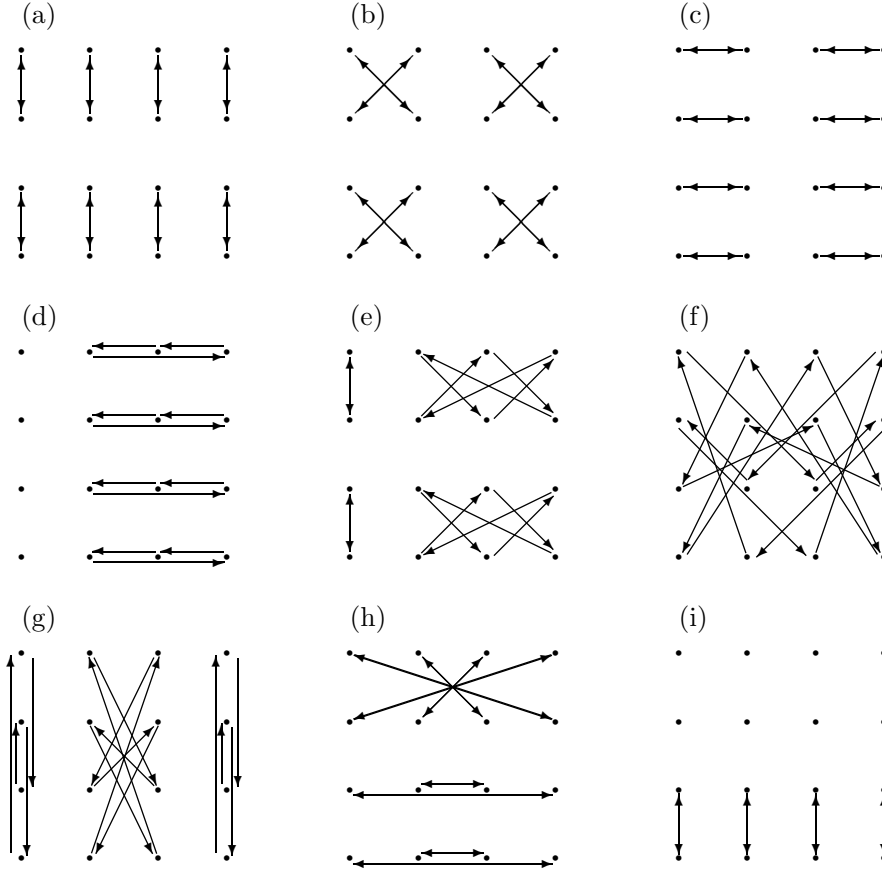


Figure 9.1: Illustration of the symmetry transformations among the 16 HW covariant SIC POMs induced by elements in the group  $EG_{\text{SYM}} = \overline{\text{EC}}(4)/\overline{D}$  (see Sec. 9.2.2). Here, every dot represents a SIC POM arranged as in Table 9.1, and every arrow starts from a SIC POM before the symmetry transformation and ends at the SIC POM after the symmetry transformation. Only one element in each conjugacy class of  $G_{\text{SYM}}$  is chosen as a representative, and the transformations induced by other elements within the same conjugacy class can be derived by permuting the columns. In the case of order-4 elements, only two out of the four conjugacy classes are chosen; the elements in the other two conjugacy classes are the inverses of the elements in the two conjugacy classes, so their transformations can be constructed by reversing the arrows. Plot (a): order-2 element in the center of  $G_{\text{SYM}}$ ; plots (b) and (c): two order-2 elements from the other two conjugacy classes, respectively; plots (d) and (e): an order-3 element and an order-6 element; plots (f) and (g): two order-4 elements from two different conjugacy classes; plot (h): the complex-conjugation operation; plot (i): the complex-conjugation operation followed by an appropriate order-2 element in  $G_{\text{SYM}}$ .

## 9.2. Structure of SIC POMs in the four-dimensional Hilbert space

---

According to Fig. 9.1, the symmetry transformations among the 16 SIC POMs can be decomposed into row transformations and column transformations. In addition to the identity, all order-3 elements and one class of order-2 elements [see plots (d) and (c) in Fig. 9.1] transform the SIC POMs within each row, and with the same effect in every row. They constitute an order-12 normal subgroup of  $G_{\text{SYM}}$ , which can also be identified with the alternating group of the four columns. The quotient group of  $G_{\text{SYM}}$  with respect to this group of row transformations acts as an order-4 cyclic subgroup [generated by the cyclic permutation of the four rows  $1 \rightarrow 3 \rightarrow 2 \rightarrow 4 \rightarrow 1$ ; see plots (f) and (g) in Fig. 9.1] of the symmetry group of the four rows. Similarly, the quotient group of  $EG_{\text{SYM}}$  acts as an order-8 subgroup of the symmetry group of the four rows.

### 9.2.3 SIC POM regrouping phenomena

By a suitable regrouping of the 256 fiducial states on the orbit of the Clifford group, 16 additional SIC POMs can be constructed, which turn out to be equivalent to the 16 original SIC POMs. This peculiar regrouping phenomenon was first noticed by Grassl [119] and rediscovered by us [283]. In this section, we show that this phenomenon is deeply rooted in the structure of the Clifford group and its normalizer pertinent to each dimension that is a multiple of 4. We also uncover all similar regrouping phenomena on the orbits cataloged in Ref. [245] and offer a unified explanation of them.

The construction of these additional SIC POMs is best illustrated when the 16 original HW covariant SIC POMs are arranged in a  $4 \times 4$  square as in Table 9.1. Under the action of the Abelian subgroup  $H = \{1, X^2, Z^2, X^2Z^2\}$  of the HW group, the 16 fiducial states in each SIC POM form four orbits of equal size. Given four fiducial states in a SIC POM connected by  $H$ , in each of the other three SIC POMs in the same row, there exist four fiducial states that are also connected by  $H$  and whose overlaps with the given four fiducial states are all equal to  $\frac{1}{5}$ —the value required to form a SIC POM in dimension 4. It turns out that the 16 states thus selected also constitute a SIC POM. In this way, four additional SIC POMs can be constructed by regrouping the fiducial

states in the four original SIC POMs in each row, that is, 16 additional SIC POMs in total. Meanwhile, inspection of the pairwise fidelities among all the 256 fiducial states shows that no more SIC POMs can be constructed by regrouping these fiducial states.

The regrouping phenomena mentioned above is closely related to the structure of the Clifford group when the dimension is a multiple of 4, in particular, the existence of two normal subgroups that are both HW groups, but in different bases, as explicated in Appendix H.2. More specifically, the 16 additional SIC POMs and the 16 original SIC POMs are covariant with respect to the two HW groups, respectively, and they can be transformed into each other by non-Clifford unitary transformations in the normalizer of the Clifford group, say, the one specified in Eq. (H.15). Actually, the roles of the two sets of SIC POMs can interchange if we start from the additional normal HW group in the Clifford group. Incidentally, analysis shows that the full symmetry group of the orbit of the 256 fiducial states happens to be the normalizer of the Clifford group (see Theorem H.5). We need to go beyond the Clifford group to understand all the symmetry operations of the 32 SIC POMs although the symmetry group of each SIC POM is a subgroup of the Clifford group.

Detailed analysis of HW covariant SIC POMs cataloged by Scott and Grassl [245] shows that, besides orbit 4a, regrouping phenomena also appear on each generic orbit in dimension 3 [279], as well as on the orbits 8b and 12b [283]. In all these cases, the symmetry group of each SIC POM contains antiunitary operations, so the orbits of the Clifford group and that of the extended Clifford group coincide.

The regrouping phenomena on the orbits 8b and 12b share a strikingly similar pattern as that on the orbit 4a as described as follows. All original SIC POMs on each orbit can be divided into sets of equal size 4, and four additional SIC POMs can be constructed by a suitable regrouping of the fiducial states of the four SIC POMs in each set. Each additional SIC POM and each original SIC POM in the set share  $d^2/4$  fiducial states. These common features are not merely a coincidence, but are deeply rooted in the structure of the Clifford group in dimensions that are multiples of 4, as mentioned above. The additional SIC POMs and the original SIC POMs are covariant

### 9.3. Two-qubit SIC POMs

---

with respect to the two normal HW groups in the Clifford group, respectively, and they can be transformed into each other by unitary transformations in the normalizer of the Clifford group. Meanwhile, every fiducial state of one HW group is simultaneously a fiducial state of the other HW group. This is really remarkable, noting that the existence of a fiducial state of one HW group is already a surprise since the equations satisfied by the fiducial state are highly over determined. On the other hand, if there exists a simultaneous fiducial state of both HW groups, then a similar regrouping phenomenon will appear. This observation can help search for potential regrouping phenomena in Hilbert spaces of other dimensions.

In dimension 3, there exists a continuous family of orbits of SIC POMs [6, 232, 275] (see Sec. 8.5.2). Each generic orbit is composed of 72 fiducial states, which constitute eight SIC POMs. By suitably regrouping these fiducial states, 24 additional SIC POMs can be constructed. However, the regrouping phenomenon exhibits quite a different nature compared with the previous three cases. For example, the additional SIC POMs are not equivalent to the original ones [279].

### 9.3 Two-qubit SIC POMs

In this section, we study the additional structure of SIC POMs when the four-dimensional Hilbert space is perceived as a tensor product of two qubit Hilbert spaces. These emergent properties are generally basis dependent, because it matters how the four-dimensional Hilbert space is tensor-factored into two two-dimensional spaces. We shall focus on the product basis and the Bell basis in the following discussion since the new features are most appealing in the two special cases.

Before discussing those properties pertinent to specific bases, we first mention a characteristic that is basis independent. The average purity of the single-qubit reduced states of states in any two-qubit SIC POM is  $\frac{4}{5}$ ; that is, the average tangle or squared concurrence of states in any two-qubit SIC POM is  $\frac{2}{5}$ . More generally, in a bipartite Hilbert space of subsystem dimensions  $d_1$  and  $d_2$ , the average purity of the reduced states in each party of states in any SIC POM is  $(d_1 + d_2)/(d_1 d_2 + 1)$ —this value is

Table 9.2: Arrangement of the components of the generalized Bloch vector of a two-qubit state.

	$r_x$	$r_y$	$r_z$
$s_x$	$C_{xx}$	$C_{xy}$	$C_{xz}$
$s_y$	$C_{yx}$	$C_{yy}$	$C_{yz}$
$s_z$	$C_{zx}$	$C_{zy}$	$C_{zz}$

equal to the average over all pure states in the bipartite Hilbert space with respect to the Haar measure [284]. This attribute follows from the fact that a SIC POM is a 2-design [232, 244, 245] (see Appendix B).

### 9.3.1 Two-qubit SIC POMs in the product basis

For a single qubit, any state can be expressed in terms of the identity operator 1 and the three Pauli operators  $\sigma_j$  for  $j = x, y, z$ ; the coefficients of expansion define the Bloch vector. In the case of two qubits, any state  $\rho$  can be expressed in terms of the pairwise tensor products among the four operators,

$$\rho = \frac{1}{4} \left( 1 \otimes 1 + \sum_{j=x,y,z} r_j 1 \otimes \sigma_j + \sum_{j=x,y,z} s_j \sigma_j \otimes 1 + \sum_{j,k=x,y,z} C_{jk} \sigma_j \otimes \sigma_k \right). \quad (9.6)$$

In analogy to the case of a single qubit, the coefficients

$$v = (r_x, r_y, r_z, s_x, s_y, s_z, C_{xx}, C_{xy}, C_{xz}, C_{yx}, C_{yy}, C_{yz}, C_{zx}, C_{zy}, C_{zz})^T \quad (9.7)$$

define the generalized Bloch vector (GBV) of  $\rho$ . Although quite common, this terminology is slightly abusive and somewhat misleading. The  $s$  column and the three columns of  $C$  in Table 9.2 transform like three-dimensional column vectors when the first qubit is rotated by local unitary transformations; likewise, the  $r$  row and the three rows of  $C$  are row vectors for local unitary transformations of the second qubit. In short, the two single-qubit Bloch vectors are *vectors*, and the two-qubit “double vector”  $C$  is a *dyadic*.

The structure of the GBVs of the 256 fiducial states is best illustrated when the components are arranged as in Table 9.2. When the standard product basis is chosen

### 9.3. Two-qubit SIC POMs

---

as the defining basis of the HW group, that is,  $|e_0\rangle = |00\rangle$ ,  $|e_1\rangle = |01\rangle$ ,  $|e_2\rangle = |10\rangle$ , and  $|e_3\rangle = |11\rangle$ , these fiducial states divide into two classes, according to the structure of their GBVs. The first class consists of the 128 fiducial states in the first eight SIC POMs, and the second class of the 128 fiducial states in the remaining eight (according to the labeling scheme described in Sec. 9.2.2). The structure of the GBV of each fiducial state in the first class is illustrated in the top tabular of Table 9.3, where

$$a, b, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3 = \pm 1, \quad A_{\pm 1} = \frac{\sqrt{1 \pm \sqrt{5}}}{\sqrt{5}}, \quad B = \frac{1}{\sqrt{5}}. \quad (9.8)$$

The eight sign factors  $a, b, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$  obey the constraint

$$ab\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3 = 1. \quad (9.9)$$

There are seven free sign factors, giving a total of 128 combinations of values, and specifying the 128 fiducial states in the first class. In addition, each SIC POM in the first class is characterized by the following three sign functions, each taking on a constant value for the fiducial states in a given SIC POM:

$$h_1 = b\alpha_2\alpha_3\beta_3, \quad h_2 = \alpha_1\alpha_2\alpha_3, \quad h_3 = ab\alpha_1. \quad (9.10)$$

Each combination of the eight sign factors that does not satisfy Eq. (9.9) specifies a Hermitian operator  $Q$  which is not positive semidefinite. Nevertheless,  $Q$  can be written as the partial transpose (with respect to the computational basis) of a fiducial state and satisfies the following 15 equations as each fiducial state does:

$$\text{tr}(QD_{k_1, k_2}QD_{k_1, k_2}^\dagger) = \frac{1}{5} \quad \text{for} \quad (k_1, k_2) \neq (0, 0). \quad (9.11)$$

These equations imply that the 16 operators generated from  $Q$  under the action of the HW group also form a 15-dimensional regular simplex in the Hilbert space of Hermitian operators.

The structure of the GBV of each fiducial state in the second class is shown in the

Table 9.3: The structure of the generalized Bloch vector of each fiducial state in the first class (top) and that in the second class (bottom) when the standard product basis is chosen as the defining basis of the HW group.

	$\beta_1 A_b$	$\beta_2 A_{-b}$	$\beta_3 B$
$\alpha_1 B$	$\alpha_1 \beta_1 A_{-b}$	$\alpha_1 \beta_2 A_b$	$\alpha_1 \beta_3 B$
$\alpha_2 A_a$	$\sqrt{2} a \alpha_2 \beta_1 A_a \delta_{a,b}$	$\sqrt{2} a \alpha_2 \beta_2 A_a \delta_{a,-b}$	$\alpha_2 \beta_3 A_{-a}$
$\alpha_3 A_{-a}$	$-\sqrt{2} a \alpha_3 \beta_1 A_{-a} \delta_{-a,b}$	$-\sqrt{2} a \alpha_3 \beta_2 A_{-a} \delta_{a,b}$	$\alpha_3 \beta_3 A_a$
	$\beta_1 A_a$	$\beta_2 A_a$	$\beta_3 B$
$\alpha_1 B$	$\alpha_1 \beta_1 A_{-a}$	$\alpha_1 \beta_2 A_{-a}$	$\alpha_1 \beta_3 B$
$\alpha_2 A_a$	$a^{(1-b)/2} \alpha_2 \beta_1 \Gamma_{-b}$	$a^{(1+b)/2} \alpha_2 \beta_2 \Gamma_b$	$\alpha_2 \beta_3 A_{-a}$
$\alpha_3 A_a$	$a^{(1+b)/2} \alpha_3 \beta_1 \Gamma_b$	$a^{(1-b)/2} \alpha_3 \beta_2 \Gamma_{-b}$	$\alpha_3 \beta_3 A_{-a}$

bottom tabular of Table 9.3, where

$$\Gamma_{\pm 1} = \frac{\sqrt{1 \pm \Gamma}}{\sqrt{5}}, \quad (9.12)$$

and  $A_{\pm 1}, B$  are defined in Eq. (9.8). There is also one constraint among the eight sign factors, namely,

$$b \alpha_1 \alpha_2 \alpha_3 \beta_1 \beta_2 \beta_3 = 1. \quad (9.13)$$

Each SIC POM in the second class is also specified by three sign functions:

$$h_1 = ab \alpha_1 \beta_3, \quad h_2 = -\alpha_1 \alpha_2 \alpha_3, \quad h_3 = b \alpha_1. \quad (9.14)$$

When the SIC POMs are arranged as in Table 9.1 and Eq. (9.5), the sign function  $h_1$  is a constant in each row, whereas  $h_2$  and  $h_3$  are constants in each column (see Table 9.4). This is one of the reasons why the numbering in Table 9.1 was done that way.

Since the standard product basis is chosen as the defining basis of the HW group, both  $Z$  and  $X^2$  are local unitary operators. Under their actions, the 16 fiducial states in each SIC POM divide into two sets of equal size, such that the eight fiducial states in each set have the same concurrence. For each SIC POM in the second class, eight fidu-



### 9.3. Two-qubit SIC POMs

---

Table 9.4: The values of the three sign functions  $h_1, h_2, h_3$  [defined in Eqs. (9.10) and (9.14)] for each HW covariant SIC POM labeled according to Sec. 9.2.2.

	$h_2 = 1$	$h_2 = 1$	$h_2 = -1$	$h_2 = -1$
	$h_3 = -1$	$h_3 = 1$	$h_3 = 1$	$h_3 = -1$
$h_1 = -1$	1	2	3	4
$h_1 = 1$	5	6	7	8
$h_1 = 1$	9	10	11	12
$h_1 = -1$	13	14	15	16

cial states have concurrence of  $\sqrt{(2 + 2\sqrt{5})}/5$ , and the other eight of  $\sqrt{(2 - 2\sqrt{5})}/5$ . What is peculiar for each SIC POM in the first class is that all 16 fiducial states have the same concurrence of  $\sqrt{2/5}$  (tangle of  $\frac{2}{5}$ ). One could say that these symmetric IC POMs are not just symmetric; they are supersymmetric. This supersymmetry is remarkable, indeed.

Since the average tangle of fiducial states in any two-qubit SIC POM is  $\frac{2}{5}$ , and since the concurrence and the entanglement of formation are both concave functions of the tangle, it follows that the average concurrence or entanglement of formation of states in a SIC POM is maximized when all states have the same tangle (or concurrence), as is the case for each SIC POM in the first class.

Fiducial states in each SIC POM in the first class can be turned into each other by just local unitary transformations. This property is particularly appealing for an experimental implementation of these POMs, because local unitary transformations are much easier to realize than global ones. As a side remark, the eight SIC POMs in the first class can be transformed into each other with local Clifford unitary transformations, and so can the eight SIC POMs in the second class.

Although all fiducial states of each SIC POM in the first class have the same concurrence, it is impossible to connect all fiducial states with only local unitary transformations in the symmetry group  $\overline{G}_{\text{sym}}$  of the SIC POM. Moreover, this conclusion is independent of the basis chosen. Seeking a contradiction, suppose the opposite is true. To connect all fiducial states in the SIC POM, the order of the local unitary transformation group is necessarily a multiple of 16. Meanwhile, the order must divide

the order of  $\overline{G}_{\text{sym}}$ , namely, 48. It follows that the local unitary transformation group has order either 16 or 48 and thus contains the HW group as a subgroup, since the latter is the only nice error basis in  $\overline{G}_{\text{sym}}$  according to Sec. 9.2.1. However, the HW group cannot be a local unitary group. This contradiction verifies our claim.

Incidentally, in each SIC POM, exactly two fiducial states share the same single-qubit reduced states for the first qubit, and the same holds for the second qubit. The end points of the Bloch vectors of the eight distinct single-qubit reduced states for each qubit form quite a regular pattern, especially for the second qubit and for each SIC POM in the first class, in which they form a cube.

In a generic bipartite Hilbert space, SIC POMs such that all fiducial states have the same Schmidt coefficients are quite rare. As far as the SIC POMs cataloged by Scott and Grassl [245] are concerned, such phenomena appear only on the orbits 4a, 6a, 12b, and 28c, and only when  $d_2 = 2$  (accordingly,  $d_1 = d/d_2 = 2, 3, 6,$  and 14). The special case of two-qubit SIC POMs is recovered when  $d_1 = 2$ . The reason behind these peculiar phenomena is still not clear. In all the cases, the concurrence is well defined. According to the discussion at the beginning of this section, the purity of the reduced density matrix of each fiducial state is  $(d_1 + 2)/(2d_1 + 1)$ , and the concurrence of each fiducial state is  $\sqrt{2(d_1 - 1)/(2d_1 + 1)}$ . On the other hand, when  $d_1$  and  $d_2$  are coprime, it is possible to choose a suitable basis such that the HW group factorizes. Then, all fiducial states in any HW covariant SIC POM are automatically equivalent under local unitary transformations [121, 210, 245].

In the eight-dimensional Hilbert space, the set of Hoggar lines [146] is covariant with respect to an alternative version of the HW group, the three-qubit Pauli group [116, 275]. Since all fiducial states are connected to each other by a local unitary group, they have the same Schmidt coefficients with respect to any bipartition of the three parties. In addition to this attribute, the set of Hoggar lines also boasts a huge symmetry group, as we shall see in Sec. 10.4.

### 9.3. Two-qubit SIC POMs

Table 9.5: The structure of the generalized Bloch vector of each fiducial state in the first class (top) and that in the second class (bottom) when the Bell basis is chosen as the defining basis of the HW group.

	$\beta_1 B$	$\sqrt{2}\beta_2 A_a \delta_{a,b}$	$\sqrt{2}\beta_3 A_{-a} \delta_{-a,b}$
$\alpha_1 B$	$\alpha_1 \beta_1 B$	$\sqrt{2}\alpha_1 \beta_2 A_{-a} \delta_{a,b}$	$\sqrt{2}\alpha_1 \beta_3 A_a \delta_{-a,b}$
$\alpha_2 A_b$	$\alpha_2 \beta_1 A_{-b}$	$b\alpha_2 \beta_2 A_a$	$b\alpha_2 \beta_3 A_{-a}$
$\alpha_3 A_b$	$\alpha_3 \beta_1 A_{-b}$	$a\alpha_3 \beta_2 A_a$	$-a\alpha_3 \beta_3 A_{-a}$
	$\beta_1 B$	$\beta_2 \Gamma_{-b}$	$\beta_3 \Gamma_b$
$\alpha_1 B$	$\alpha_1 \beta_1 B$	$-b\alpha_1 \beta_2 \Gamma_{-b}$	$b\alpha_1 \beta_3 \Gamma_b$
$\alpha_2 A_{-a}$	$\alpha_2 \beta_1 A_a$	$(-a)^{(1-b)/2} \alpha_2 \beta_2 A_{-a}$	$(-a)^{(1+b)/2} \alpha_2 \beta_3 A_{-a}$
$\alpha_3 A_a$	$\alpha_3 \beta_1 A_{-a}$	$a^{(1-b)/2} \alpha_3 \beta_2 A_a$	$a^{(1+b)/2} \alpha_3 \beta_3 A_a$

#### 9.3.2 Two-qubit SIC POMs in the Bell basis

Now consider the Bell basis as the defining basis of the HW group, that is,

$$\begin{aligned}
 |e_0\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), & |e_1\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \\
 |e_2\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), & |e_3\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).
 \end{aligned} \tag{9.15}$$

The structure of the GBV of each fiducial state in the first class (according to the classification scheme in Sec. 9.3.1) is shown in the top tabular of Table 9.5, where  $A_{\pm 1}$  and  $B$  are defined in Eq. (9.8). As in the case of the product basis, here the sign factors  $a, b, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$  may assume only the two values  $\pm 1$  and obey one constraint, namely,

$$ab\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3 = 1. \tag{9.16}$$

In addition, each SIC POM is specified by three sign functions:

$$h_1 = -b\alpha_1\beta_1\beta_2\beta_3, \quad h_2 = -\beta_1\beta_2\beta_3, \quad h_3 = ab\beta_1. \tag{9.17}$$

The structure of the GBV of each fiducial state in the second class is shown in the bottom tabular of Table 9.5, where  $\Gamma_{\pm 1}$  is defined in Eq. (9.12). Here the sign factors

$a, b, \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$  obey the constraint

$$- ab\alpha_1\alpha_2\alpha_3\beta_1\beta_2\beta_3 = 1. \quad (9.18)$$

Likewise, each SIC POM is specified by three sign functions:

$$h_1 = ab\alpha_1, \quad h_2 = -a\beta_1\beta_2\beta_3, \quad h_3 = b\beta_1. \quad (9.19)$$

The values of the three sign functions for each SIC POM are the same as that in the case of the product basis (see Table 9.4). By contrast, now fiducial states in the second class rather than the first class have the same concurrence of  $\sqrt{2/5}$ , whereas fiducial states in the first class have concurrence of either  $\sqrt{(2 + 2\sqrt{\Gamma})/5}$  or  $\sqrt{(2 - 2\sqrt{\Gamma})/5}$ .

## 9.4 Summary

We have explored the structure of HW covariant SIC POMs in the four-dimensional Hilbert space, in particular, the symmetry transformations within one SIC POM and among different SIC POMs. The symmetry group of each SIC POM is shown to be a subgroup of the Clifford group. We also constructed 16 additional SIC POMs by regrouping the 256 fiducial states and demonstrated their equivalence with the 16 original SIC POMs by deriving an explicit unitary transformation. Furthermore, we uncovered all similar regrouping phenomena of HW covariant SIC POMs and offered a unified explanation of them.

We then revealed the rich structure of these HW covariant SIC POMs when the four-dimensional Hilbert space is taken as the tensor product of two qubit Hilbert spaces. The introduction of generalized Bloch vectors allowed us to represent the fiducial states and SIC POMs in a concise way and to explore their structure in a systematic manner. In both the product basis and the Bell basis, eight of the 16 SIC POMs consist of fiducial states with the same concurrence of  $\sqrt{2/5}$ . They are thus not just symmetric IC POMs, but supersymmetric IC POMs.

# Symmetry and equivalence

---

## 10.1 Introduction

Since Zauner posed his conjecture [275], our belief in the existence of SIC POMs in arbitrary finite dimensions has strengthened considerably, thanks to the efforts of many researchers in the past decade [6, 10, 115, 116, 232, 245]. On the other hand, it is increasingly more difficult to construct new solutions with traditional approaches since they rely heavily on the computational power. Moreover, our understanding about the properties and implications of SIC POMs is far from satisfactory. There are many elusive questions in this regard. For example, what symmetry can a SIC POM possess? Even for HW covariant SIC POMs, the problem has largely remained open. Actually, the special case in dimension 3 was settled only recently [279] (see Chapter 8), although the fiducial states had been known for more than a decade [275]. Much less is known about SIC POMs covariant with respect to other nice error bases. Only a few such examples are investigated in the literature [116, 146, 232, 275]. What is worse, except for the Hoggar lines (see Sec. 8.6), it is not clear whether these examples are just HW covariant SIC POMs, but in different guises. The main difficulty lies in computing the symmetry group of a given SIC POM and in determining the equivalence relation between two SIC POMs. We have witnessed a partial success of the group-theoretic approach in the case of prime and prime-power dimensions, but to treat the general problem systematically entails some new idea.

In this Chapter, we establish a simple connection between the symmetry problem of a SIC POM and the automorphism problem of a graph constructed out of the triple products of the states in the SIC POM, based on a recent result of Appleby,

Flammia, and Fuchs [14]. By virtue of this connection, we propose an efficient algorithm for determining the symmetry group of a SIC POM. A variant of the algorithm also allows us to tackle the SIC POM equivalence problem, which can be reduced to the graph isomorphism problem. In addition to its significance in practical calculation, the graph-theoretic approach offers a fresh perspective for understanding SIC POMs, which complements the group-theoretic approach explored previously.

As an application of the graph-theoretic approach, we compute the symmetry groups of all SIC POMs known in the literature and establish complete equivalence relations among them, thereby furnishing a pretty clear picture about those known SIC POMs. Several persistent confusions concerning this subject are also clarified. This result further helps us figure out all additional nice error bases contained in the symmetry group of each SIC POM. In addition, we show by numerics that each SIC POM that can be generated by any nice error basis cataloged by Klappenecker and Rötteler [166] is equivalent to either an HW covariant SIC POM or the set of Hoggar lines. Also, any SIC POM in dimensions 2 to 7 is covariant with respect to the HW group.

### 10.2 SIC POMs and graph automorphism problem

In this section we start a graph-theoretic approach to the symmetry problem and the equivalence problem. The initial motivation for this study is to devise practical algorithms for computing the symmetry group of a SIC POM and for determining the equivalence relation between two SIC POMs. In an effort to understand the efficiency of such an algorithm, we manage to reduce the symmetry problem of SIC POMs to the automorphism problem of graphs, which has been studied for many decades in the community of graph theory [16, 17, 170, 194, 258]; see Appendix I for a brief introduction to the basic concepts in graph theory. Following the same line of thinking, we can reduce the SIC POM equivalence problem to the graph isomorphism problem. In retrospect, this connection could have been anticipated much earlier.

## 10.2. SIC POMs and graph automorphism problem

---

### 10.2.1 Unitary symmetry and permutation symmetry

Recall that the symmetry group of a SIC POM is composed of all unitary transformations that leave the SIC POM invariant. Any unitary transformation in the symmetry group induces a permutation among the outcomes of the SIC POM, henceforth denoted by  $\Pi_j$  for  $j = 0, 1, \dots, d^2 - 1$ . It is straightforward to determine the permutation once the unitary transformation is given. To tackle the reverse problem, we need to introduce some new concepts .

Following the convention in Sec. 3.2.1, we can identify the operators  $\Pi_j$  as vectors  $|\Pi_j\rangle\rangle$  in the space of Hermitian operators. Denote the reconstruction operators by  $|\Theta_j\rangle\rangle$  [see Eqs. (3.6) and (3.11)]; then we have

$$\text{tr}(\Pi_j \Theta_k) = \delta_{jk} \quad \text{for } j, k = 0, 1, \dots, d^2 - 1. \quad (10.1)$$

Now each permutation  $\sigma$  among the  $\Pi_j$ s can be represented by a superoperator,

$$\mathcal{S}_\sigma = \sum_{j=0}^{d^2-1} |\Pi_{\sigma(j)}\rangle\rangle \langle\langle \Theta_j |, \quad (10.2)$$

which satisfies  $\mathcal{S}_\sigma |\Pi_j\rangle\rangle = |\Pi_{\sigma(j)}\rangle\rangle$ . If  $\sigma$  is induced by a unitary transformation, then the action of  $\mathcal{S}_\sigma$  on the operators is equivalent to the conjugation by a unitary operator  $U_\sigma$ , which is unique up to an overall phase factor. To determine  $U_\sigma$ , let  $X'$  and  $Z'$  be the images of  $X$  and  $Z$  under the action of  $\mathcal{S}_\sigma$ , that is,  $|X'\rangle\rangle = \mathcal{S}_\sigma |X\rangle\rangle$  and  $|Z'\rangle\rangle = \mathcal{S}_\sigma |Z\rangle\rangle$ . Let  $|e'_r\rangle$  be an eigenket of  $Z'$  with eigenvalue  $\omega^r$  and define

$$U'_\sigma := \sum_{r=0}^{d-1} |e'_r\rangle \langle e_r|; \quad (10.3)$$

then the operator  $U''_\sigma := U'^\dagger_\sigma U_\sigma$  commutes with  $Z$  and is thus diagonal in the computational basis. Consequently, the operator  $X'' := U''_\sigma X U''^\dagger_\sigma$  has the form

$$X'' = \sum_{r=0}^{d-1} e^{i\theta_r} |e_{r+1}\rangle \langle e_r|, \quad (10.4)$$

where  $\prod_{r=0}^{d-1} e^{i\theta_r} = 1$ . This equation determines  $U''_\sigma$  up to an overall phase factor,

$$U''_\sigma = \sum_{r=0}^{d-1} e^{i\phi_r} |e_r\rangle\langle e_r|, \quad (10.5)$$

where

$$\phi_0 = 0, \quad \phi_r = \sum_{s=0}^{r-1} \theta_s \quad \text{for } r = 1, 2, \dots, d-1. \quad (10.6)$$

Now the unitary operator corresponding to the permutation  $\sigma$  is given by  $U_\sigma = U'_\sigma U''_\sigma$  up to a phase factor, where  $U'_\sigma$  and  $U''_\sigma$  are determined by Eqs. (10.3) and (10.5).

The above approach can also determine the unitary transformation between two different SIC POMs, except that Eq. (10.2) should be replaced by

$$\mathcal{S}_\sigma = \sum_{j=0}^{d^2-1} |\Pi'_{\sigma(j)}\rangle\langle\langle\Theta_j|, \quad (10.7)$$

where the  $\Pi'_j$ s are the outcomes of the target SIC POM.

### 10.2.2 A connection with the graph automorphism problem

According to the discussion in the previous section, to compute the symmetry group of a SIC POM, it suffices to figure out those permutations that can be induced by unitary transformations. Still, how can we determine whether a given permutation can be realized as a unitary transformation? This problem was recently solved by Appleby, Flammia, and Fuchs [14] by means of the triple products of states in the SIC POM. In this section we aim to turn their idea into a powerful practical tool for solving the symmetry and the equivalence problems. The first step along this direction is to reformulate their result in the graph-theoretic language. In this way, we can reduce the SIC POM symmetry problem to the graph automorphism problem and the equivalence problem to the isomorphism problem. This reduction is helpful not only for more efficient calculations but also for a deeper understanding of the characteristics of SIC POMs.



## 10.2. SIC POMs and graph automorphism problem

---

### 10.2.2.1 Triple products, angle tensor, and angle matrix

Following the notation in Sec. 10.2.1, we define triple products and the angle tensor as follows (see also Ref. [14]),

$$T_{jkl} = \frac{\text{tr}(\Pi_j \Pi_k \Pi_l)}{|\text{tr}(\Pi_j \Pi_k \Pi_l)|}, \quad \vartheta_{jkl} = \arg(T_{jkl}). \quad (10.8)$$

By convention, all the phases take on values between  $-\pi$  and  $\pi$ , with the two end points identified. The angle  $\vartheta_{jkl}$  is well known as the Bargmann invariant [27] or the geometric phase [1, 35], which has played an important role in various branches of physics [250]. Recently, it has also found many applications in the study of SIC POMs, such as determining the set of click probabilities in state estimation with SIC POMs [12], connecting SIC POMs with Lie algebras [14], and classifying group covariant SIC POMs in dimension 3 [279] (see Chapter 8).

By definition, the triple products satisfy the relations

$$T_{jkl} = T_{klj} = T_{ljk} = T_{jlk}^* = T_{lkj}^* = T_{kjl}^*, \quad (10.9a)$$

$$T_{jkl} = T_{mjk} T_{mkl} T_{mlj}. \quad (10.9b)$$

As for the angle tensor, we have

$$\vartheta_{jkl} = \vartheta_{klj} = \vartheta_{ljk} = -\vartheta_{jlk} = -\vartheta_{lkj} = -\vartheta_{kjl}, \quad (10.10a)$$

$$\vartheta_{jkl} = \vartheta_{mjk} + \vartheta_{mkl} + \vartheta_{mlj}. \quad (10.10b)$$

The *angle matrix*  $\Lambda^{(j)}$  is defined as the  $(d^2 - 1) \times (d^2 - 1)$  antisymmetric matrix that is composed of entries  $\Lambda_{kl}^{(j)} = \vartheta_{jkl}$  with  $k, l \neq j$ . It determines the angle tensor according to Eq. (10.10b).

If the SIC POM is generated by a nice error basis with index group  $H$  from a fiducial state  $\rho = |\psi\rangle\langle\psi|$ , then the states of the SIC POM can be labeled by the elements of the index group,  $h \rightarrow U_h \rho U_h^\dagger$  for  $h \in H$ , and so can the entries of the angle tensor and

angle matrices. Denote by  $e$  the identity of  $H$  and define  $\Lambda := \Lambda^{(e)}$ ; then we have

$$\Lambda_{g,h} = \arg(\langle \psi | U_g | \psi \rangle \langle \psi | U_g^\dagger U_h | \psi \rangle \langle \psi | U_h^\dagger | \psi \rangle), \quad g, h \neq e. \quad (10.11)$$

Group covariance implies that

$$\Lambda_{g,h} = \Lambda_{h^{-1}, h^{-1}g} = \Lambda_{g^{-1}h, g^{-1}}. \quad (10.12)$$

If  $g$  and  $h$  commute, then

$$\Lambda_{g,h} = \phi(g, h) + \Lambda_{h^{-1}, g^{-1}}, \quad (10.13)$$

where

$$e^{i\phi(g,h)} = U_g U_h^\dagger U_g^\dagger U_h. \quad (10.14)$$

For an HW covariant SIC POM, each row or column of the angle matrix can be marked by a pair of indices  $\mathbf{k} = (k_1, k_2) \in \mathbb{Z}_d^2$  or a single index  $k = dk_1 + k_2$ . Accordingly, Eqs. (10.12) and (10.13) reduce to

$$\begin{aligned} \Lambda_{\mathbf{k}, \mathbf{k}'} &= \Lambda_{-\mathbf{k}', \mathbf{k} - \mathbf{k}'} = \Lambda_{\mathbf{k}' - \mathbf{k}, -\mathbf{k}}, \\ \Lambda_{\mathbf{k}, \mathbf{k}'} &= \Lambda_{-\mathbf{k}', -\mathbf{k}} - \frac{2\pi}{d} \langle \mathbf{k}, \mathbf{k}' \rangle. \end{aligned} \quad (10.15)$$

### 10.2.2.2 The connection

Denote by  $\Pi_j$ s and  $\Pi'_j$ s the outcomes of two SIC POMs and let  $\vartheta_{jkl}$  and  $\vartheta'_{jkl}$  be the respective angle tensors. If there is a unitary transformation that maps  $\Pi_j$  to  $\Pi'_j$  for  $j = 0, 1, \dots, d^2 - 1$ , then  $\vartheta'_{jkl} = \vartheta_{jkl}$  for all  $j, k, l$ . Remarkably, Appleby, Flammia, and Fuchs [14] demonstrated that the converse is also true: Two SIC POMs are unitarily equivalent whenever their angle tensors are equal. Their result applied to the same SIC POM implies that any permutation  $\sigma$  among the outcomes can be realized by a unitary transformation if and only if  $\sigma$  preserves the angle tensor; that is,  $\vartheta_{\sigma(jkl)} = \vartheta_{jkl}$ . As a consequence, the symmetry group of a SIC POM is isomorphic to

## 10.2. SIC POMs and graph automorphism problem

---

the automorphism group of its angle tensor, which defines a 3-uniform hypergraph in graph-theoretic terms; two SIC POMs are unitarily equivalent if and only if their angle tensors are isomorphic (see Appendix I). By the same token, the extended symmetry group of a SIC POM is isomorphic to the extended automorphism group of its angle tensor; two SIC POMs are antiunitarily equivalent if and only if their angle tensors are skew isomorphic. Note that antiunitary operations reverse the sign of the angle tensor.

For practical applications, it is much easier to work with graphs defined by angle matrices rather than hypergraphs defined by angle tensors. Such simplification is possible because the angle tensors are completely determined by the angle matrices. In terms of the angle matrices, the connection between SIC problems and graph problems can be summarized as follows:

1. The (extended) stabilizer of  $\Pi_j$  is isomorphic to the (extended) automorphism group of  $\Lambda^{(j)}$ .
2.  $\Pi_j$  and  $\Pi_k$  are connected to each other by a (anti) unitary operation in the (extended) symmetry group of the SIC POM if and only if  $\Lambda^{(j)}$  and  $\Lambda^{(k)}$  are (skew) isomorphic.
3. A SIC POM is group covariant if and only if all the angle matrices  $\Lambda^{(j)}$ s are isomorphic.
4. Two group covariant SIC POMs are (anti) unitarily equivalent if and only if their angle matrices are (skew) isomorphic. The condition may be relaxed by requiring group covariance on only one of the two SIC POMs.

Consequently, to determine the symmetry group of a given SIC POM, it remains to determine the automorphism groups of the angle matrices  $\Lambda^{(j)}$  and the isomorphism relations among them. For a group covariant SIC POM, it suffices to determine the automorphism group of one angle matrix. In spite of such a great simplification, the problem is still intractable with brute force, simply because there are too many permutations to enumerate. For example, to determine the symmetry group of a group covariant SIC POM in dimension 6 in this way, it would take the age of the universe

even with the fastest computer in the world nowadays. Fortunately, there are much more efficient algorithms for this purpose, which are the focus of the next section.

### 10.2.3 An algorithm

Motivated by the symmetry problem of SIC POMs discussed in the previous section, in this section we present a simple algorithm for computing the automorphism group of a real symmetric or antisymmetric matrix. A variant of the algorithm allows determining the isomorphism relation between two such matrices. Here the matrix is identified with the adjacency matrix of a certain graph, although this identification is not essential. Actually, the algorithm had originally been written before we realized its connection with graph automorphism algorithms. In this way, nevertheless, it is much easier to visualize what the algorithm does in each step and to make contact with the vast literature on the graph automorphism problem [16, 17, 170, 185, 194, 258].

The main idea of the algorithm can be summarized as follows. The weights of the edges induce an *ordered partition* of the vertices into disjoint blocks, which specifies a necessary condition on whether two vertices can be connected by an automorphism. If the partition is *complete* in the sense that each block contains only one vertex, then the automorphism group is trivial. Otherwise, we can select a *reference vertex* from a block with more than one vertex, say, a block with the most vertices, and refine the partition according to the weights of the edges incident to the reference vertex. By repeating this process if necessary, we can make the ordered partition complete after selecting enough reference vertices. Then there exists a one-to-one correspondence between the automorphisms of the graph and the images of the sequence of reference vertices, which can be determined recursively.

To describe the algorithm in more detail, it is advisable to introduce some additional terminology. Let  $A$  be the adjacency matrix of the graph under consideration, which has  $n$  vertices. The order of the  $n$  vertices can be specified by a sequence  $\mathbf{l} = (l_1, l_2, \dots, l_n)^T$  of the  $n$  integers  $1, 2, \dots, n$ , which represent the  $n$  vertices. The adjacency matrix  $A_{\mathbf{l}}$  with respect to the order specified by  $\mathbf{l}$  can be derived from  $A$  by permuting its rows and

## 10.2. SIC POMs and graph automorphism problem

---

columns accordingly. A partition of  $n$  consists of a set of positive integers  $\lambda_1, \lambda_2, \dots, \lambda_s$  that sum up to  $n$  and is denoted by  $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_s]$ , where  $s$  is the height of the partition. The pair  $\mathbf{l}$  and  $\lambda$  specify an ordered partition of the  $n$  vertices in a self-explaining way: The first block  $B_1$  consists of the first  $\lambda_1$  vertices in  $\mathbf{l}$ , the second block  $B_2$  of the next  $\lambda_2$  vertices, and so on. The order of the vertices within the same block is not essential, although the order of natural numbers is a convenience choice.

A block containing only one vertex is called a bachelor block, and the corresponding vertex is called a bachelor vertex. Bachelor vertices will play an important role in the following algorithm. Let  $\eta_j = 1 + \sum_{k=1}^{j-1} \lambda_k$ ; then the set  $F_{\mathbf{l}}(\lambda)$  of bachelor vertices reads

$$F_{\mathbf{l}}(\lambda) = \{l_j | j \in F(\lambda)\}, \quad F(\lambda) = \{\eta_j | \lambda_j = 1, 1 \leq j \leq s\}. \quad (10.16)$$

Here  $F(\lambda)$  is the set of indices of these vertices with respect to the sequence  $\mathbf{l}$ . Equally important are the largest blocks, blocks that contain the most vertices. Define  $Q_{\mathbf{l}}(\lambda)$  as the first largest block of the ordered partition specified by  $\mathbf{l}$  and  $\lambda$ , and  $r_{\mathbf{l}}(\lambda)$  the first vertex in the block  $Q_{\mathbf{l}}(\lambda)$ . Let  $\lambda_{\max} = \max_{j=1}^s \lambda_j$  and  $k$  be the smallest number such that  $\lambda_k = \lambda_{\max}$ . Define

$$r(\lambda) := \eta_k, \quad Q(\lambda) = \{\eta_k, \eta_k + 1, \dots, \eta_k + \lambda_k - 1\}; \quad (10.17)$$

then we have

$$r_{\mathbf{l}}(\lambda) = l_{r(\lambda)}, \quad Q_{\mathbf{l}}(\lambda) := \{l_q | q \in Q(\lambda)\}. \quad (10.18)$$

Once  $\mathbf{l}$  is specified,  $F(\lambda)$  and  $F_{\mathbf{l}}(\lambda)$  provide the same amount of information; so do  $r(\lambda)$  and  $r_{\mathbf{l}}(\lambda)$ , as well as  $Q(\lambda)$  and  $Q_{\mathbf{l}}(\lambda)$ .

Now we are ready to present the algorithm for computing the automorphism group of a symmetric or antisymmetric matrix  $A$ . In the latter case, the matrix  $A$  is allowed to take nonzero diagonal entries, so that the algorithm applies to the angle matrices introduced in Sec. 10.2.2 even if they are neither symmetric nor antisymmetric in the usual sense when the angle  $\pi$  and other angles appear simultaneously. To avoid unnecessary complication, we assume that each diagonal entry of  $A$  is different from

each nondiagonal entry. This assumption is not restrictive at all since it can easily be satisfied by subtracting from  $A$  a multiple of the identity matrix if necessary, which does not affect the automorphism group.

The algorithm consists of a main algorithm and a routine called recursive ordered partition (ROP). The function of ROP is to update the ordered-partition in terms of  $\mathbf{l}$  and  $\lambda$  when a set of vertices is fixed. The set may be specified in two different ways:  $E$  specifies the vertices explicitly, whereas  $\tilde{E}$  specifies the indices of these vertices with respect to  $\mathbf{l}$ .

1. Convert  $E$  to  $\tilde{E}$  if necessary. If  $\tilde{E}$  is empty, let  $L$  be the matrix constructed from  $A_{\mathbf{l}}$  by sorting each row with respect to the partition  $\lambda$ . Otherwise, let  $L$  be the matrix composed of all columns of  $A_{\mathbf{l}}$  whose indices belong to  $\tilde{E}$  (maintain the order of the columns).
2. Let  $N$  be the matrix formed by juxtaposing  $L$  and  $\mathbf{l}$  horizontally. Sort the rows of  $N$  (according to the dictionary order) with respect to the partition  $\lambda$  and denote the resulting matrix by  $N'$ .
3. Update  $\mathbf{l}$  with the last column of  $N'$ . Let  $L'$  be the matrix composed of all columns of  $N'$  except the last one; refine the partition  $\lambda$  according to the rows of  $L'$  and denote the resulting partition by  $\lambda'$ .
4. If  $\tilde{E}$  is empty and  $\lambda'$  is identical to  $\lambda$  or if  $\lambda'$  is complete, update  $\lambda$  with  $\lambda'$  and exit the routine; otherwise, repeat the above steps after updating  $\tilde{E}$  with  $F(\lambda') \setminus (F(\lambda) \cup \tilde{E})$  and  $\lambda$  with  $\lambda'$ .

Here sorting with respect to the partition  $\lambda$  means that only elements within the same block as determined by  $\lambda$  are sorted, whereas the relative orders of elements belonging to different blocks do not change.

The main algorithm consists of two stages: The first stage is to choose a sequence of reference vertices  $r_1, r_2, \dots, r_m$  by selecting one vertex each time from a largest block under the current partition and refining the partition until it is complete. Denote by  $G_0$  the automorphism group of  $A$  and by  $G_j$  the common stabilizer of  $r_1, r_2, \dots, r_j$

## 10.2. SIC POMs and graph automorphism problem

---

for  $j = 1, 2, \dots, m$ . Let  $O_j$  be the orbit of  $r_j$  under the action of  $G_{j-1}$  and  $C_j$  be a transversal (also called set of left coset representatives) of  $G_j$  within  $G_{j-1}$ ; then there is a one-to-one correspondence between the vertices in  $O_j$  and the automorphisms in  $C_j$ . The second stage is to determine the  $O_j$ s and the  $C_j$ s recursively, thereby determining the  $G_j$ s; note that  $G_m$  is trivial and that  $C_m$  can be identified with  $G_{m-1}$ . It is not necessary and sometimes not practical to record all elements of the  $G_k$ s explicitly.

Initialization:  $\mathbf{l}^{(0)} = (1, 2, \dots, n)^T$ ,  $\lambda^{(0)} = [n]$ ,  $\tilde{E}$  is empty, and  $j = 1$ .

1. Update  $\mathbf{l} = \mathbf{l}^{(j-1)}$  and  $\lambda = \lambda^{(j-1)}$  to  $\mathbf{l}^{(j)}$  and  $\lambda^{(j)}$  with ROP.
  - (a) If  $\lambda^{(j)}$  is not complete, select  $r_j := r_{\mathbf{l}^{(j)}}(\lambda^{(j)})$  as the  $j$ th reference vertex and denote its index with respect to  $\mathbf{l}^{(j)}$  by  $\tilde{r}_j$ , which is equal to  $r(\lambda^{(j)})$ . Define  $Q_j = Q_{\mathbf{l}^{(j)}}(\lambda^{(j)})$  and  $\tilde{Q}_j = Q(\lambda^{(j)})$ ; then  $O_j$  is a subset of  $Q_j$ . Repeat this step after replacing  $\tilde{E}$  with  $\{\tilde{r}_j\}$  and  $j$  with  $j + 1$ .
  - (b) Otherwise, record the number of reference vertices  $m := j - 1$ . If  $m = 0$ , then  $G_0$  is trivial; exit the program.
2. For each  $q \in Q_m$  that is not equal to  $r_m$ , run ROP with the input  $\mathbf{l} \rightarrow \mathbf{l}^{(m)}$ ,  $\lambda \rightarrow \lambda^{(m)}$ , and  $E \rightarrow \{q\}$ ; denote the output by  $\mathbf{l}'$  and  $\lambda'$ .
  - (a) If  $\lambda'$  is complete and  $A_{\mathbf{l}'}$  is identical to  $A_{\mathbf{l}^{(m+1)}}$ , then  $\sigma(\mathbf{l}^{(m+1)}, \mathbf{l}')$  is an automorphism in  $G_{m-1}$  that maps  $r_m$  to  $q$ , where  $\sigma(\mathbf{l}, \mathbf{l}')$  denotes the permutation that maps  $l_j$  to  $l'_j$  for  $j = 1, 2, \dots, n$ .
  - (b) Otherwise, there is no such automorphism.

At the end of this step,  $G_{m-1}$  (or equivalently  $C_m$ ) and  $O_m$  can be determined. If  $m = 1$ , exit the program.

3. For each  $q \in Q_{m-1}$  and  $q \neq r_{m-1}$ , run ROP with the input  $\mathbf{l} \rightarrow \mathbf{l}^{(m-1)}$ ,  $\lambda \rightarrow \lambda^{(m-1)}$ , and  $E \rightarrow \{q\}$ ; denote the output by  $\mathbf{l}'$  and  $\lambda'$ .
  - (a) If  $\lambda' \neq \lambda^{(m)}$ , then there is no automorphism in  $G_{m-2}$  that maps  $r_{m-1}$  to  $q$ .
  - (b) Otherwise, any automorphism in  $G_{m-2}$  that maps  $r_{m-1}$  to  $q$  maps  $r_m$  to some element in  $Q_{\mathbf{l}'}(\lambda^{(m)})$ , which has the same number of elements as  $Q_m$ .

- (c) For each  $q' \in Q_{\mathbf{l}'}(\lambda^{(m)})$ , run ROP with the input  $\mathbf{l} \rightarrow \mathbf{l}'$ ,  $\lambda \rightarrow \lambda'$ , and  $E \rightarrow \{q'\}$ ; denote the output by  $\mathbf{l}''$  and  $\lambda''$ .
- i. If  $\lambda''$  is complete and  $A_{\mathbf{l}''}$  is identical to  $A_{\mathbf{l}^{(m+1)}}$ , then  $\sigma(\mathbf{l}^{(m+1)}, \mathbf{l}'')$  is an automorphism in  $G_{m-2}$  that maps  $r_{m-1}$  to  $q$ ; continue Step 3 with  $q$  updated.
  - ii. Otherwise, continue Step 3(c) with  $q'$  updated. If this condition cannot be satisfied after testing  $|Q_m| - |C_m| + 1$  elements in  $Q_{\mathbf{l}'}(\lambda^{(m)})$ , then there is no such automorphism.

At the end of this step,  $C_{m-1}$  and  $O_{m-1}$  can be determined. If  $m = 2$ , exit the program.

4. Determine  $C_{m-2}, O_{m-2}, \dots, C_1, O_1$  recursively by applying a similar procedure as in Step 3.

As we have seen, isomorphism tests are basic building blocks of the above algorithm, which is perhaps not so surprising in view of the close relation between automorphism and isomorphism. Therefore, it is straightforward to turn the automorphism algorithm into an isomorphism algorithm.

The choice of reference vertices is not unique, and many other choices work equally well as long as they are selected consistently. Step 2 can be improved by partitioning  $Q_m$  into equivalent classes according to the automorphisms already determined and testing only one element in each equivalent class. The same idea also applies to Step 3 and Step 4.

In the worst-case scenario, to determine whether  $q \in Q_j$  is connected to  $r_j$  by an automorphism in  $G_{j-1}$ , the main algorithm may need to call ROP an exponential number (in  $m - j + 1$ ) of times. Fortunately, such a situation almost never occurs in practice. In the other extreme, it suffices to call ROP  $m - j + 1$  times if the graph satisfies the condition  $O_j = Q_j$  for  $j = 1, 2, \dots, m$ ; such a graph is called a *nice graph*. For a nice graph, the automorphism group can be determined efficiently. Besides,  $A_{\mathbf{l}_C}$  provides a canonical form of  $A$ , where  $\mathbf{l}_C = \mathbf{l}^{(m+1)}$  is the sequence of the vertices after



### 10.3. HW covariant SIC POMs

---

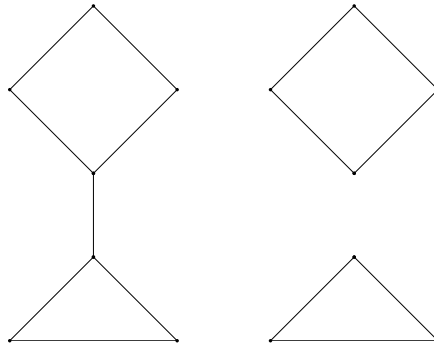


Figure 10.1: A nice graph (left plot) and a “wicked” graph (right plot).

all the reference vertices are chosen and the ordered partition is complete. Although it is generally not known a priori whether a graph is nice or not, our algorithm can test this property efficiently. Computer simulation shows that almost all randomly generated graphs are nice. The graphs defined by the angle matrices of all SIC POMs known so far are also nice, as we shall see in Sec. 10.3.2. Nevertheless, “wicked” graphs do exist: A simple example is the disjoint union of a triangle and a square (see the right plot of Fig. 10.1); it can be turned into a nice graph by adding one edge between one vertex of the triangle and one vertex of the square (see the left plot of Fig. 10.1).

### 10.3 HW covariant SIC POMs

In the past decade, there has been tremendous progress in constructing SIC POMs in small dimensions, most of which are covariant with respect to the HW group. Analytical solutions of HW covariant SIC POMs have been constructed in dimensions 2–16, 19, 24, 28, 31, 35, 37, 43, 48 [6, 9, 10, 115, 116, 117, 118, 119, 120, 232, 245, 275]; numerical solutions with high precision have been computed up to dimension 67 [232, 245]. Except in dimension 3, a comprehensive list of HW covariant SIC POMs can be found in Appendix A of Ref. [245]. For dimension 3, this appendix lists three orbits of SIC POMs out of a continuous family, which are representative of three distinct symmetry types discussed in Sec. 8.4.

In contrast with the overwhelming solutions available to us, our understanding about HW covariant SIC POMs is pretty poor. Many persistent open questions per-

tain to their symmetry properties. For example, what symmetry do they possess and what relations exist among different solutions? In view of this situation, a thorough investigation of all known solutions is highly desirable. In this section, we determine the symmetry groups of all HW covariant SIC POMs known in the literature and establish complete equivalence relations among them based on the algorithm described in Sec. 10.2. We then uncover all additional nice error bases that can generate these HW covariant SIC POMs.

### 10.3.1 SIC POMs in dimension 3 revisited

To illustrate the idea presented in Sec. 10.2, let us take HW covariant SIC POMs in dimension 3 as an example. To determine the (extended) symmetry group of a group covariant SIC POM, it suffices to determine the (extended) stabilizer of each fiducial state, which is isomorphic to the (extended) automorphism group of the angle matrix. Consider the SIC POM generated from the fiducial state  $|\psi(t)\rangle \hat{=} (0, 1, -e^{it})^T/\sqrt{2}$  [see Eq. (8.2)]; the angle matrix is given by

$$\Lambda(t) = \begin{pmatrix} 0 & \pi & -\frac{\pi}{3} & -\frac{\pi}{3} & -\frac{\pi}{3} & \frac{\pi}{3} & \frac{\pi}{3} & \frac{\pi}{3} \\ \pi & 0 & \frac{\pi}{3} & \frac{\pi}{3} & \frac{\pi}{3} & -\frac{\pi}{3} & -\frac{\pi}{3} & -\frac{\pi}{3} \\ \frac{\pi}{3} & -\frac{\pi}{3} & 0 & \frac{\pi}{3} & -\frac{\pi}{3} & \pi - 3t & \frac{\pi}{3} - 3t & -\frac{\pi}{3} - 3t \\ \frac{\pi}{3} & -\frac{\pi}{3} & -\frac{\pi}{3} & 0 & \frac{\pi}{3} & \frac{\pi}{3} - 3t & -\frac{\pi}{3} - 3t & \pi - 3t \\ \frac{\pi}{3} & -\frac{\pi}{3} & \frac{\pi}{3} & -\frac{\pi}{3} & 0 & -\frac{\pi}{3} - 3t & \pi - 3t & \frac{\pi}{3} - 3t \\ -\frac{\pi}{3} & \frac{\pi}{3} & \pi + 3t & -\frac{\pi}{3} + 3t & \frac{\pi}{3} + 3t & 0 & -\frac{\pi}{3} & \frac{\pi}{3} \\ -\frac{\pi}{3} & \frac{\pi}{3} & -\frac{\pi}{3} + 3t & \frac{\pi}{3} + 3t & \pi + 3t & \frac{\pi}{3} & 0 & -\frac{\pi}{3} \\ -\frac{\pi}{3} & \frac{\pi}{3} & \frac{\pi}{3} + 3t & \pi + 3t & -\frac{\pi}{3} + 3t & -\frac{\pi}{3} & \frac{\pi}{3} & 0 \end{pmatrix}. \quad (10.19)$$

When  $t$  is not a multiple of  $\frac{\pi}{9}$ , the off diagonal entries of  $\Lambda(t)$  take on nine distinct values:  $\pm\frac{\pi}{3}$ ,  $\pm(\frac{\pi}{3} - 3t)$ ,  $\pm(\frac{\pi}{3} + 3t)$ ,  $\pm(\pi - 3t)$ , and  $\pi$ , with multiplicities 18, 3, 3, 3, and 2, respectively. Based on the algorithm described in Sec. 10.2.3, one can show that  $\text{Aut}(\Lambda(t))$  is the order-3 group generated by the permutation (1)(2)(3 5 4)(6 7 8) in the disjoint-cycle representation, and that  $\text{Aut}_E(\Lambda(t))$  is the order-6 group generated by (1)(2)(3 5 4)(6 7 8) and (1)(2)(3 8)(4 6)(5 7). This conclusion can also be verified

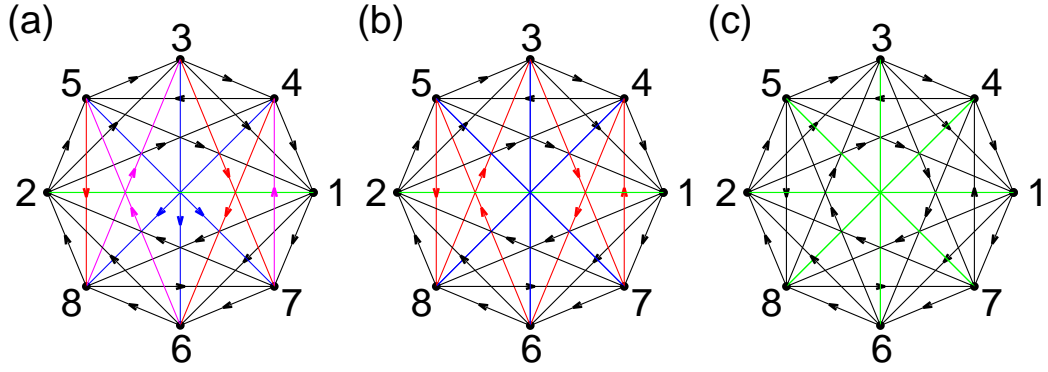


Figure 10.2: Graph representation of the angle matrix of the HW covariant SIC POM in dimension 3 generated from the fiducial state  $|\psi(t)\rangle \hat{=} (0, 1, -e^{it})^T/\sqrt{2}$ : (a) generic orbit ( $t$  is not a multiple of  $\frac{\pi}{9}$ ); (b) exceptional orbit with  $t = \frac{\pi}{3}$ ; (c) exceptional orbit with  $t = 0$ . The dot labeled by  $k = 1, 2, \dots, 8$  in each plot represents the fiducial state  $X^{k_1} Z^{k_2} |\psi(t)\rangle$  with  $3k_1 + k_2 = k$ . Black, red, magenta, and blue arrows represent angles  $\frac{\pi}{3}$ ,  $\frac{\pi}{3} - 3t$ ,  $\frac{\pi}{3} + 3t$ , and  $\pi - 3t$ , respectively; blue and green lines represent angles 0 and  $\pi$ .

by inspecting the graph defined by  $\Lambda(t)$ , as illustrated in plot (a) of Fig. 10.2. It turns out that the two generators of  $\text{Aut}_{\mathbb{E}}(\Lambda(t))$  are induced by the two extended Clifford operations given in Eq. (8.5). Therefore,  $\text{Aut}(\Lambda(t))$  and  $\text{Aut}_{\mathbb{E}}(\Lambda(t))$  are isomorphic to the stabilizer and the extended stabilizer of  $|\psi(t)\rangle$ , respectively, as expected.

For the exceptional orbit with  $t = \frac{\pi}{3}$  (orbit 3b according to Ref. [245]), because of the equalities  $\frac{\pi}{3} - 3t = \frac{\pi}{3} + 3t \pmod{2\pi}$  and  $\pi - 3t = 0$ , the angle matrix  $\Lambda(\frac{\pi}{3})$  can be represented by the graph in plot (b) of Fig. 10.2, which can be transformed from the one in plot (a) by identifying magenta arrows with red arrows and ignoring the directions of blue arrows. These changes double the order of the automorphism group:  $\text{Aut}(\Lambda(\frac{\pi}{3}))$  is the order-6 group generated by  $(1\ 2)(3\ 7\ 4\ 6\ 5\ 8)$ , and  $\Lambda(\frac{\pi}{3})$  is isomorphic to  $-\Lambda(\frac{\pi}{3})$  under the permutation  $(1)(2)(3\ 8)(4\ 6)(5\ 7)$ , as in the generic case.

For the exceptional orbit with  $t = 0$  (orbit 3c according to Ref. [245]), owing to the equalities  $\frac{\pi}{3} - 3t = \frac{\pi}{3} + 3t = \frac{\pi}{3}$  and  $\pi - 3t = \pi$ , the angle matrix  $\Lambda(0)$  can be represented by the graph in plot (c) of Fig. 10.2, which can be transformed from the one in plot (b) by identifying red arrows with black arrows and blue lines with green lines; the green lines can be deleted without affecting the automorphism group of the graph. The group  $\text{Aut}(\Lambda(0))$  has order 24 and is generated by  $(1\ 2)(3\ 7\ 4\ 6\ 5\ 8)$  and  $(1\ 3\ 2\ 6)(4\ 5\ 8\ 7)$ ; it is

isomorphic to the stabilizer of  $|\psi(0)\rangle$ , which is in turn isomorphic to the special linear group  $\text{SL}(2, \mathbb{Z}_3)$  (see Sec. 8.5.2). Compared with the graphs in plots (a) and (b), the one in plot (c) is vertex transitive in the sense that any two vertices can be mapped to each other by an automorphism; likewise, it is arrow transitive. Such a high symmetry is unique among all HW covariant SIC POMs known so far. In addition, no nontrivial automorphism can stabilize any arrow, so the order of  $\text{Aut}(\Lambda(0))$  is equal to the number of arrows in the graph, namely, 24. In contrast, each order-3 automorphism stabilizes two vertices, which form antipodal points; for example, the stabilizer of vertex 1 (or vertex 2) happens to be the automorphism group of the graph in plot (a).

Analysis shows that the graph defined by  $\Lambda(t)$  is a nice graph regardless of the value of  $t$ ; therefore, the algorithm described in Sec. 10.2.3 can offer a canonical form of the angle matrix. When  $0 < t < \frac{\pi}{9}$ , the canonical form is

$$\Lambda_{\mathbf{l}_C}(t) = \begin{pmatrix} 0 & \frac{\pi}{3} & -\frac{\pi}{3} & \frac{\pi}{3} + 3t & -\frac{\pi}{3} + 3t & \pi + 3t & \frac{\pi}{3} & -\frac{\pi}{3} \\ -\frac{\pi}{3} & 0 & \frac{\pi}{3} & -\frac{\pi}{3} + 3t & \pi + 3t & \frac{\pi}{3} + 3t & \frac{\pi}{3} & -\frac{\pi}{3} \\ \frac{\pi}{3} & -\frac{\pi}{3} & 0 & \pi + 3t & \frac{\pi}{3} + 3t & -\frac{\pi}{3} + 3t & \frac{\pi}{3} & -\frac{\pi}{3} \\ -\frac{\pi}{3} - 3t & \frac{\pi}{3} - 3t & \pi - 3t & 0 & -\frac{\pi}{3} & \frac{\pi}{3} & -\frac{\pi}{3} & \frac{\pi}{3} \\ \frac{\pi}{3} - 3t & \pi - 3t & -\frac{\pi}{3} - 3t & \frac{\pi}{3} & 0 & -\frac{\pi}{3} & -\frac{\pi}{3} & \frac{\pi}{3} \\ \pi - 3t & -\frac{\pi}{3} - 3t & \frac{\pi}{3} - 3t & -\frac{\pi}{3} & \frac{\pi}{3} & 0 & -\frac{\pi}{3} & \frac{\pi}{3} \\ -\frac{\pi}{3} & -\frac{\pi}{3} & -\frac{\pi}{3} & \frac{\pi}{3} & \frac{\pi}{3} & \frac{\pi}{3} & 0 & \pi \\ \frac{\pi}{3} & \frac{\pi}{3} & \frac{\pi}{3} & -\frac{\pi}{3} & -\frac{\pi}{3} & -\frac{\pi}{3} & \pi & 0 \end{pmatrix}, \quad (10.20)$$

where  $\mathbf{l}_C = (6, 8, 7, 5, 4, 3, 2, 1)^T$ ; when  $\frac{\pi}{9} < t < \frac{2\pi}{9}$ , the canonical form is  $\Lambda_{\mathbf{l}'_C}(t)$ , where  $\mathbf{l}'_C = (3, 4, 5, 6, 7, 8, 1, 2)^T$ ; when  $\frac{2\pi}{9} < t < \frac{\pi}{3}$ , the canonical form is  $\Lambda_{\mathbf{l}''_C}(t)$ , where  $\mathbf{l}''_C = (6, 8, 7, 4, 3, 5, 2, 1)^T$ . Straightforward calculation shows that  $\Lambda_{\mathbf{l}_C}(t) = \Lambda_{\mathbf{l}'_C}(\frac{2\pi}{9} - t) = \Lambda_{\mathbf{l}''_C}(\frac{2\pi}{9} + t)$ , so the SIC POMs on the three orbits with  $t, \frac{2\pi}{9} - t, \frac{2\pi}{9} + t$  are unitarily equivalent. The same analysis also applies when  $t$  is a multiple of  $\frac{\pi}{9}$ . In this way, the graph-theoretic approach reproduces the conclusion of Sec. 8.5.2 in a much simpler way. Meanwhile, it furnishes a new perspective for understanding the peculiar properties of SIC POMs in dimension 3.

### 10.3. HW covariant SIC POMs

---

#### 10.3.2 Symmetry and equivalence

The graph-theoretic approach illustrated in Sec. 10.3.1 applies equally well to HW covariant SIC POMs in other dimensions. Actually, we can determine the (extended) stabilizers for all the 200 HW covariant SIC POMs known in the literature (see Appendix A of Ref. [245]) on a common PC within one hour, which was impossible even for many years in the past. It turns out that the graphs defined by the angle matrices are nice for all these SIC POMs: That is one reason our approach is so efficient. In addition, except for each SIC POM on the orbit 3c, only one reference vertex is involved in the computation of  $\text{Aut}(\Lambda)$  (see Sec. 10.2.3), so the order of the stabilizer is equal to the length of the longest orbit of the fiducial states under the action of the stabilizer. Incidentally, except in dimension 3, the order of  $\text{Aut}_{\mathbb{E}}(\Lambda)$  is equal to the order of  $\text{Aut}(|\Lambda|)$ . For the orbits 3a, 3b, and 3c, the orders of  $\text{Aut}(|\Lambda|)$  are 12, 24, and 384, respectively, which are larger than the orders of  $\text{Aut}_{\mathbb{E}}(\Lambda)$ , namely, 6, 12, and 48. Again, dimension 3 is somehow peculiar.

Detailed analysis shows that, except in dimension 3, the (extended) symmetry group of each HW covariant SIC POM known so far is a subgroup of the (extended) Clifford group, which provides strong evidence in favor of Conjecture 8.10. As a consequence of Theorems 7.4 and 8.8, two such SIC POMs are unitarily or antiunitarily equivalent if and only if they are on the same orbit of the extended Clifford group. Since the peculiarity in dimension 3 was expounded in Secs. 8.5 and 10.3.1, we now have complete equivalence relations among all HW covariant SIC POMs known so far<sup>1</sup>. Although full knowledge of the angle matrix is necessary for determining the (extended) stabilizer and the (extended) symmetry group, a few angles are enough to distinguish SIC POMs on different orbits in each dimension. For example, the minimum and the maximum over the absolute values of the entries in the angle matrices suffice to differentiate almost all inequivalent orbits.

---

<sup>1</sup>In this thesis we do not consider equivalence relations under Galois field transformations, an interesting topic that deserves further study.

### 10.3.3 Nice error bases in the symmetry group

All SIC POMs known so far can be generated by nice error bases. But what nice error bases can generate SIC POMs? Up to now, only a few such examples other than the HW group are known [116, 232, 275], partially because other nice error bases are not universal as the HW group is, and they are not so familiar to many researchers in the field. In this section, we reveal that plenty of nice error bases actually appear in the symmetry group of many known HW covariant SIC POMs. In other words, a SIC POM can be covariant with respect to more than one nice error basis, some of which can be inequivalent. Based on the analysis in Sec. 10.3.2, we figure out all these nice error bases and their equivalence relations. Since the situation in dimension 3 has been discussed in detail in Sec. 8.5, here we shall focus on the SIC POMs cataloged in Appendix A of Ref. [245]. As a byproduct, our study also reveals a potential approach for constructing tight equiangular lines that are group covariant.

According to Sec. 10.3.2, the stabilizer of each fiducial state within the extended symmetry group is identical to the stabilizer within the extended Clifford group. The latter was determined in Ref. [245] (see Table I thereof) for dimensions up to 50, and incomplete information was provided for dimensions from 51 to 67. Most fiducial states are stabilized by the order-3 Zauner operation  $[F_z, \mathbf{0}]$ , where

$$F_z := \begin{pmatrix} 0 & d-1 \\ d+1 & d-1 \end{pmatrix}. \quad (10.21)$$

Although  $F_z$  and  $F_Z$  [see Eq. (7.21)] have different orders when  $d$  is even,  $[F_z, \mathbf{0}]$  and  $[F_Z, \mathbf{0}]$  always have the same order and are conjugated to each other in the Clifford group. When  $d = 9k + 3$ , solutions 12b, 21e, 30d, 39(g,h,i,j), 48(e,g), and 66a are stabilized by the order-3 Clifford operation  $[F_a, \mathbf{0}]$  [245], where

$$F_a := \begin{pmatrix} 1 & d+3 \\ d+3k & d-2 \end{pmatrix}. \quad (10.22)$$

When  $d = k^2 - 1$ , solutions 8b, 15d, 24c, 35(i,j), and 48(f,g) are stabilized by the

### 10.3. HW covariant SIC POMs

---

order-2 permutation  $[F_b, \mathbf{0}]$  [245], where

$$F_b := \begin{pmatrix} -k & d \\ d & d - k \end{pmatrix}. \quad (10.23)$$

When  $d = (3k \pm 1)^2 + 3$ , solutions 4(a), 7(b), 19(d, e), and 28(c) are stabilized by the order-2 antiunitary operation  $[F_c, \mathbf{0}]$  [245], where

$$F_c = \begin{pmatrix} \kappa & d - 2\kappa \\ d + 2\kappa & d - \kappa \end{pmatrix}, \quad \kappa = 3k^2 \pm k + 1. \quad (10.24)$$

For dimensions from 51 to 65, our calculation shows that the stabilizer of each fiducial state is generated by the Zauner operation  $[F_z, \mathbf{0}]$ ; for dimension 66, it is generated by the Clifford operation  $[F_a, \mathbf{0}]$ ; for dimension 67, it is generated by the Clifford antiunitary operation  $[(\begin{smallmatrix} 25 & 25 \\ 42 & 50 \end{smallmatrix}), \mathbf{0}]$ , whose square is the Zauner operation  $[F_z, \mathbf{0}]$ .

By virtue of Theorem H.1, one can show that, if a SIC POM is stabilized by  $[F_z, \mathbf{0}]$  when  $d$  is divisible by 3, then it is covariant with respect to two additional nice error bases generated by the following two sets of generators, respectively:

$$XZ^2, Z^3, XV_z; \quad XZ^2, Z^3, X^2V_z; \quad (10.25)$$

where  $V_z \in [F_z, \mathbf{0}]$ . Here the first two generators in each set generate the intersection of the nice error basis with the HW group. The two nice error bases are conjugated to each other under the Clifford operation  $V_{-1}$ . When  $d = 3$ , each nice error basis is an HW group, but in a different basis. Otherwise, each one has a non-Abelian index group, and the center (of the collineation group) is the order-3 group generated by  $X^{d/3}Z^{2d/3}$ . In relation to the nice error bases cataloged by Klappenecker and Rötteler [166], the two nice error bases are equivalent to the one with the index group  $G(36, 11)$  when  $d = 6$  and to the one with the index group  $G(81, 9)$  when  $d = 9$ , in the notation adopted by GAP 3 and GAP 4. In addition, the SIC POM constructed by Grassl [116] using the group  $G(36, 11)$  is actually equivalent to each HW covariant SIC POM on the orbit

6a<sup>2</sup>. Our numerical calculation further indicates that any SIC POM covariant with respect to either  $G(36, 11)$  or  $G(81, 9)$  is also covariant with respect to the HW group (see Sec. 10.5 for more details). We thus believe that the numerical SIC POMs found by Renes et al. [232] using the two groups are covariant with respect to the HW groups.

If a SIC POM is stabilized by  $[F_a, \mathbf{0}]$  when  $d = 9k + 3$ , then it is covariant with respect to eight additional nice error bases generated by the following eight sets of generators:

$$\begin{aligned}
 X, Z^3, ZV_a; & & X, Z^3, Z^2V_a; \\
 XZ, Z^3, ZV_a; & & XZ, Z^3, Z^2V_a; \\
 XZ^2, Z^3, ZV_a; & & XZ^2, Z^3, Z^2V_a; \\
 Z, X^3, XV_a; & & Z, X^3, X^2V_a;
 \end{aligned} \tag{10.26}$$

where  $V_a \in [F_a, \mathbf{0}]$ . All these nice error bases are conjugated to each other in the Clifford group: The two nice error bases in each row are conjugated to each other under the conjugation of  $V_{-1}$ ; the first three nice error bases in each column are conjugated to each other under the conjugation of  $\left[\begin{pmatrix} -2 & -3 \\ 1 & 1 \end{pmatrix}, \mathbf{0}\right]$ ; finally, the two nice error bases in the last row are conjugated to the two nice error bases in the first row under the conjugation of  $\left[\begin{pmatrix} 6k & 6k-1 \\ 1 & 1 \end{pmatrix}, \mathbf{0}\right]$ . The center of each nice error basis in Eq. (10.26) is the order-3 group generated by the  $(d/3)$ th power of the first generator; for example, it is generated by  $X^{d/3}$  for the two nice error bases in the first row.

If a SIC POM is stabilized by  $[F_b, \mathbf{0}]$  when  $d = k^2 - 1$  is even, then it is covariant with respect to three additional nice error bases generated by the following three sets of generators:

$$X, Z^2, ZV_b; \quad XZ, Z^2, ZV_b; \quad Z, X^2, XV_b; \tag{10.27}$$

where  $V_b \in [F_b, \mathbf{0}]$ . The three nice error bases are conjugated to each other under the Zauner unitary operator  $V_Z$  (or  $V_z$ ). The center of each nice error basis is the order- $(k + 1)^2$  group generated by  $X^{k-1}$  and  $Z^{k-1}$ . When  $d = 8$ , the three nice error bases

---

<sup>2</sup>We have verified the equivalence using a SIC POM provided by Markus Grassl in private communication since we cannot confirm the correctness of the fiducial state specified in Ref. [116].



### 10.3. HW covariant SIC POMs

---

are equivalent to those with the index group  $G(64, 3)$  as listed in Ref. [166].

Incidentally,  $V_b$  simultaneously stabilizes  $(k+1)^2$  fiducial states in the SIC POM, which belong to an eigenspace of dimension  $(k^2+k)/2$ . These fiducial states constitute a set of tight equiangular lines according to Eq. (7.1), which is covariant with respect to the group generated by  $X^{k-1}$  and  $Z^{k-1}$ . In dimension 35, there are two orbits (35i and 35j; see Ref. [245]) of inequivalent SIC POMs that are stabilized by  $V_b$ ; analysis shows that the two corresponding sets of tight equiangular lines are also not equivalent. Likewise, the two sets of lines associated with orbits 48f and 48g are not equivalent. The above observation suggests a new approach for constructing tight equiangular lines that are group covariant. The potential of this line of thinking deserves further exploration.

As far as the SIC POMs cataloged in Ref. [245] are concerned, the nice error bases specified in Eqs. (10.25), (10.26), and (10.27) exhaust all additional nice error bases except for the orbits 3c and 48g. For the orbit 3c, the symmetry group of each SIC POM contains four Sylow 3-subgroups, and each Sylow 3-subgroup contains two additional nice error bases, which are conjugated to the two nice error bases in Eq. (10.25), so the symmetry group contains nine nice error bases in total (see Sec. 8.5). For the orbit 48g, the stabilizer is the order-24 group generated by the Clifford antiunitary operation  $[(\begin{smallmatrix} 4 & 37 \\ 25 & 63 \end{smallmatrix}), \mathbf{0}]$  (see Ref. [245]), which contains the two Clifford operations  $[F_a, \mathbf{0}]$  and  $[F_b, \mathbf{0}]$ . In addition to the nice error bases in Eqs. (10.26) and (10.27), the SIC POM is also covariant with respect to 12 nice error bases generated by the following 12 sets of generators:

$$\begin{array}{ll}
 X, Z^4, ZV_F; & X, Z^4, Z^{-1}V_F; \\
 XZ^3, Z^4, ZV_F; & XZ^3, Z^4, Z^{-1}V_F; \\
 X^2Z, X^4, XV_F; & X^2Z, X^4, X^{-1}V_F; \\
 XZ, Z^4, ZV_F; & XZ, Z^4, Z^{-1}V_F; \\
 XZ^2, Z^4, ZV_F; & XZ^2, Z^4, Z^{-1}V_F; \\
 Z, X^4, XV_F; & Z, X^4, X^{-1}V_F;
 \end{array} \tag{10.28}$$

where  $F = (\begin{smallmatrix} 4 & 37 \\ 25 & 63 \end{smallmatrix})^6 = (\begin{smallmatrix} 13 & 8 \\ 8 & 5 \end{smallmatrix}) \pmod{96}$ . All of them are conjugated to each other under the Clifford group: The two nice error bases in each row are conjugated to each other

under the conjugation of  $V_{-1}$ ; the nice error bases in the first three rows are connected by the conjugation of  $\left[\left(\begin{smallmatrix} 4 & 37 \\ 25 & 63 \end{smallmatrix}\right), \mathbf{0}\right]^2$ , and so are the ones in the last three rows; finally, the two sets of nice error bases are connected by the conjugation of  $\left[\left(\begin{smallmatrix} 5 & 1 \\ 1 & 58 \end{smallmatrix}\right), \mathbf{0}\right]$ . The center of each nice error basis is the order-16 group generated by  $X^{12}$  and  $Z^{12}$ . In total, each SIC POM on the orbit 48g is covariant with respect to 24 nice error bases including the HW group, the number being the largest over all HW covariant SIC POMs known so far.

In summary, when the dimension is not divisible by 3, each HW covariant SIC POM known so far, except for those on the orbit 8b, is covariant with respect to only one nice error basis, namely, the HW group. In marked contrast, when the dimension is divisible by 3, each one is covariant with respect to at least three nice error bases, which may compose two, three, or four equivalent classes depending on the dimension and the orbit. Except in the case of dimension 3, all the additional nice error bases have non-Abelian index groups, as expected from Theorem 8.8.

## 10.4 Hoggar lines

The set of Hoggar lines was first constructed by Hoggar [146] more than a decade ago by complexifying 64 lines in the four-dimensional space over the quaternion. Shortly after its discovery, it was shown to be covariant with respect to the three-qubit Pauli group by Zauner [275]. In Sec. 8.6, we proved that it is not covariant with respect to the usual HW group, thereby revealing the unique status of the set of Hoggar lines in the study of SIC POMs. Beyond this point, however, little is known about its properties.

In this section, we determine the symmetry group of the Hoggar lines and the nice error bases contained in the symmetry group. As a byproduct, our study uncovers two types of tight equiangular lines embedded in the Hoggar lines, both of which are group covariant. We also demonstrate that the SIC POM in dimension 8 constructed by Grassl [116] is actually equivalent to the Hoggar lines, thereby clarifying a persistent confusion about SIC POMs that are not covariant with respect to the HW groups.

The Hoggar lines can be generated by the three-qubit Pauli group from the fiducial

## 10.4. Hoggar lines

---

state [275]

$$|\psi_8\rangle \hat{=} \frac{1}{\sqrt{6}}(1 + i, 0, -1, 1, -i, -1, 0, 0)^T. \quad (10.29)$$

To determine the (extended) stabilizer of this fiducial state, we need to inspect the angle matrix (see Sec. 10.2). Calculation shows that the off diagonal entries of the angle matrix  $\Lambda$  assume four possible values:  $0, \pi, -\pi/2, \pi/2$  with multiplicities 24, 6, 16, 16 for each row (or each column). The group  $\text{Aut}(\Lambda)$  has order 6048, which is exceptionally large compared with the corresponding value for any HW covariant SIC POM known so far. It acts transitively on the vertices and edges (with the same weight) of the graph defined by  $\Lambda$ , as is the case for each SIC POM on the orbit 3c (see Sec. 10.3.1). In addition,  $\Lambda$  is isomorphic to  $-\Lambda$ . Accordingly, each fiducial state of the Hoggar lines is stabilized by 6048 unitary operations and the same number of antiunitary operations within the extended symmetry group. Two ordered triples of fiducial states can be mapped to each other in the symmetry group if and only if they have the same triple products.

The stabilizer of  $|\psi_8\rangle$  is the order-6048 group generated by the following two operators:

$$U_7 \hat{=} \frac{\omega^5}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 1 & 0 & -i & 0 & 0 & 0 \\ 0 & 0 & i & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & -i & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -i & 0 \\ -i & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & -i & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & i \end{pmatrix}, \quad U_{12} \hat{=} \frac{\omega^3}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & i & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & i & 0 & 0 \\ 1 & -i & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & i & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -i \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -i \end{pmatrix}, \quad (10.30)$$

where  $\omega = e^{2\pi i/8}$ . The two generators have orders 7 and 12, respectively. Their phases have been chosen such that  $|\psi_8\rangle$  is an eigenket with eigenvalue 1. The extended stabilizer is generated by the above two elements and the order-2 antiunitary operator

$$V = \hat{J} \text{diag}(-i, i, 1, 1, -1, 1, -i, -i), \quad (10.31)$$

Table 10.1: Conjugacy classes of the stabilizer of the fiducial state  $|\psi_8\rangle$  of the Hoggar lines. The class representatives are defined in Eq. (10.32). Also presented are the number of fiducial states stabilized by each class representative and the dimension of the eigenspace to which these states belong.

Representative	1	$U_2$	$U_{3a}$	$U_{3b}$	$U_{4a}$	$U_{4a}^\dagger$	$U_{4b}$	$U_6$	$U_7$	$U_7^\dagger$	$U_8$	$U_8^\dagger$	$U_{12}$	$U_{12}^\dagger$
Order	1	2	3	3	4	4	4	6	7	7	8	8	12	12
# Conjugates	1	63	56	672	63	378	504	864	864	864	756	756	504	504
States stabilized	64	16	1	4	4	4	4	1	1	1	2	2	1	1
Eigenspace dim.	8	6	2	4	3	4	4	2	2	2	2	2	1	1

where  $\hat{J}$  is the complex-conjugation operator. Calculation shows that  $U_7$ ,  $U_{12}$ , and  $V$  stabilize the three-qubit Pauli group, which implies that the extended symmetry group of the Hoggar lines is a subgroup of the extended Clifford group (of the three-qubit Pauli group). There are 240 SIC POMs on the orbit of the (extended) Clifford group, given that the Clifford group has order 92897280 [56].

The conjugacy classes of elements in the stabilizer are shown in Table 10.1, where

$$\begin{aligned}
 U_{4a} &= U_{12}^3, & U_2 &= U_{4a}^2, & U_6 &= U_{12}^2, & U_{3a} &= U_6^2, \\
 U_{3b} &= U_6 U_7^2, & U_{4b} &= U_7 U_2 U_7^\dagger U_{4a}, & U_8 &= U_6^\dagger U_7^2.
 \end{aligned}
 \tag{10.32}$$

In addition,  $U_{4a}$ ,  $U_7$ ,  $U_8$ , and  $U_{12}$  are conjugated to their respective inverses in the extended stabilizer. Therefore, two unitary operations are conjugated to each other under the extended stabilizer if and only if they have the same order and the same number of conjugates.

A closer look at Table 10.1 reveals two types of tight equiangular lines [see Eq. (7.1)] embedded in the Hoggar lines. The 16 fiducial states stabilized by  $U_2$  form tight equiangular lines in dimension 6, which are covariant with respect to the group generated by

$$\sigma_z \otimes 1 \otimes 1, \quad 1 \otimes \sigma_z \otimes 1, \quad 1 \otimes 1 \otimes \sigma_z, \quad 1 \otimes 1 \otimes \sigma_x.
 \tag{10.33}$$

The four states stabilized by  $U_{4a}$  or, equivalently, by  $U_{4b}$  form tight equiangular lines in dimension 3, which are covariant with respect to the group generated by  $\sigma_z \otimes 1 \otimes 1$

## 10.4. Hoggar lines

---

and  $1 \otimes \sigma_z \otimes 1$ . These lines can be mapped into a real Hilbert space by the unitary transformation

$$\text{diag}(\omega^3, \omega, 1, 1, i, 1, \omega^3, \omega^3), \quad (10.34)$$

and are thus equivalent to the four diagonals of the cube.

Preliminary analysis shows that the set of Hoggar lines is covariant with respect to 35344 nice error bases in total, which constitute 21 equivalent classes and 27 conjugacy classes in its symmetry group. These numbers are exceptionally large compared with those for any HW covariant SIC POM known so far. It turns out that the three-qubit Pauli group is the only nice error basis with an Abelian index group and, meanwhile, the only order-64 normal subgroup in the symmetry group [277]. Hence, there is only one orbit of SIC POMs of the (extended) Clifford group that are equivalent to the Hoggar lines according to Theorem 7.4. Incidentally, every known SIC POM, except those in dimension 3, is covariant with respect to one and only one nice error basis with an Abelian index group. This observation suggests that Conjecture 8.12 might be generalized to cover all group covariant SIC POMs, not just HW covariant ones.

Dimension 8 boasts a large number of inequivalent nice error bases [166]; it is of great interest whether these nice error bases can generate SIC POMs, especially those not known before. According to Grassl [116], a nice error basis with the non-Abelian index group  $G(64, 78)$  can generate SIC POMs. It is generated by the following three operators:

$$\begin{pmatrix} 0 & i & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -i & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & i \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & i & 0 & 0 \end{pmatrix}. \quad (10.35)$$

One of the fiducial states is

$$\frac{1}{\sqrt{6}}(0, 1, \omega, 0, 1, \sqrt{2}, 0, \omega)^T. \quad (10.36)$$

He also showed that each fiducial state in the SIC POM thus generated has a trivial stabilizer within the normalizer of the nice error basis<sup>3</sup>, which has order 1024.

It seems that the set of Hoggar lines is much more symmetric than the SIC POM constructed by Grassl if one considers symmetry operations only within the normalizers of the corresponding nice error bases, which has been the choice in most studies in the past decade because of the difficulty in determining the full symmetry group. Surprisingly, it turns out that the former can be transformed into the latter by the monomial unitary operator

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \omega^3 & 0 \\ 0 & 0 & 0 & 0 & -i & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega^5 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \omega^3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -i \\ 0 & 0 & \omega & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (10.37)$$

## 10.5 Quest for new SIC POMs

In the previous sections, we inspected all SIC POMs known in the literature to the best of our knowledge and found that they are covariant with respect to either the HW group or the three-qubit Pauli group. Does there exist a SIC POM that is not covariant with respect to either of the two groups?

In an effort to answer the above question, we performed a comprehensive numerical

---

<sup>3</sup>Markus Grassl considered the normalizer of the unitary group but not the normalizer of the collineation group. According to our convention, the normalizer has order 4096, and the stabilizer of each fiducial state has order 4.

## 10.5. Quest for new SIC POMs

---

analysis of all the SIC POMs that can be generated by those nice error bases cataloged by Klappenecker and Rötteler [166]<sup>4</sup>. In numerical optimization, a simple steepest-ascent algorithm was applied to minimize the frame potential [232]

$$\Phi = \sum_{h \in H} |\langle \psi | U_h | \psi \rangle|^4, \quad (10.38)$$

where  $H$  is the index group of the pertinent nice error basis. For each nice error basis, the optimization was repeated 10000 times. A SIC POM was obtained if  $\Phi$  reaches the threshold  $2d/(d+1)$  within numerical error. In that case, a variant of the algorithm described in Sec. 10.2.3 was employed to compare the solution with known ones for the given dimension.

Surprising or not, all the SIC POMs we found were equivalent to known solutions, as indicated in Table 10.2. The group  $G(64, 8)$  in the table was already investigated by Renes et al. [232], and the group  $G(64, 78)$  by Grassl [116] (see Sec. 10.4). In addition, we inspected several nice error bases that are tensor products of the HW groups, but did not find any SIC POM except for the Hoggar lines. Unfortunately, we still cannot answer the question posed at the beginning of this section.

Furthermore, we searched and analyzed SIC POMs in dimensions 2 to 8 without the assumption on group covariance. In contrast with the previous case, the main difference in numerical optimization was the replacement of the frame potential by

$$\Phi' = \sum_{j,k=0}^{d^2-1} |\langle \psi_j | \psi_k \rangle|^4 \quad (10.39)$$

and, accordingly, the threshold by  $2d^3/(d+1)$  [232]. Our investigation indicates that, in dimension 2 and dimensions 4 to 7, all SIC POMs are covariant with respect to the HW groups. In dimension 2, this conclusion agrees with analytical analysis. In dimension 3, it is difficult to minimize the frame potential with a simple steepest-ascent

---

<sup>4</sup>By virtue of the computer algebra system Magma, Markus Grassl has investigated many of these nice error bases for generating analytical SIC POMs for a long time. We are grateful to him for stimulating discussions and are happy to acknowledge his pioneering works on this subject.

Table 10.2: SIC POMs generated by the nice error bases cataloged by Klappenecker and Rötteler [166] (it should be noted that some distinct nice error bases listed in Ref. [166] are equivalent according to our criterion in Appendix F). HW groups and those nice error bases that cannot generate SIC POMs are not included. Each index group  $G(d^2, k)$  is represented by the single number  $k$ . All these SIC POMs are equivalent to either HW covariant SIC POMs or the Hoggar lines. Here “a” and “b” indicate the orbits of equivalent SIC POMs according to the labeling scheme in Ref. [245], and “H” represents the Hoggar lines.

$d$	Index group	Class	Orbit	$d$	Index group	Class	Orbit	$d$	Index group	Class	Orbit
6	11	1	a	8	68	1	H	8	138	2	H
8	3	1	b	8	69	1	H	8	138	3	H
8	3	2	b	8	71	1	H	8	138	4	H
8	8	1	H	8	71	2	H	8	193	1	H
8	60	1	H	8	74	1	H	8	195	1	H
8	60	2	H	8	74	2	H	8	202	1	H
8	60	3	H	8	75	1	H	8	202	2	H
8	60	4	H	8	75	2	H	8	202	3	H
8	60	5	H	8	77	1	H	8	202	4	H
8	60	6	H	8	78	1	H	8	202	5	H
8	60	7	H	8	78	2	H	9	9	1	a,b
8	62	1	H	8	90	1	H	9	9	2	a,b
8	67	1	H	8	90	2	H	9	9	3	a,b
8	67	2	H	8	91	1	H	9	9	4	a,b
8	67	4	H	8	91	2	H				

algorithm because of the existence of a continuous family of SIC POMs. Nevertheless, preliminary analysis favors the same conclusion. Beyond dimension 7, it is increasingly more difficult to hit the global minimum of the frame potential, and it is not easy to reach a reliable conclusion. It seems that group structure is not merely a convenience in constructing SIC POMs, at least for small dimensions, but the reason is still not clear. Further study is indispensable to unravel the mystery.

## 10.6 Summary and open questions

We have established a simple connection between the symmetry problem of a SIC POM and the automorphism problem of a graph constructed out of the triple products of the states in the SIC POM. Based on this connection, we developed an efficient algorithm for computing the symmetry group of the SIC POM, which is much faster than any



## 10.6. Summary and open questions

---

algorithm known before. The same idea also applies to determining the equivalence relation between two SIC POMs. In addition to providing a powerful tool for solving the symmetry and the equivalence problems, the graph-theoretic approach furnishes a fresh perspective for understanding the intriguing properties of SIC POMs.

Furthermore, we determined the symmetry groups of all SIC POMs known in the literature and established complete equivalence relations among them. Our study indicated that the set of Hoggar lines is the only known SIC POM that is not covariant with respect to the usual HW group. Except in dimension 3, the symmetry group of any HW covariant SIC POM known in the literature is a subgroup of the Clifford group. As a consequence, two such SIC POMs are unitarily or antiunitarily equivalent if and only if they are on the same orbit of the extended Clifford group. It seems that there is a deep reason for these observations, but the mystery is yet to be unraveled.

There are many elusive questions that deserve further study:

1. Do SIC POMs exist in every finite dimension?
2. Are all SIC POMs (strong) group covariant?
3. Are all group covariant SIC POMs strong group covariant?
4. Do HW covariant SIC POMs exist in every finite dimension? (This question is closely related to Zauner's conjecture [6, 232, 275].)
5. Does there exist any group covariant SIC POM that is not covariant with respect to the HW group or the three-qubit Pauli group?
6. Does there exist any continuous family of inequivalent SIC POMs in some dimension not equal to 3?
7. Does there exist any SIC POM in some dimension not equal to 3 that is covariant with respect to more than one HW group (or more than one nice error basis with an Abelian index group)?
8. Can every (strong) group covariant SIC POM be generated by a group composed of monomial matrices?

9. Is the symmetry group of any HW covariant SIC POM in every dimension not equal to 3 necessarily a subgroup of the Clifford group? (The answer is positive for any SIC POM known so far according to Sec. 10.3.2 and for any SIC POM in a prime dimension according to Theorem 8.4; see also Conjecture 8.10.)

# Several distance and distinguishability measures

---

Most figures of merit used in quantum state estimation are based on certain distance or distinguishability measures. In this appendix, we briefly review several common candidates, such as the HS distance, the trace distance, the Bures distance, and the fidelity; see Refs. [30, 97, 206] for more details.

## A.1 Hilbert–Schmidt distance and trace distance

The *Hilbert–Schmidt (HS) distance* between two quantum states  $\rho$  and  $\sigma$  is induced by the HS inner product between operators,

$$\|\rho - \sigma\|_{\text{HS}} = \sqrt{\text{tr}(\rho - \sigma)^2}. \quad (\text{A.1})$$

It is the Euclidean distance between  $\rho$  and  $\sigma$  viewed as vectors in the space of Hermitian operators.

The *trace distance* between  $\rho$  and  $\sigma$  is defined as

$$\|\rho - \sigma\|_{\text{tr}} = \frac{1}{2} \text{tr} |\rho - \sigma|. \quad (\text{A.2})$$

It is one of the most common figures of merit used in quantum state estimation, especially in experiments, because it has a nice operational interpretation, which is best manifested in a state-discrimination problem. Suppose Alice prepares one of the two states  $\rho$  and  $\sigma$  with equal probability and asks Bob to discriminate between the two

---

## Appendix A. Several distance and distinguishability measures

---

states. Bob can make a measurement with the two outcomes  $\Pi$  and  $1 - \Pi$  and declares that the state is  $\rho$  if the outcome  $\Pi$  occurs and  $\sigma$  otherwise. The maximal probability that Bob gets the right answer is given by [141]

$$\max_{0 \leq \Pi \leq 1} \frac{1}{2} [\text{tr} \rho \Pi + \text{tr} \{\sigma(1 - \Pi)\}] = \max_{0 \leq \Pi \leq 1} \frac{1}{2} [1 + \text{tr} \{\Pi(\rho - \sigma)\}] = \frac{1}{2} [1 + \|\rho - \sigma\|_{\text{tr}}], \quad (\text{A.3})$$

where we have applied the equality [206]

$$\|\rho - \sigma\|_{\text{tr}} = \max_{0 \leq \Pi \leq 1} \text{tr} [\Pi(\rho - \sigma)]. \quad (\text{A.4})$$

The proof follows from the observation that  $\rho - \sigma$  can be expressed as the difference between two positive operators with orthogonal supports. Therefore, the trace distance between two given states determines how well they can be distinguished from each other by the optimal measurement. In addition, it is also equal to the maximal trace distance between the probability distributions resulting from the same measurements on the two states, respectively [206].

### A.2 Fidelity and Bures distance

The *fidelity* between  $\rho$  and  $\sigma$  is defined as [30, 206, 257]<sup>1</sup>

$$F(\rho, \sigma) = \left( \text{tr} \sqrt{\rho^{1/2} \sigma \rho^{1/2}} \right)^2, \quad (\text{A.5})$$

or equivalently,

$$F(\rho, \sigma) = \left( \text{tr} |\rho^{1/2} \sigma^{1/2}| \right)^2. \quad (\text{A.6})$$

The second definition makes it clear that the fidelity is symmetric with respect to the two states. When  $\sigma$  is a pure state, say,  $|\psi\rangle\langle\psi|$ , the formulas can be simplified,  $F(\rho, |\psi\rangle\langle\psi|) = \langle\psi|\rho|\psi\rangle$ . The meaning of the fidelity is best manifested in a charac-

---

<sup>1</sup>It should be noted that some authors define the fidelity without the square [206], which is called the root fidelity according to our definition.

## A.2. Fidelity and Bures distance

---

terization due to Uhlmann [257], according to which  $F(\rho, \sigma)$  is equal to the maximal transition probability between the purifications of  $\rho$  and  $\sigma$ ; that is,

$$F(\rho, \sigma) = \max_{|\psi_\rho\rangle, |\psi_\sigma\rangle} |\langle\psi_\rho|\psi_\sigma\rangle|^2, \quad (\text{A.7})$$

where  $|\psi_\rho\rangle$  is a purification of  $\rho$ , and  $|\psi_\sigma\rangle$  of  $\sigma$ . This formula implies that  $F(\rho, \sigma) \leq 1$  and that the maximum is saturated if and only if  $\rho = \sigma$ . Similar to the trace distance, the fidelity between two quantum states also has a nice characterization based on the classical fidelity between probability distributions:  $F(\rho, \sigma)$  is equal to the minimal fidelity between the probability distributions that arise from the same measurements on the two states, respectively [97, 206].

The fidelity allows defining the *Bures distance*  $D_B$  [30, 50],

$$D_B^2(\rho, \sigma) = 2 - 2\sqrt{F(\rho, \sigma)}. \quad (\text{A.8})$$

When both  $\rho$  and  $\sigma$  are diagonal, the Bures distance reduces to the Hellinger distance between the diagonals of  $\rho$  and  $\sigma$ . When  $\rho \hat{=} \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$  has full rank and  $\sigma$  is infinitesimally apart, the Bures distance is explicitly given by [153]

$$D_B^2(\rho, \rho + d\rho) = \frac{1}{2} \sum_{j,k} \frac{|\langle j|d\rho|k\rangle|^2}{\lambda_j + \lambda_k}. \quad (\text{A.9})$$

Like its classical counterpart, the infinitesimal Bures distance has a clear operational meaning as it determines how well two nearby quantum states can be distinguished. In addition, it enables defining a monotone Riemannian metric in the state space that is equivalent to the metric defined by the SLD Fisher information matrix [45] (see Chapter 5).

# Weighted $t$ -designs

---

Consider a weighted set of states  $\{|\psi_j\rangle, w_j\}$  with  $0 < w_j \leq 1$  and  $\sum_j w_j = d$ . Given a positive integer  $t$ , the order- $t$  *frame potential*  $\Phi_t$  is defined as [232, 244]

$$\Phi_t = \sum_{j,k} w_j w_k |\langle \psi_j | \psi_k \rangle|^{2t} = \text{tr}(B_t^2), \quad B_t = \sum_j w_j (|\psi_j\rangle \langle \psi_j|)^{\otimes t}. \quad (\text{B.1})$$

Note that  $B_t$  is supported on the  $t$ -partite symmetric subspace, whose dimension is  $\binom{d+t-1}{t}$ . The frame potential  $\Phi_t$  is bounded from below by  $d^2 \binom{d+t-1}{t}^{-1}$ , and the bound is saturated if and only if  $B_t = d \binom{d+t-1}{t}^{-1} S_t$ , where  $S_t$  is the projector onto the  $t$ -partite symmetric subspace. The weighted set  $\{|\psi_j\rangle, w_j\}$  is a (complex projective) *weighted  $t$ -design* if the lower bound is saturated; it is a  $t$ -design if, in addition, all the weights  $w_j$  are equal [144, 145, 232, 234, 244]. By definition, a weighted  $t$ -design is also a weighted  $t'$ -design for  $t' < t$ .

For any pair of positive integers  $d$  and  $t$ , there exists a (weighted)  $t$ -design with a finite number of elements [248]. The number is bounded from below by [144, 244]

$$\binom{d + \lceil t/2 \rceil - 1}{\lceil t/2 \rceil} \binom{d + \lfloor t/2 \rfloor - 1}{\lfloor t/2 \rfloor}, \quad (\text{B.2})$$

which is equal to  $d, d^2, d^2(d+1)/2$  for  $t = 1, 2, 3$ , respectively. Any resolution of the identity consisting of pure states is a weighted 1-design. SIC POMs [232, 244, 245, 275] and complete sets of MUB [83, 155, 272] are prominent examples of 2-designs. The complete set of MUB for  $d = 2$  is also a 3-design. Our interest in weighted  $t$ -design mainly stems from their applications in studying quantum state estimation [128, 234, 244, 281] (see also Chapter 3) and SIC POMs [165, 230, 232, 244, 275] (see also Sec. 7.1).

## Proof of Lemma 4.1

---

The idea of the proof follows from the proof of Lemma 5.1 in Chapter VI of Ref. [147]. Let  $\mathbf{u}$  and  $\mathbf{v}$  be two  $m \times 1$  vectors such that  $\mathbf{v}$  belongs to the range of  $B^\dagger$ . Let  $\mathbf{a} = A^\dagger \mathbf{u}$  and  $\mathbf{b} = B^\dagger \mathbf{v}$ ; then we have

$$\mathbf{a}^\dagger \mathbf{a} = \mathbf{u}^\dagger A A^\dagger \mathbf{u}, \quad \mathbf{b}^\dagger \mathbf{b} = \mathbf{v}^\dagger B B^\dagger \mathbf{v}, \quad \mathbf{a}^\dagger \mathbf{b} = \mathbf{u}^\dagger A B^\dagger \mathbf{v} = \mathbf{u}^\dagger \mathbf{v}. \quad (\text{C.1})$$

The Cauchy inequality applied to the equation yields

$$(\mathbf{u}^\dagger A A^\dagger \mathbf{u})(\mathbf{v}^\dagger B B^\dagger \mathbf{v}) \geq (\mathbf{u}^\dagger \mathbf{v})^2. \quad (\text{C.2})$$

Setting  $\mathbf{v} = (B B^\dagger)^+ \mathbf{u}$  gives rise to

$$\mathbf{u}^\dagger A A^\dagger \mathbf{u} \geq \mathbf{u}^\dagger (B B^\dagger)^+ \mathbf{u}, \quad (\text{C.3})$$

which implies that  $A A^\dagger \geq (B B^\dagger)^+$ . A necessary condition for saturating the inequality is  $A^\dagger \mathbf{u} \propto B^\dagger (B B^\dagger)^+ \mathbf{u}$  for arbitrary  $\mathbf{u}$ ; that is,  $A^\dagger \propto B^\dagger (B B^\dagger)^+$  and  $A \propto (B B^\dagger)^+ B$ . Since  $A B^\dagger$  is a projector by assumption, it follows that  $A = (B B^\dagger)^+ B$ , which happens to be the pseudoinverse of  $B^\dagger$  [34]. Now the inequality is indeed saturated.

If  $A B^\dagger = 1$ , then  $(B B^\dagger)$  is invertible. The second part of the lemma follows from the fact that  $(B B^\dagger)^+ = (B B^\dagger)^{-1}$ .

# Supplementary materials about adaptive measurements

---

## D.1 Derivation of Eqs. (5.23) and (5.24)

In order to derive Eqs. (5.23) and (5.24), in this appendix we determine the minimum of  $\text{tr}(WC)$  under the constraint  $\text{tr}(J^{-1}C^{-1}) = 1$ , assuming that  $J$ ,  $C$ , and  $W$  are positive definite  $n \times n$  matrices (see also Ref. [107]).

To start with, consider the special case in which  $J$  is the identity. Choose a suitable basis such that  $W$  is diagonal; then  $\text{tr}(WC)$  depends on only the diagonal elements of  $C$ . Now  $C$  must be diagonal to attain the minimum of  $\text{tr}(WC)$ . Otherwise, its diagonal is majorized by its eigenvalues, which implies that  $\sum_{j=1}^n 1/C_{jj} < \text{tr}(C^{-1}) = 1$ . Therefore, we can construct a diagonal matrix  $C' \propto \text{diag}(C_{11}, \dots, C_{nn})$  such that  $\text{tr}(C'^{-1}) = 1$  and  $\text{tr}(WC') < \text{tr}(WC)$ , contrary to the assumption. As a consequence, the minimum of  $\text{tr}(WC)$  is equal to  $(\text{tr} \sqrt{W})^2$  and it is achieved when  $C = \text{tr}(\sqrt{W})/\sqrt{W}$ .

As for the general situation, we have

$$\begin{aligned} \min_{\text{tr}(J^{-1}C^{-1})} \text{tr}(WC) &= \min_{\text{tr}\{(J^{1/2}CJ^{1/2})^{-1}\}=1} \text{tr}\{(J^{-1/2}WJ^{-1/2})(J^{1/2}CJ^{1/2})\} \\ &= (\text{tr} \sqrt{J^{-1/2}WJ^{-1/2}})^2 = (\text{tr} \sqrt{W^{1/2}J^{-1}W^{1/2}})^2. \end{aligned} \quad (\text{D.1})$$

The minimum is attained when

$$C = J^{-1/2} \frac{\text{tr} \sqrt{J^{-1/2}WJ^{-1/2}}}{\sqrt{\text{tr} \sqrt{J^{-1/2}WJ^{-1/2}}}} J^{-1/2}. \quad (\text{D.2})$$



## D.2 Connection with pure-state estimation

In this appendix, following the analysis in Sec. 5.4.3, we discuss the characteristics of the minimal MSH and MSB, as well as the optimal estimation strategies when the true state is close to the boundary of the state space, thereby making contact with the problem of pure-state estimation. Our study shows that, in the pure-state limit, the minimal MSH converges to the value in pure-state estimation, whereas the minimal MSB does not.

According to Eq. (5.78), when  $\lambda_1 = 1/d_1$  and  $\lambda_2 = 0$ , the minimal scaled MSH is attained at  $x_{\text{SH}}^{\text{Opt}} = 1$ ,

$$\mathcal{E}_{\text{SH}}^{\text{Opt}} = \frac{(d_1 + 1)^2(d_1 - 1)}{d_1} + 2d_1d_2. \quad (\text{D.3})$$

Here the first term is identical with the minimal scaled MSH when the true state is the completely mixed state on the Hilbert space of dimension  $d_1$ ; the second term accounts for the additional uncertainty due to the increase in the dimension of the Hilbert space. According to Eq. (5.36), the GM bound is

$$\mathcal{E}_{\text{SH}}^{\text{GM}} = \frac{(d_1^2 + \sqrt{2}d_1d_2 - 1)^2}{d_1(d_1 - 1)}. \quad (\text{D.4})$$

For pure states (when  $d_1 = 1$ ), the minimal scaled MSH  $2(d - 1)$  coincides with the bound and is also equal to the minimal value in pure-state estimation [192] (see also Chapter 20 in Ref. [133]). By contrast, the minimal scaled MSH is strictly larger than the bound when  $d_1 \geq 1$ . Given  $\gamma := d_2/d_1$ , the ratio  $\mathcal{E}_{\text{SH}}^{\text{Opt}}/\mathcal{E}_{\text{SH}}^{\text{GM}}$  increases monotonically with  $d_1$  and  $d_2$  if  $d_1 \geq 3$ . In the large-dimension limit, we have

$$\lim_{d_1, d_2 \rightarrow \infty} \frac{\mathcal{E}_{\text{SH}}^{\text{Opt}}}{\mathcal{E}_{\text{SH}}^{\text{GM}}} = \frac{(\gamma + 1)(2\gamma + 1)}{(\sqrt{2}\gamma + 1)^2}. \quad (\text{D.5})$$

The maximal ratio  $(4 + 3\sqrt{2})/8$  is attained when  $\gamma = 1/\sqrt{2}$ . Further analysis shows that this value is also the global maximum of  $\mathcal{E}_{\text{SH}}^{\text{Opt}}/\mathcal{E}_{\text{SH}}^{\text{GM}}$  over states that are proportional to projectors. Therefore, the GM bound for the scaled MSH can be saturated approximately for these states.

## Appendix D. Supplementary materials about adaptive measurements

---

According to Eq. (5.78), when  $\lambda_1 = 1/d_1$  and  $\lambda_2 = 0$ ,  $\mathcal{E}_{\text{SB}}^{\text{Opt}}$  and  $x_{\text{SB}}^{\text{Opt}}$  are given by

$$\begin{aligned}\mathcal{E}_{\text{SB}}^{\text{Opt}} &= \frac{1}{4} \left[ (d_1 + 1)^2 (d_1 - 1) + d_2 \left( 2d_1 + \sqrt{d_2^2 + d_2 - 1} \right)^2 \right], \\ x_{\text{SB}}^{\text{Opt}} &= \frac{2d_1}{2d_1 + \sqrt{d_2^2 + d_2 - 1}}.\end{aligned}\tag{D.6}$$

The minimal scaled MSB is strictly larger than the GM bound  $\mathcal{E}_{\text{SB}}^{\text{GM}} = (d+1)^2(d-1)/4$  (see Sec. 5.3.3). In the pure-state limit, it is strictly larger than the minimal value in pure-state estimation. For given  $\gamma$ , numerical calculation suggests that  $\mathcal{E}_{\text{SB}}^{\text{Opt}}/\mathcal{E}_{\text{SB}}^{\text{GM}}$  increases monotonically with  $d_1$  and  $d_2$ . In the large-dimension limit, we have

$$\lim_{d_1, d_2 \rightarrow \infty} \frac{\mathcal{E}_{\text{SB}}^{\text{Opt}}}{\mathcal{E}_{\text{SB}}^{\text{GM}}} = 1 + \frac{\gamma}{(1 + \gamma)^2};\tag{D.7}$$

the maximal ratio  $\frac{5}{4}$  is attained when  $\gamma = 1$ .

Now consider the family of states in Eq. (5.38). When  $1 - s$  is small, we have

$$\begin{aligned}\mathcal{E}_{\text{SH}}^{\text{Opt}}(s) &\approx \mathcal{E}_{\text{SH}}^{\text{Opt}}(s = 1) + 4d_2 \sqrt{\frac{1 + d_1(d_2^2 + d_2 - 1)}{2(d_1 + d_2)}} \sqrt{1 - s}, \\ x_{\text{SH}}^{\text{Opt}}(s) &\approx 1 - \frac{1}{d_1} \sqrt{\frac{1 + d_1(d_2^2 + d_2 - 1)}{2(d_1 + d_2)}} \sqrt{1 - s},\end{aligned}\tag{D.8}$$

both of which vary rapidly as  $s$  decreases. As for the scaled MSB, we have

$$\mathcal{E}_{\text{SB}}^{\text{Opt}}(s) \approx \mathcal{E}_{\text{SB}}^{\text{Opt}}(s = 1) - b(1 - s), \quad x_{\text{SB}}^{\text{Opt}}(s) \approx x_{\text{SB}}^{\text{Opt}}(s = 1) + \frac{u}{v}(1 - s),\tag{D.9}$$

both of which vary smoothly. Here

$$\begin{aligned}b &= \frac{d_1^2 d_2 [(3d_1^2 + d_2^2 + d_2 - d_1)c + 4d_1 c^{3/2} - 4d_1 - 4c^{1/2}]}{2c(d_1 + d_2)(2d_1 + c^{1/2})}, \\ u &= d_1 [-(d_1 + 1)^2 (d_1 - 1) c^{3/2} + c^{3/2} (d_2 - d_1) (c + 4d_1^2 + 4d_1 c^{1/2}) \\ &\quad + (d_2 + 1)^2 (d_2 - 1) (2d_1 + c^{1/2})^2 (4d_1 + c^{1/2})], \\ v &= c(d_1 + d_2) [16d_1^4 + 32d_1^3 c^{1/2} + 24d_1^2 c + 8d_1 c^{3/2} + c^2], \\ c &= d_2^2 + d_2 - 1.\end{aligned}\tag{D.10}$$

### D.3. Discontinuity of the minimal scaled MSB

---

In the opposite extreme  $s = 0$ , that is, when the true state is completely mixed, both the minimal scaled MSH and the minimal scaled MSB are achieved when  $x = d/(d+1)$ , and they saturate the GM bounds. When  $s$  is very small, we have

$$\begin{aligned}\mathcal{E}_{\text{SH}}^{\text{Opt}}(s) &\approx \frac{(d+1)^2(d-1)}{d} - \frac{d_2(d^3+d+2)}{d_1d(d+2)}s^2, \\ x_{\text{SH}}^{\text{Opt}}(s) &\approx \frac{d}{d+1} - \begin{cases} \frac{d(d_2-d_1)}{d_1(d+1)(d+2)}s & \text{if } d_2 \neq d_1, \\ \frac{d_1(2d_1^2-5d_1-1)}{2(d_1+1)(2d_1+1)^2}s^2 & \text{if } d_2 = d_1, \end{cases}\end{aligned}\quad (\text{D.11})$$

$$\begin{aligned}\mathcal{E}_{\text{SB}}^{\text{Opt}}(s) &\approx \frac{1}{4}(d+1)^2(d-1) + \frac{d_2d^2(d-2)(d+1)}{8d_1(d+2)}s^2, \\ x_{\text{SB}}^{\text{Opt}}(s) &\approx \frac{d}{d+1} - \begin{cases} \frac{d(d_2-d_1)}{d_1(d+1)(d+2)}s & \text{if } d_2 \neq d_1, \\ \frac{d_1(d_1-1)(5d_1+2)}{2(d_1+1)(2d_1+1)^2}s^2 & \text{if } d_2 = d_1. \end{cases}\end{aligned}\quad (\text{D.12})$$

In contrast to the scenario in the pure-state limit, all these quantities vary smoothly with  $s$ . More specifically,  $\mathcal{E}_{\text{SH}}^{\text{Opt}}(s)$  decreases slowly with  $s$ , whereas  $\mathcal{E}_{\text{SB}}^{\text{Opt}}(s)$  increases slowly with  $s$ .

### D.3 Discontinuity of the minimal scaled MSB at the boundary of the state space

In this appendix, we reveal an exotic feature of state estimation that emerges when the dimension of the Hilbert space is larger than two: the discontinuity of the minimal scaled MSB at the boundary of the state space. This feature reflects higher complexity and richer structure of the state-estimation problem beyond the qubit setting. The implications of this discontinuity are also discussed briefly.

Consider a state with the three distinct eigenvalues

$$\frac{1-\eta}{d_1}, \quad \eta\left(\frac{s}{d_2} + \frac{1-s}{d_2+d_3}\right), \quad \frac{\eta(1-s)}{d_2+d_3}, \quad (\text{D.13})$$

whose multiplicities are  $d_1, d_2, d_3$ , respectively. When  $s = 0$ , in the limit  $\eta \rightarrow 0$ , the minimal scaled MSB is given by Eq. (D.6) with  $d_2$  replaced by  $d_2 + d_3$ . When  $s > 0$ ,

## Appendix D. Supplementary materials about adaptive measurements

---

assuming that Conjecture 5.2 holds, then the Fisher information polytope has five extremal points, which correspond to the five partitions  $\{\{1\}, \{2, 3\}\}$ ,  $\{\{2\}, \{1, 3\}\}$ ,  $\{\{3\}, \{1, 2\}\}$ ,  $\{\{1\}, \{2\}, \{3\}\}$ , and  $\{\{1, 2, 3\}\}$ , respectively. To simplify the notation, we use the five numbers 1, 2, 3, 4 and 5 to represent the five partitions. In the limit  $\eta \rightarrow 0$  and then  $s \rightarrow 1$ , the cost function of the scaled MSB turns out to be

$$\begin{aligned} \mathcal{E}_{\text{SB}}(x) = & \frac{1}{4} \left[ d_1(d_1^2 + d_1 - 1) + \frac{d_2(d_2^2 + d_2 - 1)}{x_1 + x_2 + x_4} + \frac{d_3(d_3^2 + d_3 - 1)}{x_3 + x_4} + \frac{4d_1^2 d_2}{x_3 + x_5} \right. \\ & \left. + \frac{4d_1^2 d_3}{x_2 + x_5} + \frac{4d_2^2 d_3}{x_1} - 1 \right]. \end{aligned} \quad (\text{D.14})$$

The minimum under the constraint  $x_2 = x_4 = 0$  is given by

$$\begin{aligned} \mathcal{E}_{\text{SB}} = & \frac{1}{4} [d_1(d_1^2 + d_1 - 1) + (\sqrt{u} + \sqrt{w})^2 - 1], \\ x_1 = & \frac{\sqrt{u}}{\sqrt{w} + \sqrt{u}}, \quad x_3 = \frac{\sqrt{d_3^2 + d_3 - 1}}{v} \frac{\sqrt{w}}{\sqrt{w} + \sqrt{u}}, \quad x_5 = \frac{2d_1}{v} \frac{\sqrt{w}}{\sqrt{w} + \sqrt{u}}, \end{aligned} \quad (\text{D.15})$$

where

$$u = d_2(d_2^2 + d_2 - 1) + 4d_2^2 d_3, \quad v = 2d_1 + \sqrt{d_3^2 + d_3 - 1}, \quad w = d_3 v^2 + 4d_1^2 d_2. \quad (\text{D.16})$$

When  $d_1 = d_2 = d_3$ , numerical calculation shows that this constrained minimum of  $\mathcal{E}_{\text{SB}}(x)$  is actually a global minimum. In the special case  $d_1 = d_2 = d_3 = 1$ , we have

$$\begin{aligned} \mathcal{E}_{\text{SB}} = & \frac{1}{2} (9 + \sqrt{65}), \\ x_1 = & \frac{\sqrt{5}}{\sqrt{5} + \sqrt{13}}, \quad x_3 = \frac{\sqrt{13}}{3(\sqrt{5} + \sqrt{13})}, \quad x_5 = \frac{2\sqrt{13}}{3(\sqrt{5} + \sqrt{13})}. \end{aligned} \quad (\text{D.17})$$

Compared with the minimal scaled MSB  $(9 + 4\sqrt{5})/2$  corresponding to  $s = 0$  [see Eq. (D.6)], the one corresponding to  $s = 1$  is smaller, which indicates that the minimal scaled MSB is not continuous in the pure-state limit. Note that his conclusion is independent of the validity of Conjecture 5.2. The left plot of Fig. D.1 shows that the minimal scaled MSB can assume any value between  $(9 + \sqrt{65})/2$  and  $(9 + 4\sqrt{5})/2$  depending on the way the limit is taken. Similar analysis also applies to states in higher

### D.3. Discontinuity of the minimal scaled MSB

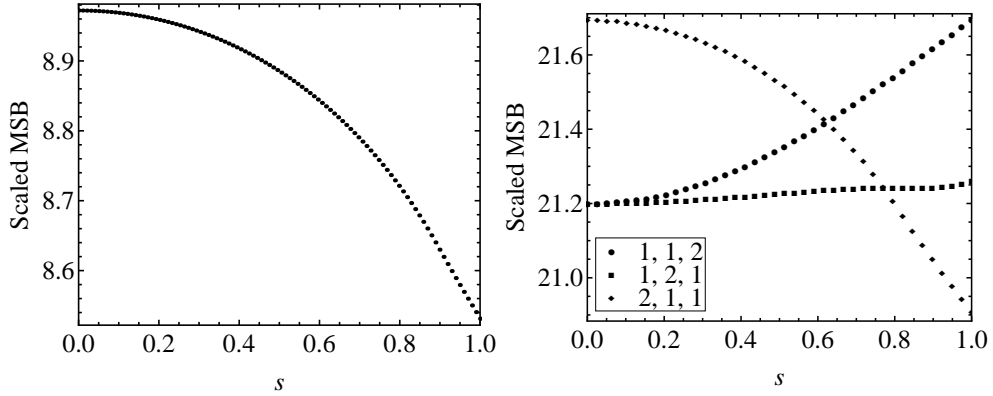


Figure D.1: Minimal scaled MSB for states determined by Eq. (D.13) in the limit  $\eta \rightarrow 0$  for four different sets of values of  $d_1, d_2, d_3$ : 1, 1, 1 (left plot); 1, 1, 2; 1, 2, 1; and 2, 1, 1 (right plot). The figure implies that the minimal scaled MSB is not continuous at the boundary of the state space.

dimensions and to states of larger ranks (but smaller than  $d - 1$ ). Nevertheless, the dependence of the minimal scaled MSB on  $s$  can be qualitatively different, as illustrated in the right plot of Fig. D.1.

The discontinuity of the minimal scaled MSB is closely related to the divergence of the weight matrix  $J/4$  at the boundary of the state space, which reflects a paranoid requirement on the precision in estimating certain parameters. It can be eliminated by replacing  $J$  with  $J^\alpha$  for  $0 \leq \alpha < 1$ . In that case, the minimal scaled WMSE and the optimal estimation strategy can vary rapidly near the boundary of the state space as if the scaled MSH is the figure of merit (see Sec. 5.4.3 and Appendix D.2). We emphasize that the limit  $\alpha \rightarrow 1$  and the pure-state limit do not commute.

It should be noted that the conclusion on the discontinuity of the minimal scaled MSB is applicable only in the asymptotic limit. In practice, the discontinuity does not appear since the sample size is always finite. Nevertheless, the optimal measurement scheme and the minimal error heavily depend on the true state and the figure of merit when the true state is close to the boundary of the state space. One root of this phenomenon is the inevitable information trade-off among noncommutative observables. This problem poses a serious challenge to obtaining reliable and efficient estimators in practice. On the other hand, it motivates the study of collective measurements, the most general measurements in quantum mechanics, which are the focus of Chapter 6.

# Technical details about Chapter 6

---

## E.1 Proof of Eq. (6.49) for Slater-determinant states

Any  $N$ -partite Slater-determinant state can be written as

$$|\Psi\rangle = U^{\otimes N}(|1\rangle \wedge \cdots \wedge |N\rangle) = \sum_{k_1, \dots, k_N} a_{k_1, \dots, k_N} |k_1, \dots, k_N\rangle, \quad (\text{E.1})$$

where

$$a_{k_1, \dots, k_N} = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} \text{sgn}(\sigma) U_{k_1 \sigma(1)} \cdots U_{k_N \sigma(N)} \quad (\text{E.2})$$

satisfy the equation

$$a_{j_1 j_2 \dots j_N} a_{m k_2 \dots k_N} - \sum_{l=1}^N a_{j_1 \dots j_{l-1} m j_{l+1} \dots j_N} a_{j_l k_2 \dots k_N} = 0 \quad (\text{E.3})$$

thanks to the permutation symmetry. Now the proof of Eq. (6.49) proceeds as follows:

$$\begin{aligned} \text{tr}\{\Lambda_N(\rho)(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} &= N^2 \text{tr}\{V(1, N+1)(1 \otimes \rho^{\otimes(2N-1)})(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} \\ &= N^2 \sum_{\substack{m, j_2, \dots, j_N \\ n, k_2, \dots, k_N}} \lambda_m \lambda_{j_2} \cdots \lambda_{j_N} \lambda_{k_2} \cdots \lambda_{k_N} a_{m j_2 \dots j_N}^* a_{n k_2 \dots k_N}^* a_{n j_2 \dots j_N} a_{m k_2 \dots k_N} \\ &= N \sum_{\substack{m, j_2, \dots, j_N \\ n, k_2, \dots, k_N}} \left[ \lambda_m \lambda_{j_2} \cdots \lambda_{j_N} \lambda_{k_2} \cdots \lambda_{k_N} a_{m j_2 \dots j_N}^* a_{n k_2 \dots k_N}^* \right. \\ &\quad \left. \times (a_{n j_2 \dots j_N} a_{m k_2 \dots k_N} - a_{n m \dots j_N} a_{j_2 k_2 \dots k_N} - \cdots - a_{n j_2 \dots m} a_{j_N k_2 \dots k_N}) \right] \\ &= N \sum_{\substack{m, j_2, \dots, j_N \\ n, k_2, \dots, k_N}} \lambda_m \lambda_{j_2} \cdots \lambda_{j_N} \lambda_{k_2} \cdots \lambda_{k_N} |a_{m j_2 \dots j_N}|^2 |a_{n k_2 \dots k_N}|^2 \\ &= N \langle \Psi | \rho^{\otimes N} | \Psi \rangle \langle \Psi | (1 \otimes \rho^{\otimes(N-1)}) | \Psi \rangle. \end{aligned} \quad (\text{E.4})$$

## E.2 Proof of Conjecture 6.4 for symmetric and bipartite antisymmetric subspaces

**Lemma E.1** *Any  $N$ -partite symmetric state  $|\Psi\rangle$  satisfies the inequality*

$$\mathrm{tr}\{\Lambda_N(\rho)(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} \leq N^2\langle\Psi|(1 \otimes \rho^{\otimes(N-1)})|\Psi\rangle\langle\Psi|\rho^{\otimes N}|\Psi\rangle; \quad (\text{E.5})$$

*the inequality is saturated if and only if  $|\Psi\rangle$  is a tensor power of a single-particle state.*

The inequality in the lemma is equivalent to

$$\mathrm{tr}\{V(1, 2N)(1 \otimes \rho^{\otimes(2N-1)})(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} \leq \mathrm{tr}\{(1 \otimes \rho^{\otimes(2N-1)})(|\Psi\rangle\langle\Psi|)^{\otimes 2}\}. \quad (\text{E.6})$$

Suppose that  $|\Psi\rangle = \sum_{j_1, \dots, j_N} a_{j_1 \dots j_N} |j_1, \dots, j_N\rangle$ , where the coefficients  $a_{j_1 \dots j_N}$  are invariant under permutations of the indices. Then we have

$$\begin{aligned} & \mathrm{tr}\{V(1, 2N)(1 \otimes \rho^{\otimes(2N-1)})(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} \\ &= \sum_{j_1, k_1, \dots, j_N, k_N} \lambda_{j_2} \cdots \lambda_{j_N} \lambda_{k_1} \cdots \lambda_{k_N} a_{k_N j_2 \dots j_N}^* a_{k_1 \dots k_{N-1} j_1} a_{j_1 \dots j_N} a_{k_1 \dots k_N} \\ &= \frac{1}{2} \sum_{j_1, k_1, \dots, j_N, k_N} \lambda_{j_2} \cdots \lambda_{j_N} \lambda_{k_1} \cdots \lambda_{k_N} (a_{k_N j_2 \dots j_N}^* a_{k_1 \dots k_{N-1} j_1} a_{j_1 \dots j_N} a_{k_1 \dots k_N} \\ & \quad + a_{k_N k_2 \dots k_{N-1} k_1}^* a_{j_N j_2 \dots j_{N-1} j_1}^* a_{j_1 k_2 \dots k_{N-1} k_1} a_{j_N j_2 \dots j_{N-1} k_N}) \\ &= \frac{1}{2} \sum_{j_1, k_1, \dots, j_N, k_N} \lambda_{j_2} \cdots \lambda_{j_N} \lambda_{k_1} \cdots \lambda_{k_N} (a_{k_N j_2 \dots j_N}^* a_{k_1 \dots k_{N-1} j_1}^* a_{j_1 \dots j_N} a_{k_1 \dots k_N} \\ & \quad + a_{j_1 \dots j_N}^* a_{k_1 \dots k_N}^* a_{k_N j_2 \dots j_N} a_{k_1 \dots k_{N-1} j_1}) \\ &\leq \frac{1}{2} \sum_{j_1, k_1, \dots, j_N, k_N} \lambda_{j_2} \cdots \lambda_{j_N} \lambda_{k_1} \cdots \lambda_{k_N} (|a_{k_N j_2 \dots j_N} a_{k_1 \dots k_{N-1} j_1}|^2 + |a_{j_1 \dots j_N} a_{k_1 \dots k_N}|^2) \\ &= \sum_{j_1, k_1, \dots, j_N, k_N} \lambda_{j_2} \cdots \lambda_{j_N} \lambda_{k_1} \cdots \lambda_{k_N} |a_{j_1 \dots j_N} a_{k_1 \dots k_N}|^2 \\ &= \mathrm{tr}\{(1 \otimes \rho^{\otimes(2N-1)})(|\Psi\rangle\langle\Psi|)^{\otimes 2}\}. \end{aligned} \quad (\text{E.7})$$

Obviously, the inequality is saturated when  $|\Psi\rangle$  is a tensor power of a single-particle state. On the other hand, suppose that the inequality is saturated at  $|\Psi\rangle$  for some

state  $\rho$  of full rank; then the same holds when  $\rho$  is completely mixed. Therefore,

$$\mathrm{tr}\{V(1, 2N)(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} = \mathrm{tr}\{(|\Psi\rangle\langle\Psi|)^{\otimes 2}\}, \quad (\text{E.8})$$

which implies that  $|\Psi\rangle^{\otimes 2}$  is invariant under any permutation of the  $2N$  parties, given that  $|\Psi\rangle$  is symmetric. This possibility can happen if and only if  $|\Psi\rangle$  is a tensor power of a single-particle state.

**Lemma E.2** *Any bipartite antisymmetric state  $|\Psi\rangle$  satisfies the inequality*

$$2 \mathrm{tr}\{V(2, 3)(\rho^{\otimes 2} \otimes 1 \otimes \rho)(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} \leq \mathrm{tr}\{(\rho^{\otimes 2} \otimes 1 \otimes \rho)(|\Psi\rangle\langle\Psi|)^{\otimes 2}\}, \quad (\text{E.9})$$

*and the inequality is saturated if and only if  $|\Psi\rangle$  is a Slater-determinant state.*

Suppose that  $|\Psi\rangle = \sum_{jk} a_{jk}|jk\rangle$  with  $a_{jk} = -a_{kj}$ . Then we have

$$\begin{aligned} & \mathrm{tr}\{(\rho^{\otimes 2} \otimes 1 \otimes \rho)(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} - 2 \mathrm{tr}\{V(2, 3)(\rho^{\otimes 2} \otimes 1 \otimes \rho)(|\Psi\rangle\langle\Psi|)^{\otimes 2}\} \\ &= \sum_{jkmn} \lambda_j \lambda_k \lambda_n a_{jm}^* a_{kn}^* a_{jm} a_{kn} - 2 \sum_{jkmn} \lambda_j \lambda_k \lambda_n a_{jm}^* a_{kn}^* a_{jk} a_{mn} \\ &= \sum_{jkmn} \lambda_j \lambda_k \lambda_n a_{jm}^* a_{kn}^* (a_{jm} a_{kn} - a_{jk} a_{mn} + a_{jn} a_{mk}) \\ &= \frac{1}{3} \sum_{jkmn} \lambda_j \lambda_k \lambda_n \left[ a_{jm}^* a_{kn}^* (a_{jm} a_{kn} - a_{jk} a_{mn} + a_{jn} a_{mk}) \right. \\ & \quad \left. + a_{km}^* a_{jn}^* (a_{km} a_{jn} - a_{kj} a_{mn} + a_{kn} a_{mj}) \right. \\ & \quad \left. + a_{nm}^* a_{kj}^* (a_{nm} a_{kj} - a_{nk} a_{mj} + a_{nj} a_{mk}) \right] \\ &= \frac{1}{3} \sum_{jkmn} \lambda_j \lambda_k \lambda_n |a_{jm} a_{kn} - a_{jk} a_{mn} - a_{jn} a_{km}|^2 \geq 0. \end{aligned} \quad (\text{E.10})$$

If  $|\Psi\rangle$  is a Slater-determinant state, then the inequality is saturated according to Eq. (E.4). Suppose, on the other hand, that the inequality is saturated at  $|\Psi\rangle$  for some state  $\rho$  of full rank; then it is saturated at  $|\Psi\rangle$  when  $\rho$  is the completely mixed state, which implies that each reduced state of  $|\Psi\rangle$  has purity  $\frac{1}{2}$ . Therefore,  $|\Psi\rangle$  is a Slater-determinant state.



### E.3 Proof of Theorem 6.7

When  $N$  is sufficiently large, the major contribution to  $t^N(1/d)$  stems from irreducible components with  $|\mu_j - N/d| \sim \sqrt{N}$  and  $|\mu_j - \mu_k| \sim \sqrt{N}$ . In that case, the expressions of  $d_\mu = N!/h_\mu$  and  $D_\mu = y_\mu/h_\mu$  can be simplified by means of the approximation

$$\begin{aligned} h_\mu &= \frac{\prod_{j=1}^d (d + \mu_j - j)!}{\prod_{j < k} (\mu_j - \mu_k + k - j)} \approx \left( \prod_j \mu_j! \right) \left( \prod_{j < k} \frac{N}{d} \frac{1}{\mu_j - \mu_k} \right), \\ y_\mu &= \prod_{j=1}^d \frac{(d + \mu_j - j)!}{(d - j)!} \approx \left( \frac{N}{d} \right)^{d(d-1)/2} \left( \prod_{j=1}^{d-1} \frac{1}{j!} \right) \left( \prod_j \mu_j! \right). \end{aligned} \quad (\text{E.11})$$

$$\begin{aligned} \frac{t^N(1/d)}{N} &= \frac{d}{Nd^N} \sum_{\mu} d_\mu D_\mu \left( \mu^2 - \frac{N^2}{d} \right) \\ &\approx \frac{1}{\prod_{j=1}^{d-1} j!} \sum_{\substack{\mu_1 \geq \mu_2 \geq \dots \geq \mu_d \\ |\mu| = N}} \left\{ \frac{N!}{d^N \prod_j \mu_j!} \left[ \prod_{j < k} \frac{d}{N} (\mu_j - \mu_k)^2 \right] \sum_j \frac{d}{N} \left( \mu_j - \frac{N}{d} \right)^2 \right\} \\ &\approx \frac{1}{\prod_{j=1}^d j!} \sum_{|\mu| = N} \left\{ \frac{N!}{d^N \prod_j \mu_j!} \left[ \prod_{j < k} \frac{d}{N} (\mu_j - \mu_k)^2 \right] \sum_j \frac{d}{N} \left( \mu_j - \frac{N}{d} \right)^2 \right\} \\ &\approx \frac{1}{\prod_{j=1}^d j!} \sum_n p_N(n) \sum_{|\mu| = n} \left\{ \frac{n!}{d^n \prod_j \mu_j!} \left[ \prod_{j < k} \frac{d}{N} (\mu_j - \mu_k)^2 \right] \frac{d}{N} \sum_j \left( \mu_j - \frac{n}{d} \right)^2 \right\}, \end{aligned} \quad (\text{E.12})$$

where  $p_N(\cdot)$  is a probability distribution of Gaussian shape centered at  $N$  and with variance  $N/d$ . In the large- $N$  limit, the summation may be replaced by the integral

$$\int \left( \prod_j dx_j \right) \Phi(x) \sum_j \left( x_j - \frac{1}{d} \sum_k x_k \right)^2, \quad (\text{E.13})$$

where  $x_j = \sqrt{d/N}(\mu_j - N/d)$  and the probability distribution

$$\Phi(x) = \frac{1}{(2\pi)^{d/2} \prod_{j=1}^d j!} \exp\left(-\frac{1}{2} \sum_j x_j^2\right) \prod_{j < k} (x_j - x_k)^2 \quad (\text{E.14})$$

satisfies

$$\int \left( \prod_j dx_j \right) \Phi(x) x_k x_l = (d+1)\delta_{kl} - 1, \quad (\text{E.15})$$

according to Eqs. (17.6.7), (17.8.5) and (17.8.8) in Ref. [196]. Therefore, the values in Eqs. (E.12) and (E.13) are both equal to  $d^2 - 1$ , and the theorem follows.

## E.4 Proof of Theorem 6.8

Theorem 6.8 follows from Theorem 6.7 when  $r = 0$  and from Eq. (6.90) in the pure-state limit. When  $0 < r < 1$  and  $N$  is large, the main contribution to  $I^N(r)$  stems from irreducible components with  $[\mu_1, \mu_2] \approx [N(1+r)/2, N(1-r)/2]$ . In that case, the Fisher information matrix  $I_\mu$  is diagonal and  $I_{\mu,11} = I_{\mu,22}$ . According to Eq. (6.83),

$$\begin{aligned}
 I_{\mu,11}(r) &\approx f_{\mu,1}(r) \approx \frac{q}{2r^2} \left(\frac{1+r}{2}\right)^{\mu_1} \left(\frac{1-r}{2}\right)^{\mu_2}, \\
 I_{\mu,33}(r) &\approx f_{\mu,3}(r) \approx \frac{1}{4r^3} \left(\frac{1+r}{2}\right)^{\mu_1} \left(\frac{1-r}{2}\right)^{\mu_2} \\
 &\quad \times \left[ \left(\frac{\mu_2}{|\rho|} - 2N\right)^2 \frac{1+r}{2} + 2(q+1) \left(\frac{\mu_2}{|\rho|} - 2N\right) + \frac{2(q^2+2q)}{1+r} \right] \\
 &= \frac{1}{4r^3} \left(\frac{1+r}{2}\right)^{\mu_1} \left(\frac{1-r}{2}\right)^{\mu_2} \\
 &\quad \times \left[ \frac{2r^2}{(1-r)|\rho|} \left(\mu_2 - \frac{1-r}{2}N\right)^2 + 2\left(\frac{\mu_2}{|\rho|} - 2N\right) + \frac{4q}{1+r} \right].
 \end{aligned} \tag{E.16}$$

These approximations become exact in the large- $N$  limit. Therefore,

$$\begin{aligned}
 \lim_{N \rightarrow \infty} \frac{I_{11}^N(r)}{N} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{\mu} d_{\mu} I_{\mu,11}(r) \\
 &= \lim_{N \rightarrow \infty} \frac{1}{2r^2} \sum_{\substack{\mu_1 + \mu_2 = N \\ \mu_2 \leq \mu_1}} \frac{N! q(q+1)}{N(\mu_1+1)! \mu_2!} \left(\frac{1+r}{2}\right)^{\mu_1} \left(\frac{1-r}{2}\right)^{\mu_2} \\
 &= \lim_{N \rightarrow \infty} \sum_{\mu_1 + \mu_2 = N} \binom{N}{\mu_1} \left(\frac{1+r}{2}\right)^{\mu_1} \left(\frac{1-r}{2}\right)^{\mu_2} \frac{q(q+1)}{2N(\mu_1+1)r^2} = \frac{1}{1+r}, \\
 \lim_{N \rightarrow \infty} \frac{I_{33}^N(r)}{N} &= \lim_{N \rightarrow \infty} \sum_{\mu_1 + \mu_2 = N} \binom{N}{\mu_1} \left(\frac{1+r}{2}\right)^{\mu_1} \left(\frac{1-r}{2}\right)^{\mu_2} \frac{q+1}{4(\mu_1+1)r^3} \\
 &\quad \times \frac{1}{N} \left[ \frac{2r^2}{(1-r)|\rho|} \left(\mu_2 - \frac{1-r}{2}N\right)^2 + 2\left(\frac{\mu_2}{|\rho|} - 2N\right) + \frac{4q}{1+r} \right] \\
 &= \frac{1}{2r^2(1+r)} \left[ \frac{2r^2}{(1-r)} + 4\left(\frac{1}{1+r} - 1\right) + \frac{4r}{1+r} \right] = \frac{1}{1-r^2},
 \end{aligned} \tag{E.17}$$

where we note that the binomial distribution is highly peaked in the large- $N$  limit.

# Nice error bases

---

A *nice error basis* [166, 168, 169] is a set of  $d \times d$  unitary matrices  $\{M(h)|h \in H\}$  parameterized by the elements of an order- $d^2$  *index group*  $H$ , such that

1.  $M(e)$  is the identity matrix;
2.  $\text{tr}\{M(h)\} = d\delta_{h,e}$ ;
3.  $M(g)M(h) \propto M(gh)$ ;

where  $e$  is the identity of  $H$ . Nice error bases were originally introduced by Knill [169] in the study of quantum codes and quantum computation. Later, they were investigated in more detail by Klappenecker and Rötteler [166, 168], who cataloged all nice error bases with non-Abelian index groups up to dimension 10. In general, a set of  $d^2$  unitary matrices parameterized by the group  $H$  is a nice error basis if and only if it forms a faithful irreducible projective representation of  $H$  [166, 169].

Let  $G$  be the group generated by a nice error basis parameterized by  $H$ ; then its collineation group  $\overline{G}$  is isomorphic to  $H$ . Since  $G$  or  $\overline{G}$  determines the nice error basis up to overall phase factors, it is also called a nice error basis when there is no confusion.

Two nice error bases  $\overline{G}_1$  and  $\overline{G}_2$  are (unitarily) *equivalent* if there exists a unitary transformation that maps all elements of  $\overline{G}_1$  to that of  $\overline{G}_2$ . Note that this equivalence criterion is less restrictive than the one of Klappenecker and Rötteler [166]. Our definition is more suitable if one is concerned with nice error bases as a whole rather than projective representations of the index groups.

## $p$ -groups and Sylow's theorem

---

Let  $p$  be a prime. A nontrivial finite group is a  $p$ -group if its order is a power of  $p$  [172]. A  $p$ -group is endowed with the following basic properties [69, 172]:

1. Any  $p$ -group has a nontrivial center.
2. The normalizer of a proper subgroup of a  $p$ -group is strictly larger than the subgroup.
3. Any irreducible representation of a  $p$ -group is monomial.

Suppose  $G$  is a finite group whose order  $|G|$  is divisible by  $p^k$  but not by  $p^{k+1}$ , where  $k$  is a positive integer. Then a *Sylow  $p$ -subgroup* of  $G$  is any subgroup of  $G$  whose order is  $p^k$  [172]. The significance of Sylow  $p$ -subgroups is summarized in Sylow's theorem, which is one of the most profound theorems in finite group theory [172].

**Theorem G.1** (Sylow) *Suppose  $G$  is a finite group whose order  $|G|$  is divisible by  $p^k$  but not by  $p^{k+1}$ . Then the following statements hold.*

1. *There exists a Sylow  $p$ -subgroup  $P$  of  $G$ .*
2. *The number of Sylow  $p$ -subgroups divides  $|G|/|P|$  and is equal to  $1 \pmod{p}$ .*
3. *All Sylow  $p$ -subgroups of  $G$  are conjugated to each other.*
4. *Any  $p$ -subgroup of  $G$  is a subgroup of at least one of the Sylow  $p$ -subgroups.*

# Supplementary information about the Clifford group

---

In this appendix, we offer several additional results on the Clifford group. First, we derive a simple formula for computing the absolute square of the trace of a Clifford operator, which is crucial to investigating the symmetry group of an HW covariant SIC POM and the nice error bases in the symmetry group (see Chapter 8 and Sec. 10.3.3). Second, we determine the normalizer of the Clifford group, which is indispensable to understanding the regrouping phenomena of SIC POMs [119, 278, 283] (see Sec. 9.2.3). Third, in the case of a prime dimension, we figure out all subgroups in the Clifford group that are unitarily equivalent to the HW group, which are helpful in appreciating the peculiarity of HW covariant SIC POMs in dimension 3 [279] (see Sec. 8.5).

## H.1 Trace of a Clifford unitary operator

The traces of Clifford operators play a crucial role in determining the symmetry group of an HW covariant SIC POM and the additional nice error bases in the symmetry group, but it is usually quite tedious to compute them. Fortunately, in most cases, it suffices to know their absolute values, for which purpose we can offer a simple recipe.

Let  $M$  be a  $2 \times 2$  matrix over  $\mathbb{Z}_d$ . The *kernel* and *range* of  $M$  are defined as follows:

$$\ker(M) := \{\chi | \chi \in \mathbb{Z}_d^2 \text{ and } M\chi = \mathbf{0}\}, \quad \text{range}(M) := \{M\chi | \chi \in \mathbb{Z}_d^2\}. \quad (\text{H.1})$$

By definition,  $\ker(M)$  and  $\text{range}(M)$  are subgroups of  $\mathbb{Z}_d^2$ , and the product of their

---

## Appendix H. Supplementary information about the Clifford group

---

orders is  $d^2$ . Let  $[F, \chi]$  and  $[F_1, \chi_1]$  be two Clifford operations in dimension  $d$  such that  $F$  and  $F_1$  commute; then we have

$$[F_1, \chi_1] \circ [F, \chi] \circ [F_1, \chi_1]^{-1} = [F, (\mathbf{1} - F)\chi_1 + F_1\chi]. \quad (\text{H.2})$$

Hence,  $[F, \chi]$  is conjugated to  $[F, \mathbf{0}]$  when  $\chi \in \text{range}(\mathbf{1} - \tilde{F})$ , where  $\tilde{F} := F \pmod{d}$ . For example,  $[F, \chi]$  is always conjugated to  $[F, \mathbf{0}]$  when  $(\mathbf{1} - F)$  is invertible. When  $d$  is odd, the condition  $\chi \in \text{range}(\mathbf{1} - \tilde{F})$  is also necessary for  $[F, \chi]$  to be conjugated to  $[F, \mathbf{0}]$ . Otherwise, it is not. To illustrate, let  $F = \begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix}$  and  $\chi = \begin{pmatrix} d/2 \\ 0 \end{pmatrix}$ ; then  $[F, \chi]$  is conjugated to  $[F, \mathbf{0}]$ , although  $\text{range}(\mathbf{1} - \tilde{F}) = \{\mathbf{0}\}$ . This complication arises because a Clifford operation has several different representations when  $d$  is even (see Sec. 7.3.3).

Our main interest in  $\ker(\mathbf{1} - \tilde{F})$  and  $\text{range}(\mathbf{1} - \tilde{F})$  is motivated by their connection with the absolute squares of the traces of certain Clifford operations.

**Theorem H.1** *Suppose  $[F, \chi]$  is a Clifford operation in dimension  $d$ . If  $d$  is odd, or if  $d$  is even and  $\begin{pmatrix} sd/2 \\ td/2 \end{pmatrix} \in \text{range}(\mathbf{1} - \tilde{F})$  for  $s, t = 0, 1$ , where  $\tilde{F} = F \pmod{d}$ , then*

$$|\text{tr}([F, \chi])|^2 = \begin{cases} |\ker(\mathbf{1} - \tilde{F})| & \text{if } \chi \in \text{range}(\mathbf{1} - \tilde{F}), \\ 0 & \text{otherwise.} \end{cases} \quad (\text{H.3})$$

*Otherwise, there exists  $F' \in \text{SL}(2, \mathbb{Z}_{\bar{d}})$  such that  $F' = F \pmod{d}$  and*

$$|\text{tr}([F', \chi])|^2 = \begin{cases} |\ker(\mathbf{1} - \tilde{F})| & \text{if } \chi \in \text{range}(\mathbf{1} - \tilde{F}), \\ 0 & \text{otherwise.} \end{cases} \quad (\text{H.4})$$

Here  $|\text{tr}([F, \chi])|$  denotes  $|\text{tr}(U)|$  for any  $U \in [F, \chi]$ .

**Proof.** Represent  $D_{\mathbf{k}}$  with a column vector as in the proof of Theorem 7.1 and let  $P$  be the matrix formed by juxtaposing the column representations of the  $D_{\mathbf{k}}$ s. Then the conjugation by  $U \in [F, \chi]$  can be written as multiplication by  $U \otimes U^*$ ,

$$(U \otimes U^*)P = PM_{F, \chi}, \quad (\text{H.5})$$

where  $M_{F, \chi}$  is a monomial matrix determined by  $F$  and  $\chi$ . Since the  $D_{\mathbf{k}}$ s span the

## H.1. Trace of a Clifford unitary operator

---

matrix space, the matrix  $P$  is invertible. As a consequence,

$$|\mathrm{tr}([F, \chi])|^2 = \mathrm{tr} M_{F, \chi} = \sum_{\mathbf{k} \in \ker(\mathbf{1} - \tilde{F})} \omega^{\langle \chi, F\mathbf{k} \rangle} \tau^{\langle \mathbf{k}, (\mathbf{1} - F)\mathbf{k} \rangle}, \quad (\text{H.6})$$

where in deriving the second equality we have applied Eqs. (7.9) and (7.15). Setting  $\chi = \mathbf{0}$  yields

$$|\mathrm{tr}([F, \mathbf{0}])|^2 = \sum_{\mathbf{k} \in \ker(\mathbf{1} - \tilde{F})} \tau^{\langle \mathbf{k}, (\mathbf{1} - F)\mathbf{k} \rangle}. \quad (\text{H.7})$$

When  $d$  is odd,  $\tau^{\langle \mathbf{k}, (\mathbf{1} - F)\mathbf{k} \rangle} = 1$  for any  $\mathbf{k} \in \ker(\mathbf{1} - \tilde{F})$ ; therefore,

$$|\mathrm{tr}([F, \mathbf{0}])|^2 = |\ker(\mathbf{1} - \tilde{F})|. \quad (\text{H.8})$$

If  $\chi \in \mathrm{range}(\mathbf{1} - \tilde{F})$ , then  $|\mathrm{tr}([F, \chi])|^2 = |\ker(\mathbf{1} - \tilde{F})|$  since  $[F, \chi]$  is conjugated to  $[F, \mathbf{0}]$ .

In addition,

$$\sum_{\chi \in \mathrm{range}(\mathbf{1} - \tilde{F})} |\mathrm{tr}([F, \chi])|^2 = |\mathrm{range}(\mathbf{1} - \tilde{F})| \times |\ker(\mathbf{1} - \tilde{F})| = d^2. \quad (\text{H.9})$$

On the other hand,

$$\sum_{\chi \in \mathbb{Z}_d^2} |\mathrm{tr}([F, \chi])|^2 = d \mathrm{tr}(V_F V_F^\dagger) = d^2, \quad (\text{H.10})$$

since the HW group is a nice error basis. Hence,  $[F, \chi]$  is traceless if  $\chi \notin \mathrm{range}(\mathbf{1} - \tilde{F})$ .

This observation completes the proof of the theorem when  $d$  is odd.

In the case  $d$  is even, if  $\begin{pmatrix} sd/2 \\ td/2 \end{pmatrix} \in \mathrm{range}(\mathbf{1} - \tilde{F})$  for  $s, t = 0, 1$ , then any  $\mathbf{k} \in \ker(\mathbf{1} - \tilde{F})$  is divisible by 2, and  $\langle \mathbf{k}, (\mathbf{1} - F)\mathbf{k} \rangle$  is divisible by  $\bar{d}$ . Consequently, Eqs. (H.8)–(H.10) still holds, and so does the theorem.

Otherwise, to prove the theorem, it remains to find  $F' \in \mathrm{SL}(2, \mathbb{Z}_{\bar{d}})$  such that  $F' = F \pmod{d}$ , and that  $\langle \mathbf{k}, (\mathbf{1} - F')\mathbf{k} \rangle$  is divisible by  $\bar{d}$  for all  $\mathbf{k} \in \ker(\mathbf{1} - \tilde{F})$ . Thanks to the Chinese remainder theorem, it suffices to consider the case when  $d$  is a power of 2.

Since any subgroup of  $\mathbb{Z}_d^2$  has rank at most two,  $\ker(\mathbf{1} - \tilde{F})$  can be generated by either one element or two elements. Applying a suitable conjugation on  $F$  if necessary,

## Appendix H. Supplementary information about the Clifford group

---

we may assume that  $\ker(\mathbf{1} - \tilde{F})$  is generated by  $\begin{pmatrix} a \\ 0 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ b \end{pmatrix}$ , where  $a$  and  $b$  are divisors of  $d$  including 1 and  $d$ , with  $a$  being divisible by  $b$ . Since  $\begin{pmatrix} sd/2 \\ td/2 \end{pmatrix} \in \text{range}(\mathbf{1} - \tilde{F})$  if  $b$  is divisible by 2, in which case the theorem has already been proved, it remains to consider the case  $b = 1$ . Now the assumption  $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \ker(\mathbf{1} - \tilde{F})$  implies that

$$F = \begin{pmatrix} 1 + rd & sd \\ j & 1 + (js + r)d \end{pmatrix}, \quad (\text{H.11})$$

where  $r, s = 0, 1$  and  $j \in \mathbb{Z}_{\bar{d}}$ . If  $j$  is divisible by  $d$ , then the choice  $F' = \mathbf{1}$  satisfies the conditions required in the theorem. Otherwise,  $F' = \begin{pmatrix} 1 & 0 \\ j & 1 \end{pmatrix}$  does, since

$$\left\langle \begin{pmatrix} k_1 a \\ k_2 \end{pmatrix}, (\mathbf{1} - F') \begin{pmatrix} k_1 a \\ k_2 \end{pmatrix} \right\rangle = j k_1^2 a^2 \quad (\text{H.12})$$

is divisible by  $\bar{d}$ , given that  $ja$  is divisible by  $d$  and  $a$  is even according to the assumption. □

Theorem H.1 implies that the absolute square of the trace of any Clifford operation is either zero or a divisor of  $d^2$ . Together with Theorem 7.1, this conclusion means that, within an HW covariant SIC POM, the number of fiducial states stabilized by each Clifford operation is either zero or a divisor of  $d^2$ . If  $d$  is odd and the Clifford operation  $[F, \chi]$  stabilizes a fiducial state, then  $\chi \in \text{range}(\mathbf{1} - F)$  and  $[F, \chi]$  is conjugated to  $[F, \mathbf{0}]$ . Therefore, on the orbit of any given fiducial state, there exists a fiducial state that is stabilized by a symplectic operation. This observation can help figure out the potential stabilizer of a fiducial state and search for a fiducial state with specific symmetry. On the other hand, if  $[F, \chi]$  is a nontrivial Clifford operation that belongs to a nice error basis, then  $\chi \notin \text{range}(\mathbf{1} - F)$ . This result is useful for determining the nice error bases in the Clifford group or in the symmetry group of an HW covariant SIC POM. When  $d$  is even, the above conclusions still hold if  $\begin{pmatrix} sd/2 \\ td/2 \end{pmatrix} \in \text{range}(\mathbf{1} - \tilde{F})$  for  $s, t = 0, 1$ ; otherwise, we need to replace  $F$  with suitable  $F'$  as required in Theorem H.1.

**Corollary H.2** *Let  $[F, \chi]$  be a Clifford operation of order  $n$  and  $k$  an integer that is coprime with  $n$ . Then  $|\text{tr}([F, \chi]^k)|^2 = |\text{tr}([F, \chi])|^2$ .*



## H.2. Normalizer of the Clifford group

---

Let  $\zeta$  be a primitive  $n$ th root of unity. Then there exists a polynomial  $f(x)$  with integer coefficients such that  $|\operatorname{tr}([F, \chi])|^2 = f(\zeta)$ . In other words,  $\zeta$  is a root of the polynomial  $g(x) = f(x) - |\operatorname{tr}([F, \chi])|^2$ , whose coefficients are integers owing to Theorem H.1. The polynomial  $g(x)$  is divisible by the minimal polynomial of  $\zeta$ , namely, the  $n$ th cyclotomic polynomial [75]. Since  $\zeta^k$  is a primitive  $n$ th root of unity when  $k$  is coprime with  $n$ , it is also a root of the cyclotomic polynomial and thus a root of the polynomial  $g(x)$ . Therefore,  $|\operatorname{tr}([F, \chi]^k)|^2 = f(\zeta^k) = |\operatorname{tr}([F, \chi])|^2$ .

**Corollary H.3** *Let  $[F, \chi]$  be a Clifford operation with nonzero trace. Then  $|\operatorname{tr}([F, \chi]^k)|^2$  is divisible by  $|\operatorname{tr}([F, \chi])|^2$  for any integer  $k$ . If  $d$  is odd, then  $[F, \chi]^k$  has a nonzero trace.*

According to Theorem H.1,  $|\operatorname{tr}([F, \chi])|^2 = |\ker(\mathbf{1} - \tilde{F})|$ . If  $d$  is odd, then  $[F, \chi]$  is conjugated to  $[F, \mathbf{0}]$  and we have  $|\operatorname{tr}([F, \chi]^k)|^2 = |\ker(\mathbf{1} - \tilde{F}^k)|$ . Otherwise,  $|\operatorname{tr}([F, \chi]^k)|^2$  is equal to either zero or  $|\ker(\mathbf{1} - \tilde{F}^k)|$ . Now the corollary follows from the fact that  $\ker(\mathbf{1} - \tilde{F})$  is a subgroup of  $\ker(\mathbf{1} - \tilde{F}^k)$ .

**Corollary H.4** *Let  $F \in \operatorname{SL}(2, \mathbb{Z}_{\bar{d}})$ ; then the following statements are equivalent.*

1.  $2 - \operatorname{tr}(F)$  is coprime with  $\bar{d}$ .
2.  $\mathbf{1} - \tilde{F}$  is invertible; that is,  $\ker(\mathbf{1} - \tilde{F}) = \{\mathbf{0}\}$  and/or  $\operatorname{range}(\mathbf{1} - \tilde{F}) = \mathbb{Z}_{\bar{d}}^2$ .
3.  $[F, \chi]$  is conjugated to  $[F, \mathbf{0}]$  for all  $\chi \in \mathbb{Z}_{\bar{d}}^2$ .
4.  $|\operatorname{tr}([F, \chi])|^2 = 1$  for all  $\chi \in \mathbb{Z}_{\bar{d}}^2$ .
5.  $|\operatorname{tr}([F, \chi])|^2 = 1$  for some  $\chi \in \mathbb{Z}_{\bar{d}}^2$ .

## H.2 Normalizer of the Clifford group

In this appendix, we briefly discuss the normalizer of the Clifford group; more details will be presented elsewhere [277]. The motivation behind this study is to explicate the regrouping phenomena of SIC POMs [119, 278, 283] (see Sec. 9.2.3), but the result may be of general interest.

---

## Appendix H. Supplementary information about the Clifford group

---

**Theorem H.5** *When  $d$  is not a multiple of 4, there is only one normal subgroup in the Clifford group that is unitarily equivalent to the HW group, and the normalizer of the Clifford group is itself. Otherwise, there are two such groups, which are conjugated to each other in the normalizer of the Clifford group. The additional normal HW group is generated by  $[F'_{1,d}, \binom{0}{1}]$  and  $[F'_{2,d}, \binom{1}{1}]$ , where*

$$F'_{1,d} := \begin{pmatrix} 1 - a\frac{d}{2} & 0 \\ a\frac{d}{2} & 1 + \frac{d}{2} \end{pmatrix}, \quad F'_{2,d} := \begin{pmatrix} 1 + \frac{d^2}{4} & \frac{d}{2} \\ \frac{d}{2} & 1 \end{pmatrix} \quad (\text{H.13})$$

with  $a = 1$  if  $d$  is divisible by 8 and  $a = -1$  otherwise.

The two generators of the additional normal HW group are explicitly given by

$$\begin{aligned} X' &:= \frac{1}{\sqrt{2}} e^{-i\pi/4} \tau \sum_{s=0}^{d-1} i^{s^2} \omega^s (|e_{s+1}\rangle\langle e_s| - i|e_{s+d/2+1}\rangle\langle e_s|) \in \left[ F'_{2,d}, \binom{1}{1} \right], \\ Z' &:= \sum_{s=0, s \text{ even}}^{d-1} \omega^s |e_s\rangle\langle e_s| - i \sum_{s=1, s \text{ odd}}^{d-1} \omega^s |e_{s+d/2}\rangle\langle e_s| \in \left[ F'_{1,d}, \binom{0}{1} \right]. \end{aligned} \quad (\text{H.14})$$

The two normal HW groups in the Clifford group can be mapped to each other by the unitary operator

$$U_d := \sum_{s=0, s \text{ even}}^{d-1} |e_s\rangle\langle e_s| + \frac{1}{\sqrt{2}} e^{i\pi/4} \sum_{s=1, s \text{ odd}}^{d-1} (|e_s\rangle\langle e_s| + i|e_{s+d/2}\rangle\langle e_s|), \quad (\text{H.15})$$

which, together with Clifford operators, generates the normalizer of the Clifford group.

More specifically,

$$U_d X' U_d^\dagger = \tau \sum_{s=0}^{d-1} \omega^s |e_{s+1}\rangle\langle e_s| = \tau X Z, \quad U_d Z' U_d^\dagger = \sum_{s=0}^{d-1} \omega^s |e_s\rangle\langle e_s| = Z. \quad (\text{H.16})$$

The square of  $U_d$  is a Clifford operator,

$$U_d^2 = \sum_{s=0, s \text{ even}}^{d-1} |e_s\rangle\langle e_s| - \sum_{s=1, s \text{ odd}}^{d-1} |e_{s+d/2}\rangle\langle e_s| \in \left[ \left( \begin{pmatrix} 1 - a\frac{d}{2} & 0 \\ d & 1 + \frac{d}{2} \end{pmatrix}, \binom{0}{0} \right) \right]. \quad (\text{H.17})$$

### H.3 HW groups in the Clifford group in a prime dimension

In this appendix, we determine all subgroups in the Clifford group in any prime dimension that are also HW groups, but in different bases, which play an important role in understanding the structure of SIC POMs in prime dimensions, especially in the case of dimension 3 [279] (see Sec. 8.5). When necessary, the HW group defined by Eq. (7.6) is referred to as the *standard HW group*.

In prime dimension  $p$ , each HW group in the Clifford group is a  $p$ -group and is thus contained in a Sylow  $p$ -subgroup of the Clifford group [172] (see Appendix G). In correspondence with the  $p + 1$  Sylow  $p$ -subgroups in  $\text{SL}(2, \mathbb{Z}_p)$  [154, 279], there are  $p + 1$  Sylow  $p$ -subgroups in the Clifford group, whose intersection is the standard HW group. Since all Sylow  $p$ -subgroups are conjugated to each other, to determine the HW groups in the Clifford group, it suffices to focus on a specific Sylow  $p$ -subgroup, say, the one generated by  $X$  and  $V_L$  [see Eq. (7.22)], which is denoted by  $\overline{P}_1$  henceforth.

The group  $\overline{P}_1$  has order  $p^3$ , and its center is the cyclic group generated by  $Z$ . The quotient of  $\overline{P}_1$  with respect to its center is isomorphic to  $\mathbb{Z}_p^2$ . There are  $p + 1$  order- $p^2$  subgroups in  $\overline{P}_1$ , namely,  $\langle Z, V_L^j X \rangle$  for  $j = 0, \dots, p - 1$  and  $\langle Z, V_L \rangle$ ; all of them are normal subgroups of  $\overline{P}_1$  and contain its center. The first  $p$  of them are unitarily equivalent to the HW group, which can be verified by examining the commutator between  $Z$  and  $V_L^j X$ . This is not the case for the group  $\langle Z, V_L \rangle$  since all its elements are diagonal. In summary, there are  $p(p + 1) + 1$  order- $p^2$  subgroups in the Clifford group, out of which  $p^2$  are unitarily equivalent to the HW group. Recall that the intersection of the  $p + 1$  Sylow  $p$ -subgroups is the standard HW group.

Define

$$U \cong \begin{cases} \text{diag}(1, e^{-i\pi/4}) & \text{if } p = 2, \\ \text{diag}(1, e^{-2i\pi/9}, e^{-4i\pi/9}) & \text{if } p = 3, \\ \text{diag}(1, \tau^{a_1}, \dots, \tau^{a_{p-1}}) & \text{if } p \geq 5, \end{cases} \quad (\text{H.18})$$

where  $a_k = \frac{1}{6}k(k+1)(2k+1)$  for  $k = 1, 2, \dots, p-1$ . Then  $U$  realizes a cyclic permutation

---

## Appendix H. Supplementary information about the Clifford group

---

among the HW groups in  $\overline{P}_1$ ,

$$U^j Z U^{j\dagger} = Z, \quad U^j X U^{j\dagger} = V_L^j X \quad \text{for } j = 0, \dots, p-1, \quad (\text{H.19})$$

where  $V_L'$  is defined as

$$V_L' \hat{=} \begin{cases} e^{i\pi/4} V_L & \text{if } p = 2, \\ e^{4i\pi/9} V_L & \text{if } p = 3, \\ V_L & \text{if } p \geq 5; \end{cases} \quad (\text{H.20})$$

it is unimodular thanks to the specific choice of the phase factor.

The  $p^2 - 1$  additional HW groups in the Clifford group form a single conjugacy class if  $p = 3$  or  $3|(p-2)$ , and three classes if  $3|(p-1)$ . That is, the  $p-1$  additional HW groups in  $\overline{P}_1$  form the corresponding number of classes under the conjugation of the Clifford group. Note that all Sylow  $p$ -subgroups are conjugated to each other. This claim is obvious when  $p = 2$ , so it remains to consider the odd-prime case.

Suppose the two HW groups  $\langle Z, V_L^m X \rangle$  and  $\langle Z, V_L^j X \rangle$  are conjugated to each other in the Clifford group; then they are conjugated to each other under some symplectic operation that belongs to the normalizer of  $\overline{P}_1$  and thus assumes the form  $[(\begin{smallmatrix} \alpha & 0 \\ \gamma & \alpha^{-1} \end{smallmatrix}), \mathbf{0}]$ . Up to an overall phase factor, the image of  $V_L^m X$  under the conjugation is  $V = V_L^{\alpha^{-2}m} X^\alpha Z^{m'}$ , where  $m' \in \mathbb{Z}_p$ , whose specific value is not relevant here. Note that  $V \in \langle Z, V_L^j X \rangle$  if and only if  $\alpha^{-3}m = j$ . If  $p = 3$  or  $3|(p-2)$ , then  $\alpha^{-3}$  can take on any value in  $\mathbb{Z}_p^*$ , so there exists  $\alpha$  satisfying  $\alpha^{-3}m = j$  for any pair  $m, j \in \mathbb{Z}_p^*$ . As a consequence, the  $(p-1)$  HW groups  $\langle Z, V_L^j X \rangle$  for  $j = 1, \dots, p-1$  are conjugated to each other in the Clifford group. If  $3|(p-1)$ , then  $\alpha^{-3}$  can take on only one third possible values in  $\mathbb{Z}_p^*$ , so the  $(p-1)$  HW groups form three classes of equal size  $\frac{p-1}{3}$ .

According to the above analysis, the intersection of the standard Clifford group and the normalizer of any HW group other than the standard one in  $\overline{P}_1$  is composed of the elements

$$\left[ \left( \begin{array}{cc} \alpha & 0 \\ \gamma & \alpha^{-1} \end{array} \right), \left( \begin{array}{c} k_1 \\ k_2 \end{array} \right) \right] \quad \text{with } \alpha^3 = 1. \quad (\text{H.21})$$

It contains  $\overline{P}_1$  as a subgroup of index 3 if  $3|(p-1)$  but is identical with  $\overline{P}_1$  otherwise.

# Some basic concepts in graph theory

---

A *graph*  $G$  is composed of a collection of vertices and a collection of edges, in which each edge connects two vertices [42]. Each vertex can be represented graphically as a dot, and each edge as an arrow or a line depending on whether it is directed or not. A graph is *directed* if all edges are directed and is *undirected* if no edge is directed. A *weighted graph* is a graph in which each edge is associated with a weight; by introducing enough number of distinct weights, we may assume, without loss of generality, that every pair of vertices is connected by an edge. An ordinary graph may be seen as a special example of a weighted graph in which there are only two kinds of weights, say, 0 and 1.

A graph with  $n$  vertices can be represented by an  $n \times n$  matrix, called the *adjacency matrix* [42]. The  $i$ - $j$ th entry of the adjacency matrix is the weight of the edge connecting vertex  $i$  and vertex  $j$  if the edge is undirected or directed from  $i$  to  $j$ ; otherwise, it is the additive inverse of the weight. By definition, the adjacency matrix is symmetric for an undirected graph, but antisymmetric for a directed graph.

An *isomorphism* between two graphs  $G_1$  and  $G_2$  is a one-to-one mapping from the vertices of  $G_1$  to those of  $G_2$  that preserves directions and weights of all edges [170, 258]. A *skew isomorphism* is a one-to-one mapping that preserves the weights, but reverses the directions. Two graphs are (skew) isomorphic if there exists an (skew) isomorphism between them. Let  $A_1$  and  $A_2$  be the adjacency matrices of  $G_1$  and  $G_2$ . Then  $G_1$  and  $G_2$  are (skew) isomorphic if and only if there exists a permutation matrix  $P$  such that  $(-1)PA_1P^T = A_2$ . A *canonical form* of a graph  $G$  is a graph  $\text{Canon}(G)$  that is isomorphic to  $G$ , such that  $\text{Canon}(G_1) = \text{Canon}(G_2)$  if and only if  $G_1$  is isomorphic to  $G_2$ .

One of the best known canonical forms is the lexicographically smallest graph within the isomorphism class, which is the graph with lexicographically smallest adjacency matrix [16, 17]. The definitions of (skew) isomorphism and canonical forms can also be generalized to adjacency matrices and even arbitrary matrices, whether they are associated with graphs or not.

An (skew) *automorphism* of a graph  $G$  is an (skew) isomorphism between  $G$  and itself [170, 258]. The *automorphism group* of  $G$  is composed of all these automorphisms and is denoted by  $\text{Aut}(G)$ . Similarly, the *extended automorphism group*  $\text{Aut}_E(G)$  is composed of all automorphisms and skew automorphisms. The (extended) automorphism group of a matrix can be defined similarly. For the convenience of discussion, we may identify the (extended) automorphism group of a graph with that of its adjacency matrix.

The *graph automorphism problem* and the *graph isomorphism problem* are two fundamental problems in graph theory. The former is to compute the automorphism group of a graph, while the latter is to determine the isomorphism relation between two graphs. Although no polynomial-time (in the number of vertices) algorithm is known for either problem, they can be solved efficiently for almost all graphs in practice. A common approach to determining the isomorphism relation is to compute canonical forms [16, 17, 170, 185, 194, 258].

A *hypergraph* [261] is a generalization of a graph in which some edges may take more than two vertices. Let  $k \geq 2$  be an integer. A hypergraph is *k-uniform* if every edge takes  $k$  vertices. In that case, it can be represented by a  $k$ -dimensional array. The concepts of isomorphism and automorphism can be generalized to hypergraphs in a straightforward way.

# Bibliography

- [1] Y. Aharonov and J. Anandan. Phase change during a cyclic quantum evolution. *Phys. Rev. Lett.*, 58:1593, 1987.
- [2] O. Albouy and M. R. Kibler. A unified approach to SIC-POVMs and MUBs. *J. Russ. Laser Res.*, 28(5):429–438, 2007.
- [3] E. M. Alfsen. *Compact Convex Sets and Boundary Integrals*. Springer-Verlag, New York, 1971.
- [4] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O’Brien, M. A. Nielsen, and A. G. White. Ancilla-assisted quantum process tomography. *Phys. Rev. Lett.*, 90:193601, 2003.
- [5] M. Altunbulak and A. Klyachko. The Pauli principle revisited. *Commun. Math. Phys.*, 282:287–322, 2008.
- [6] D. M. Appleby. Symmetric informationally complete-positive operator valued measures and the extended Clifford group. *J. Math. Phys.*, 46:052107, 2005.
- [7] D. M. Appleby. Properties of the extended Clifford group with applications to SIC-POVMs and MUBs, 2009. Available at <http://arxiv.org/abs/0909.5233>.
- [8] D. M. Appleby. SIC-POVMs and MUBs: Geometrical relationships in prime dimension. *AIP Conf. Proc.*, 1101:223, 2009.
- [9] D. M. Appleby. The analytical solutions for  $d = 31, 37, 43$  were reported by Appleby in private communication, 2011.
- [10] D. M. Appleby, I. Bengtsson, S. Brierley, M. Grassl, D. Gross, and J. Å Larson. The monomial representations of the Clifford group. *Quant. Inf. Comput.*, 12:0404–0431, 2012.
- [11] D. M. Appleby, I. Bengtsson, and S. Chaturvedi. Spectra of phase point operators in odd prime dimensions and the extended Clifford group. *J. Math. Phys.*, 49:012102, 2008.
- [12] D. M. Appleby, H. B. Dang, and C. A. Fuchs. Symmetric informationally complete quantum states as analogues to orthonormal bases and minimum-uncertainty states, 2007. Available at <http://arxiv.org/abs/0707.2071>.
- [13] D. M. Appleby, Å. Ericsson, and C. A. Fuchs. Properties of QBist state spaces. *Found. Phys.*, 41:564, 2011.
- [14] D. M. Appleby, S. T. Flammia, and C. A. Fuchs. The Lie algebraic significance of symmetric informationally complete measurements. *J. Math. Phys.*, 52:022202, 2011.
- [15] M. A. Armen, J. K. Au, J. K. Stockton, A. C. Doherty, and H. Mabuchi. Adaptive homodyne measurement of optical phase. *Phys. Rev. Lett.*, 89:133602, 2002.

## Bibliography

---

- [16] L. Babai and L. Kucera. Canonical labeling of graphs in average linear time. *Proc. 20th Annual IEEE Symposium on Foundations of Computer Science*, page 39, 1979.
- [17] L. Babai and E. M. Luks. Canonical labeling of graphs. *Proceedings of the 15th Annual ACM Symposium on Theory of Computing*, page 171, 1983.
- [18] E. Bagan, M. Baig, R. Muñoz-Tapia, and A. Rodriguez. Collective versus local measurements in a qubit mixed-state estimation. *Phys. Rev. A*, 69:010304(R), 2004.
- [19] E. Bagan, M. A. Ballester, R. D. Gill, A. Monras, and R. Muñoz-Tapia. Optimal full estimation of qubit mixed states. *Phys. Rev. A*, 73:032301, 2006.
- [20] E. Bagan, M. A. Ballester, R. D. Gill, R. Muñoz-Tapia, and O. Romero-Isart. Separable measurement estimation of density matrices and its fidelity gap with collective protocols. *Phys. Rev. Lett.*, 97:130501, 2006.
- [21] T. Baier and D. Petz. Complementarity and state estimation. *Rep. Math. Phys.*, 65(2):203–214, 2010.
- [22] M. A. Ballester. Estimating the spectrum of a density matrix with LOCC. *J. Phys. A: Math. Gen.*, 39(7):1645, 2006.
- [23] W. Band and J. L. Park. The empirical determination of quantum states. *Found. Phys.*, 1:133–144, 1970.
- [24] W. Band and J. L. Park. A general method of empirical state determination in quantum physics: Part II. *Found. Phys.*, 1:339–357, 1971.
- [25] W. Band and J. L. Park. Quantum state determination: Quorum for a particle in one dimension. *Am. J. Phys.*, 47(2):188–191, 1979.
- [26] E. Bannai and E. Bannai. A survey on spherical designs and algebraic combinatorics on spheres. *Eur. J. Combinator.*, 30(6):1392–1425, 2009.
- [27] V. Bargmann. Note on Wigner’s theorem on symmetry operations. *J. Math. Phys.*, 5:862, 1964.
- [28] O. E. Barndorff-Nielsen and R. D. Gill. Fisher information in quantum statistics. *J. Phys. A: Math. Gen.*, 33(24):4481, 2000.
- [29] I. Bengtsson. MUBs, polytopes, and finite geometries. *AIP Conf. Proc.*, 750(1):63–69, 2005.
- [30] I. Bengtsson and K. Życzkowski. *Geometry of Quantum States: An Introduction to Quantum Entanglement*. Cambridge University Press, Cambridge, UK, 2006.
- [31] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895, 1993.



## Bibliography

---

- [32] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59(2):1070–1091, 1999.
- [33] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69(20):2881–2884, 1992.
- [34] D. S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory*. Princeton University Press, Princeton, New Jersey, 2005.
- [35] M. V. Berry. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A*, 392:45, 1984.
- [36] J. Bertrand and P. Bertrand. A tomographic approach to Wigner’s function. *Found. Phys.*, 17(4):397–405, 1987.
- [37] H. F. Blichfeldt. *Finite Collineation Groups*. University of Chicago Press, Chicago, IL, 1917.
- [38] R. Blume-Kohout. Hedged maximum likelihood quantum state estimation. *Phys. Rev. Lett.*, 105:200504, 2010.
- [39] R. Blume-Kohout. Optimal, reliable estimation of quantum states. *New J. Phys.*, 12(4):043034, 2010.
- [40] R. Blume-Kohout, J. O. S. Yin, and S. J. van Enk. Entanglement verification with finite data. *Phys. Rev. Lett.*, 105:170501, 2010.
- [41] N. Bohr. The quantum postulate and the recent development of atomic theory. *Nature*, 121:580–590, 1928.
- [42] J. A. Bondy and U. S. R. Murty. *Graph Theory with Applications*. Macmillan, London, 1976.
- [43] M. Bourennane, M. Eibl, C. Kurtsiefer, S. Gaertner, H. Weinfurter, O. Gühne, P. Hyllus, D. Bruß, M. Lewenstein, and A. Sanpera. Experimental detection of multipartite entanglement using witness operators. *Phys. Rev. Lett.*, 92(8):087902, 2004.
- [44] F. A. Bovino, G. Castagnoli, A. K. Ekert, P. Horodecki, C. M. Alves, and A. V. Sergienko. Direct measurement of nonlinear properties of bipartite quantum states. *Phys. Rev. Lett.*, 95:240407, 2005.
- [45] S. L. Braunstein and C. M. Caves. Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.*, 72:3439–3443, 1994.
- [46] Č. Brukner and A. Zeilinger. Operationally invariant information in quantum measurements. *Phys. Rev. Lett.*, 83:3354–3357, 1999.

## Bibliography

---

- [47] D. Bruß, M. Christandl, A. K. Ekert, B.-G. Englert, D. Kaszlikowski, and C. Macchiavello. Tomographic quantum cryptography: Equivalence of quantum and classical key distillation. *Phys. Rev. Lett.*, 91:097901, 2003.
- [48] D. Bruß, A. K. Ekert, and C. Macchiavello. Optimal universal quantum cloning and state estimation. *Phys. Rev. Lett.*, 81:2598–2601, 1998.
- [49] D. Bruß and C. Macchiavello. Optimal state estimation for  $d$ -dimensional quantum systems. *Phys. Lett. A*, 253(5-6):249–251, 1999.
- [50] D. J. C. Bures. An extension of Kakutani’s theorem on infinite product measures to tensor product of semifinite  $w^*$ -algebras. *Trans. Am. Math. Soc.*, 135:199–212, 1969.
- [51] M. D. de Burgh, N. K. Langford, A. C. Doherty, and A. Gilchrist. Choice of measurement sets in qubit tomography. *Phys. Rev. A*, 78:052122, 2008.
- [52] P. Busch. Informationally complete-sets of physical quantities. *Int. J. Theor. Phys.*, 30:1217, 1991.
- [53] V. Bužek, G. Adam, and G. Drobný. Reconstruction of Wigner functions on different observation levels. *Ann. Phys.*, 245(1):37–97, 1996.
- [54] V. Bužek, G. Drobný, G. Adam, R. Derka, and P. Knight. Reconstruction of quantum states of spin systems via the Jaynes principle of maximum entropy. *J. Mod. Opt.*, 44(11-12):2607–2627, 1997.
- [55] V. Bužek, G. Drobný, R. Derka, G. Adam, and H. Wiedemann. Quantum state reconstruction from incomplete data. *Chaos, Solitons & Fractals*, 10(6):981–1074, 1999.
- [56] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum error correction via codes over  $GF(4)$ . *IEEE Trans. Inf. Theory*, 44(4):1369, 1998.
- [57] E. J. Candès, J. K. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inf. Theory*, 52(2):489–509, 2006.
- [58] E. J. Candès, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math.*, 59:1207–1223, 2006.
- [59] E. J. Candès and T. Tao. Decoding by linear programming. *IEEE Trans. Inf. Theory*, 51(12):4203–4215, 2005.
- [60] E. J. Candès and M. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.
- [61] P. G. Casazza. The art of frame theory. *Taiw. J. Math.*, 4:129, 2000.
- [62] C. M. Caves. Available at <http://info.phys.unm.edu/~caves/reports/reports.html>.

## Bibliography

---

- [63] N. N. Čencov. *Statistical Decision Rules and Optimal Inferences*, volume 53 of *Transl. Math. Monogr.* Am. Math. Soc., Providence, 1982.
- [64] L. Chen, H. Zhu, and T.-C. Wei. Connections of geometric measure of entanglement of pure symmetric states to quantum state estimation. *Phys. Rev. A*, 83(1):012305, 2011.
- [65] A. M. Childs, I. L. Chuang, and D. W. Leung. Realization of quantum process tomography in NMR. *Phys. Rev. A*, 64:012314, 2001.
- [66] I. L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *J. Mod. Opt.*, 44(11-12):2455–2467, 1997.
- [67] S. Colin, J. Corbett, T. Durt, and D. Gross. About SIC POVMs and discrete Wigner distributions. *J. Opt. B: Quantum Semiclass. Opt.*, 7(12):S778, 2005.
- [68] H. Cramér. *Mathematical Methods of Statistics*. Princeton University Press, Princeton, New Jersey, 1946.
- [69] C. W. Curtis and I. Reiner. *Representation Theory of Finite Groups and Associative Algebras*. Interscience Publishers, New York, 1962.
- [70] G. M. D’Ariano and P. Lo Presti. Quantum tomography for measuring experimentally the matrix elements of an arbitrary quantum operation. *Phys. Rev. Lett.*, 86:4195–4198, 2001.
- [71] G. M. D’Ariano, P. Lo Presti, and M. F. Sacchi. Bell measurements and observables. *Phys. Lett. A*, 272:32, 2000.
- [72] G. M. D’Ariano, C. Macchiavello, and M. G. A. Paris. Detection of the density matrix through optical homodyne tomography without filtered back projection. *Phys. Rev. A*, 50:4298–4302, 1994.
- [73] G. M. D’Ariano and P. Perinotti. Optimal data processing for quantum measurements. *Phys. Rev. Lett.*, 98:020403, 2007.
- [74] G. M. D’Ariano, P. Perinotti, and M. F. Sacchi. Informationally complete measurements and group representation. *J. Opt. B: Quantum Semiclass. Opt.*, 6:S487, 2004.
- [75] R. A. Dean. *Classical Abstract Algebra*. Harper & Row, New York, 1990.
- [76] P. Delsarte, J. M. Goethals, and J. J. Seidel. Bounds for systems of lines, and Jacobi polynomials. *Philips Res. Rep.*, 30:91, 1975.
- [77] P. Delsarte, J. M. Goethals, and J. J. Seidel. Spherical codes and designs. *Geom. Dedicata*, 6:363–388, 1977.
- [78] R. Derka, V. Bužek, and A. K. Ekert. Universal algorithm for optimal estimation of quantum states from finite ensembles via realizable generalized measurement. *Phys. Rev. Lett.*, 80:1571–1575, 1998.

## Bibliography

---

- [79] D. L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory*, 52:1289–1306, 2006.
- [80] J. Du, M. Sun, X. Peng, and T. Durt. Realization of entanglement-assisted qubit-covariant symmetric-informationally-complete positive-operator-valued measurements. *Phys. Rev. A*, 74:042341, 2006.
- [81] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Trans. Am. Math. Soc.*, 72:341, 1952.
- [82] T. J. Dunn, I. A. Walmsley, and S. Mukamel. Experimental determination of the quantum-mechanical state of a molecular vibrational mode using fluorescence tomography. *Phys. Rev. Lett.*, 74:884–887, 1995.
- [83] T. Durt, B.-G. Englert, I. Bengtsson, and K. Życzkowski. On mutually unbiased bases. *Int. J. Quant. Inf.*, 8:535, 2010.
- [84] T. Durt, C. Kurtsiefer, A. Lamas-Linares, and A. Ling. Wigner tomography of two-qubit states and quantum cryptography. *Phys. Rev. A*, 78:042338, 2008.
- [85] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777, 1935.
- [86] A. K. Ekert. Quantum cryptography based on Bell’s theorem. *Phys. Rev. Lett.*, 67:661, 1991.
- [87] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek. Direct estimations of linear and nonlinear functionals of a quantum state. *Phys. Rev. Lett.*, 88:217901, 2002.
- [88] F. Embacher and H. Narnhofer. Strategies to measure a quantum state. *Ann. Phys.*, 311(1):220–244, 2004.
- [89] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme. Symmetrized characterization of noisy quantum processes. *Science*, 317(5846):1893–1896, 2007.
- [90] B.-G. Englert. Fringe visibility and which-way information: An inequality. *Phys. Rev. Lett.*, 77:2154–2157, 1996.
- [91] B.-G. Englert, D. Kaszlikowski, H. K. Ng, W. K. Chua, J. Řeháček, and J. Anders. Efficient and robust quantum key distribution with minimal state tomography, 2004. Available at <http://arxiv.org/abs/quant-ph/0412075>.
- [92] U. Fano. Description of states in quantum mechanics by density matrix and operator techniques. *Rev. Mod. Phys.*, 29:74–93, 1957.
- [93] R. A. Fisher. On the mathematical foundations of theoretical statistics. *Phil. Trans. R. Soc. Lond. A*, 222:309–368, 1922.
- [94] R. A. Fisher. Theory of statistical estimation. *Math. Proc. Cambr. Phil. Soc.*, 22(05):700–725, 1925.

## Bibliography

---

- [95] S. T. Flammia. On SIC-POVMs in prime dimensions. *J. Phys. A: Math. Gen.*, 39:13483, 2006.
- [96] S. T. Flammia and Y.-K. Liu. Direct fidelity estimation from few Pauli measurements. *Phys. Rev. Lett.*, 106:230501, 2011.
- [97] C. A. Fuchs. *Distinguishability and Accessible Information in Quantum Theory*. PhD thesis, The University of New Mexico, 1996. Available at <http://arxiv.org/abs/quant-ph/9601020>.
- [98] C. A. Fuchs. Quantum mechanics as quantum information, mostly. *J. Mod. Opt.*, 50(6-7):987–1023, 2003.
- [99] C. A. Fuchs. QBism, the perimeter of quantum Bayesianism, 2010. Available at <http://arxiv.org/abs/1003.5209>.
- [100] C. A. Fuchs and M. Sasaki. Squeezing quantum information through a classical channel: Measuring the “quantumness” of a set of quantum states. *Quant. Inf. Comput.*, 3(5):377–404, 2003.
- [101] C. A. Fuchs and R. Schack. Quantum-Bayesian coherence, 2009. Available at <http://arxiv.org/abs/0906.2187>.
- [102] C. A. Fuchs and J. van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Trans. Inf. Theory*, 45:1216, 1999.
- [103] A. Fujiwara. *A geometrical study in quantum information systems*. PhD thesis, University of Tokyo, 1995. Part II of this thesis was reprinted in Ref. [133].
- [104] A. Fujiwara and H. Nagaoka. Quantum Fisher metric and estimation for pure state models. *Phys. Lett. A*, 201:119–124, 1995.
- [105] A. Fujiwara and H. Nagaoka. An estimation theoretical characterization of coherent states. *J. Math. Phys.*, 40(9):4227–4239, 1999.
- [106] W. Gale, E. Guth, and G. T. Trammell. Determination of the quantum state by measurements. *Phys. Rev.*, 165:1434–1436, 1968.
- [107] R. D. Gill and S. Massar. State estimation for large ensembles. *Phys. Rev. A*, 61:042312, 2000.
- [108] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum-enhanced measurements: Beating the standard quantum limit. *Science*, 306:1330, 2004.
- [109] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum metrology. *Phys. Rev. Lett.*, 96:010401, 2006.
- [110] V. Giovannetti, S. Lloyd, and L. Maccone. Advances in quantum metrology. *Nat. Photon.*, 5:222, 2011.
- [111] N. Giri and W. von Waldenfels. An algebraic version of the central limit theorem. *Probability Theory and Related Fields*, 42:129–134, 1978.

## Bibliography

---

- [112] C. Godsil and A. Roy. Equiangular lines, mutually unbiased bases, and spin models. *Eur. J. Combinator.*, 30:246–262, 2009.
- [113] D. Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, 1997. Available at <http://arxiv.org/abs/quant-ph/9705052>.
- [114] D. Gottesman. Theory of fault-tolerant quantum computation. *Phys. Rev. A*, 57:127, 1998.
- [115] M. Grassl. On SIC-POVMs and MUBs in dimension 6. In *Proceedings of the 2004 ERATO Conference on Quantum Information Science*, pages 60–61, Tokyo, 2004. Available at <http://arxiv.org/abs/quant-ph/0406175>.
- [116] M. Grassl. Tomography of quantum states in small dimensions. *Electron. Notes Discrete Math.*, 20:151, 2005.
- [117] M. Grassl. Finding equiangular lines in complex space. In *MAGMA 2006 Conference*, Technische Universität Berlin, 2006. Available at <http://magma.maths.usyd.edu.au/Magma2006/>.
- [118] M. Grassl. Computing equiangular lines in complex space. *Lect. Notes Comput. Sci.*, 5393:89, 2008.
- [119] M. Grassl. Seeking symmetries of SIC-POVMs. In *Seeking SICs: A Workshop on Quantum Frames and Designs*, Perimeter Institute, Waterloo, 2008. Available at <http://pirsa.org/08100069/>.
- [120] M. Grassl. The analytical solution for  $d = 28$  was reported by Markus Grassl in private communication, 2011.
- [121] D. Gross. Culs-de-sac and open ends. In *Seeking SICs: A Workshop on Quantum Frames and Designs*, Perimeter Institute, Waterloo, 2008. Available at <http://pirsa.org/08100075/>.
- [122] D. Gross, K. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *J. Math. Phys.*, 48(5):052104, 2007.
- [123] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401, 2010.
- [124] M. Guță, B. Janssens, and J. Kahn. Optimal estimation of qubit states with continuous time measurements. *Commun. Math. Phys.*, 277:127–160, 2008.
- [125] M. Guță and A. Jenčová. Local asymptotic normality in quantum statistics. *Commun. Math. Phys.*, 276:341–379, 2007.
- [126] M. Guță and J. Kahn. Local asymptotic normality for qubit states. *Phys. Rev. A*, 73:052108, 2006.
- [127] J. Haantjes. Equilateral point-sets in elliptic two- and three-dimensional spaces. *Nieuw Arch. Wisk.*, 22:355–362, 1948.

## Bibliography

---

- [128] A. Hayashi, T. Hashimoto, and M. Horibe. Reexamination of optimal quantum state estimation of pure states. *Phys. Rev. A*, 72:032325, 2005.
- [129] M. Hayashi. A linear programming approach to attainable Cramér-Rao type bounds. In O. Hirota, A. S. Holevo, and C. A. Caves, editors, *Quantum Communication, Computing, and Measurement*, New York, 1997. Plenum. Reprinted in Ref. [133].
- [130] M. Hayashi. Asymptotic estimation theory for a finite-dimensional pure state model. *J. Phy. A: Math. Gen.*, 31(20):4633, 1998.
- [131] M. Hayashi. Asymptotic quantum estimation theory for the displaced thermal states family. In P. Kumar, G. M. D’Ariano, and O. Hirota, editors, *Quantum Communication, Computing, and Measurement*, New York, 2000. Kluwer/Plenum. Reprinted in Ref. [133].
- [132] M. Hayashi. Two quantum analogues of Fisher information from a large deviation viewpoint of quantum estimation. *J. Phys. A: Math. Gen.*, 35(36):7689, 2002.
- [133] M. Hayashi, editor. *Asymptotic Theory of Quantum Statistical Inference*. World Scientific, Singapore, 2005.
- [134] M. Hayashi. *Quantum Information: An Introduction*. Springer, Berlin, 2006.
- [135] M. Hayashi. Quantum estimation and the quantum central limit theorem. *American Mathematical Society Translations Series 2*, 227:95–123, 2009. Original Japanese version was published in *Bulletin of Mathematical Society of Japan, Sugaku*, 55(4):368–391, 2003.
- [136] M. Hayashi and K. Matsumoto. Statistical model with measurement degree of freedom and quantum physics. In M. Hayashi, editor, *Asymptotic theory of quantum statistical inference*, Singapore, 2005. World scientific. Original Japanese version was published in *Surikaiseki Kenkyusho Kokyuroku*, 1055:96-110, 1998.
- [137] M. Hayashi and K. Matsumoto. Asymptotic performance of optimal state estimation in qubit system. *J. Math. Phys.*, 49(10):102101, 2008.
- [138] W. Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeit. für Physik*, 43:172, 1927.
- [139] C. W. Helstrom. Minimum mean-squared error of estimates in quantum statistics. *Phys. Lett. A*, 25(2):101–102, 1967.
- [140] C. W. Helstrom. The minimum variance of estimates in quantum signal detection. *IEEE Trans. Inf. Theory*, 14(2):234–242, 1968.
- [141] C. W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, New York, 1976.
- [142] C. W. Helstrom and R. S. Kennedy. Noncommuting observables in quantum detection and estimation theory. *IEEE Trans. Inf. Theory*, 20(1):16–24, 1974.

## Bibliography

---

- [143] B. L. Higgins, D. W. Berry, S. D. Bartlett, H. M. Wiseman, and G. J. Pryde. Entanglement-free Heisenberg-limited phase estimation. *Nature*, 450(7168):393–396, 2007.
- [144] S. G. Hoggar.  $t$ -designs in projective spaces. *Eur. J. Combinator.*, 3:233–254, 1982.
- [145] S. G. Hoggar.  $t$ -designs with general angle set. *Eur. J. Combinator.*, 13(4):257–271, 1992.
- [146] S. G. Hoggar. 64 lines from a quaternionic polytope. *Geom. Dedicata*, 69:287, 1998.
- [147] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland, Amsterdam, 1982.
- [148] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: Necessary and sufficient conditions. *Phys. Lett. A*, 223:1, 1996.
- [149] P. Horodecki and A. K. Ekert. Method for direct detection of quantum entanglement. *Phys. Rev. Lett.*, 89:127902, 2002.
- [150] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865, 2009.
- [151] S. D. Howard, A. R. Calderbank, and W. Moran. The finite Heisenberg-Weyl groups in radar and communications. *EURASIP J. Appl. Signal Processing*, 2006:85685, 2006.
- [152] Z. Hradil. Quantum-state estimation. *Phys. Rev. A*, 55:R1561, 1997.
- [153] M. Hübner. Explicit computation of the Bures distance for density matrices. *Phys. Lett. A*, 163:239–242, 1992.
- [154] J. E. Humphreys. Representations of  $SL(2, p)$ . *Am. Math. Monthly*, 82:21, 1975.
- [155] I. D. Ivanović. Geometrical description of quantal state determination. *J. Phys. A: Math. Gen.*, 14:3241, 1981.
- [156] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White. Measurement of qubits. *Phys. Rev. A*, 64(5):052312, 2001.
- [157] E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630, 1957.
- [158] E. T. Jaynes. Information theory and statistical mechanics. II. *Phys. Rev.*, 108:171–190, 1957.
- [159] R. Jozsa. Entanglement and quantum computation, 1997. Available at <http://arxiv.org/abs/quant-ph/9707034>.
- [160] J. Kahn and M. Guță. Local asymptotic normality for finite dimensional quantum systems. *Commun. Math. Phys.*, 289:597–652, 2009.



## Bibliography

---

- [161] A. Kalev, J. Shang, and B.-G. Englert. Experimental proposal for symmetric minimal two-qubit state tomography. *Phys. Rev. A*, 85:052115, 2012.
- [162] A. Kalev, J. Shang, and B.-G. Englert. Symmetric minimal quantum tomography by successive measurements. *Phys. Rev. A*, 85:052116, 2012.
- [163] M. Khatirinejad. On Weyl-Heisenberg orbits of equiangular lines. *J. Algebr. Comb.*, 28:333–349, 2008.
- [164] M. Khatirinejad. *Regular Structures of Lines in Complex Spaces*. PhD thesis, Simon Fraser University, 2008.
- [165] I. H. Kim. Quantumness, generalized design and symmetric informationally complete POVM. *Quantum Inf. Comput.*, 7(8):730–737, 2007.
- [166] A. Klappenecker and M. Rötteler. Beyond stabilizer codes I: Nice error bases. *IEEE Trans. Inf. Theory*, 48:2392, 2002. Supplementary information including a catalogue of nice error bases available at <http://www.cs.tamu.edu/faculty/klappi/ueb/ueb.html>.
- [167] A. Klappenecker and M. Rötteler. Mutually unbiased bases are complex projective 2-designs. In *IEEE International Symposium on Information Theory*, pages 1740–1744, Adelaide, Australia, 2005.
- [168] A. Klappenecker and M. Rötteler. On the monomiality of nice error bases. *IEEE Trans. Inf. Theory*, 51:1084, 2005.
- [169] E. Knill. Group representations, error bases and quantum codes. *Los Alamos National Laboratory Report LAUR-96-2807*, 1996. Available at <http://arxiv.org/abs/quant-ph/9608049>.
- [170] D. L. Kreher and D. R. Stinson. *Combinatorial Algorithms: Generation, Enumeration, and Search*. CRC Press, Boca Raton, 1999.
- [171] C. Kurtsiefer, T. Pfau, and J. Mlynek. Measurement of the Wigner function of an ensemble of helium atoms. *Nature*, 386:150, 1997.
- [172] H. Kurzweil and B. Stellmacher. *The Theory of Finite Groups: An Introduction*. Springer, New York, 2004.
- [173] E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer, 1998.
- [174] D. Leibfried, D. M. Meekhof, B. E. King, C. Monroe, W. M. Itano, and D. J. Wineland. Experimental determination of the motional quantum state of a trapped atom. *Phys. Rev. Lett.*, 77:4281–4285, 1996.
- [175] P. W. H. Lemmens and J. J. Seidel. Equiangular lines. *J. Algebra*, 24:494, 1973.
- [176] U. Leonhardt. *Measuring the Quantum State of Light*. Cambridge University Press, Cambridge, UK, 1997.

## Bibliography

---

- [177] U. Leonhardt, M. Munroe, T. Kiss, T. Richter, and M. G. Raymer. Sampling of photon statistics and density matrix using homodyne detection. *Opt. Commun.*, 127(1–3):144–160, 1996.
- [178] U. Leonhardt, H. Paul, and G. M. D’Ariano. Tomographic reconstruction of the density matrix via pattern functions. *Phys. Rev. A*, 52:4899–4907, 1995.
- [179] D. W. Leung. *Towards Robust Quantum Computation*. PhD thesis, Stanford University, 2000.
- [180] D. W. Leung. Choi’s proof as a recipe for quantum process tomography. *J. Math. Phys.*, 44(2):528–533, 2003.
- [181] Y. C. Liang, D. Kaszlikowski, B.-G. Englert, L.-C. Kwek, and C. H. Oh. Tomographic quantum cryptography. *Phys. Rev. A*, 68:022324, 2003.
- [182] G. J. Lidstone. Note on the general case of the Bayes-Laplace formula for inductive or a posteriori probabilities. *Trans. Fac. Actuaries*, 8:182–192, 1920.
- [183] A. Ling, A. Lamas-Linares, and C. Kurtsiefer. Accuracy of minimal and optimal qubit tomography for finite-length experiments, 2008. Available at <http://arxiv.org/abs/0807.0991>.
- [184] J. H. van Lint and J. J. Seidel. Equilateral point sets in elliptic geometry. *Proc. Kon. Nederl. Akad. Wet. Ser. A*, 69:335–348, 1966.
- [185] E. M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *J. Comput. Syst. Sci.*, 25(1):42–65, 1982.
- [186] A. I. Lvovsky and M. G. Raymer. Continuous-variable optical quantum-state tomography. *Rev. Mod. Phys.*, 81:299, 2009.
- [187] G. Mackiw. The linear group  $SL(2, 3)$  as a source of examples. *The Mathematical Gazette*, 81:64–67, 1997.
- [188] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and its Applications*. Academic Press, New York, 1979.
- [189] F. de Martini, A. Mazzei, M. Ricci, and G. M. D’Ariano. Exploiting quantum parallelism of entanglement for a complete experimental quantum characterization of a single-qubit device. *Phys. Rev. A*, 67:062307, 2003.
- [190] I. Marvian and R. W. Spekkens. A generalization of Schur-Weyl duality with applications in quantum estimation, 2011. Available at <http://arxiv.org/abs/1112.0638>.
- [191] S. Massar and S. Popescu. Optimal extraction of information from finite quantum ensembles. *Phys. Rev. Lett.*, 74:1259–1263, 1995.
- [192] K. Matsumoto. *A Geometrical Approach to Quantum Estimation Theory*. PhD thesis, Graduate School of Mathematical Sciences, University of Tokyo, 1997. Part of this thesis was reprinted in Ref. [133].

## Bibliography

---

- [193] K. Matsumoto. A new approach to the Cramér–Rao-type bound of the pure-state model. *J. Phys. A: Math. Gen.*, 35(13):3111, 2002.
- [194] B. D. McKay. Practical graph isomorphism. *Congressus Numerantium*, 30:45, 1981.
- [195] Z. E. D. Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. M. Tabia, C. A. Fuchs, and A. M. Steinberg. Experimental characterization of qutrits using symmetric informationally complete positive operator-valued measurements. *Phys. Rev. A*, 83:051801(R), 2011.
- [196] M. L. Mehta. *Random Matrices*. Elsevier, Amsterdam, 2004.
- [197] M. W. Mitchell, C. W. Ellenor, S. Schneider, and A. M. Steinberg. Diagnosis, prescription, and prognosis of a Bell-state filter by quantum process tomography. *Phys. Rev. Lett.*, 91:120402, 2003.
- [198] M. Mohseni and D. A. Lidar. Direct characterization of quantum dynamics. *Phys. Rev. Lett.*, 97:170501, 2006.
- [199] M. Mohseni and D. A. Lidar. Direct characterization of quantum dynamics: General theory. *Phys. Rev. A*, 75:062331, 2007.
- [200] M. Mohseni, A. T. Rezakhani, and D. A. Lidar. Quantum-process tomography: Resource analysis of different strategies. *Phys. Rev. A*, 77:032322, 2008.
- [201] H. Nagaoka. A new approach to Cramér-Rao bounds for quantum state estimation. *IEICE Technical Report*, IT 89-42:9–14, 1989. Reprinted in Ref. [133].
- [202] H. Nagaoka. On the parameter estimation problem for quantum statistical models. In *Proceedings of 12th Symposium on Information Theory and its Applications (SITA)*, pages 577–582, 1989. Reprinted in Ref. [133].
- [203] H. Nagaoka. A generalization of the simultaneous diagonalization of Hermitian matrices and its relation to quantum estimation theory. In M. Hayashi, editor, *Asymptotic Theory of Quantum Statistical Inference*, Singapore, 2005. World Scientific. Originally published in *Trans. Jap. Soc. Indust. Appl. Math.* **1**, 43–56, 1991 (in Japanese).
- [204] J. von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, New Jersey, 1955. Translated from the German edition by R. T. Beyer.
- [205] R. G. Newton and B.-L. Young. Measurability of the spin density matrix. *Ann. Phys.*, 49(3):393–402, 1968.
- [206] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, UK, 2000.
- [207] J. L. O’Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White. Quantum process tomography of a controlled-not gate. *Phys. Rev. Lett.*, 93:080502, 2004.

## Bibliography

---

- [208] M. G. A. Paris and J. Řeháček, editors. *Quantum State Estimation*, volume 649 of *Lecture Notes in Physics*. Springer, Berlin, 2004.
- [209] J. L. Park and W. Band. A general theory of empirical state determination in quantum physics: Part I. *Found. Phys.*, 1:211–226, 1971.
- [210] T. Paterek, M. Pawłowski, M. Grassl, and Č Brukner. On the connection between mutually unbiased bases and orthogonal Latin squares. *Phys. Scr.*, T140:014031, 2010.
- [211] W. Pauli. Die allgemeinen Prinzipien der Wellenmechanik. In H. Geiger and K. Scheel, editors, *Handbuch der Physik*, Berlin, 1933. Springer. English translation: W. Pauli, *General Principles of Quantum Mechanics*, Springer, Berlin, 1980.
- [212] A. M. Perelomov. *Generalized coherent states and their applications*. Springer, Berlin, 1986.
- [213] A. Peres. Neumark’s theorem and quantum inseparability. *Found. Phys.*, 20:1441–1453, 1990.
- [214] A. Peres and W. K. Wootters. Optimal detection of quantum information. *Phys. Rev. Lett.*, 66:1119–1122, 1991.
- [215] D. Petz. *An Invitation to the Algebra of Canonical Commutation Relations*, volume 2 of *Leuven Notes in Mathematical and Theoretical Physics*. Leuven University Press, Leuven, Belgium, 1990.
- [216] D. Petz. Monotone metrics on matrix spaces. *Linear Algebra Appl.*, 244:81–96, 1996.
- [217] D. Petz and L. Ruppert. Efficient quantum tomography needs complementary and symmetric measurements, 2010. Available at <http://arxiv.org/abs/1011.5210v2>.
- [218] D. Petz and L. Ruppert. Optimal quantum-state tomography with known parameters. *J. Phys. A: Math. Theor.*, 45(8):085306, 2012.
- [219] D. Petz and C. Sudár. Geometries of quantum states. *J. Math. Phys.*, 37(6):2662–2673, 1996.
- [220] A. R. Plastino, D. Manzano, and J. S. Dehesa. Separability criteria and entanglement measures for pure states of  $N$  identical fermions. *Europhys. Lett.*, 86:20005, 2009.
- [221] J. F. Poyatos, J. I. Cirac, and P. Zoller. Complete characterization of a quantum process: The two-bit quantum gate. *Phys. Rev. Lett.*, 78:390–393, 1997.
- [222] C. Procesi. *Lie Groups: An Approach through Invariants and Representations*. Springer, New York, 2007.
- [223] E. Prugovečki. Information-theoretical aspects of quantum measurement. *Int. J. Theor. Phys.*, 16:321, 1977.

## Bibliography

---

- [224] C. R. Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, 37(3):81–91, 1945.
- [225] R. Raussendorf and H. J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86(22):5188–5191, 2001.
- [226] J. Řeháček, B.-G. Englert, and D. Kaszlikowski. Minimal qubit tomography. *Phys. Rev. A*, 70:052321, 2004.
- [227] J. Řeháček and Z. Hradil. Invariant information and quantum state estimation. *Phys. Rev. Lett.*, 88:130401, 2002.
- [228] J. Řeháček, Z. Hradil, and M. Ježek. Iterative algorithm for reconstruction of entangled states. *Phys. Rev. A*, 63(4):040303(R), 2001.
- [229] J. Řeháček, Z. Hradil, E. Knill, and A. I. Lvovsky. Diluted maximum-likelihood algorithm for quantum tomography. *Phys. Rev. A*, 75(4):042108, 2007.
- [230] J. M. Renes. *Frames, Designs, and Spherical Codes in Quantum Information Theory*. PhD thesis, The University of New Mexico, 2004.
- [231] J. M. Renes. Equiangular spherical codes in quantum cryptography. *Quantum Inf. Comput.*, 5:81, 2005.
- [232] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves. Symmetric informationally complete quantum measurements. *J. Math. Phys.*, 45:2171, 2004. Supplementary information including the fiducial kets available at <http://www.cquic.org/papers/reports/>.
- [233] A. Roy. *Complex lines with restricted angles*. PhD thesis, University of Waterloo, 2006.
- [234] A. Roy and A. J. Scott. Weighted complex projective 2-designs from bases: Optimal state determination by orthogonal measurements. *J. Math. Phys.*, 48:072110, 2007.
- [235] A. Roy and A. J. Scott. Unitary designs and codes. *Des. Codes Cryptogr.*, 53(1):13–31, 2009.
- [236] A. Royer. Measurement of the Wigner function. *Phys. Rev. Lett.*, 55:2745–2748, 1985.
- [237] A. Royer. Measurement of quantum states and the Wigner function. *Found. Phys.*, 19:3–32, 1989.
- [238] P. Rungta, V. Bužek, C. M. Caves, M. Hillery, and G. J. Milburn. Universal state inversion and concurrence in arbitrary dimensions. *Phys. Rev. A*, 64:042315, 2001.
- [239] P. Rungta, W. J. Munro, K. Nemoto, P. Deuar, G. J. Milburn, and C. M. Caves. Qudit entanglement. In H. J. Carmichael, R. J. Glauber, and M. O. Scully, editors, *Directions in Quantum Optics: A Collection of Papers Dedicated to the Memory of Dan Walls*, page 149. Springer, Berlin, 2000.

## Bibliography

---

- [240] R. Salazar, D. Goyeneche, A. Delgado, and C. Saavedra. Constructing symmetric informationally complete positive-operator-valued measures in Bloch space. *Phys. Lett. A*, 376(4):325 – 329, 2012.
- [241] V. Scarani, S. Iblisdir, N. Gisin, and A. Acín. Quantum cloning. *Rev. Mod. Phys.*, 77:1225–1256, 2005.
- [242] E. Schrödinger. Die gegenwärtige Situation in der Quantenmechanik. *Die Naturwissenschaften*, 23:807–812, 823–828, 844–849, 1935.
- [243] J. Schwinger. Unitary Operator Bases. *Proc. Natl. Acad. Sci. USA*, 46(4):570–579, 1960.
- [244] A. J. Scott. Tight informationally complete quantum measurements. *J. Phys. A: Math. Gen.*, 39:13507, 2006.
- [245] A. J. Scott and M. Grassl. Symmetric informationally complete positive-operator-valued measures: A new computer study. *J. Math. Phys.*, 51:042203, 2010. Supplementary information including the fiducial kets available at <http://arxiv.org/abs/0910.5784>.
- [246] A. J. Scott, J. Walgate, and B. C. Sanders. Optimal fingerprinting strategies with one-sided error. *Quantum Inf. Comput.*, 7:243–264, 2007.
- [247] M. O. Scully, B.-G. Englert, and H. Walther. Quantum optical tests of complementarity. *Nature*, 351(6322):111–116, 1991.
- [248] P. D. Seymour and T. Zaslavsky. Averaging sets: A generalization of mean values and spherical designs. *Adv. Math.*, 52:213, 1984.
- [249] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White. Efficient measurement of quantum dynamics via compressive sensing. *Phys. Rev. Lett.*, 106:100401, 2011.
- [250] A. Shapere and F. Wilczek, editors. *Geometric Phases in Physics*. World Scientific, Singapore, 1989.
- [251] D. A. Sibley. Certain finite linear groups of prime degree. *J. Algebra*, 32:286, 1974.
- [252] P. B. Slater. Increased efficiency of quantum state estimation using non-separable measurements. *J. Phys. A: Math. Gen.*, 34(35):7029–7046, 2001.
- [253] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Phys. Rev. Lett.*, 70:1244–1247, 1993.
- [254] R. Tarrach and G. Vidal. Universality of optimal measurements. *Phys. Rev. A*, 60:R3339–R3342, 1999.

## Bibliography

---

- [255] Y. S. Teo, H. Zhu, and B.-G. Englert. Product measurements and fully symmetric measurements in qubit-pair tomography: A numerical study. *Opt. Commun.*, 283:724–729, 2010.
- [256] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček, and Z. Hradil. Quantum-state reconstruction by maximizing likelihood and entropy. *Phys. Rev. Lett.*, 107:020404, 2011.
- [257] A. Uhlmann. The “transition probability” in the state space of a  $*$ -algebra. *Rep. Math. Phys.*, 9(2):273–279, 1976.
- [258] G. Valiente. *Algorithms on Trees and Graphs*. Springer, Berlin, 2002.
- [259] G. Vidal, J. I. Latorre, P. Pascual, and R. Tarrach. Optimal minimal measurements of mixed states. *Phys. Rev. A*, 60:126, 1999.
- [260] K. Vogel and H. Risken. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A*, 40:2847–2849(R), 1989.
- [261] V. I. Voloshin. *Introduction to Graph and Hypergraph Theory*. Nova Kroschka Books, 2009.
- [262] Z.-W. Wang, Y.-S. Zhang, Y.-F. Huang, X.-F. Ren, and G.-C. Guo. Experimental realization of direct characterization of quantum dynamics. *Phys. Rev. A*, 75:044304, 2007.
- [263] L. R. Welch. Lower bounds on the maximum cross correlation of signals. *IEEE Trans. Inf. Theory*, 20(3):397–399, 1974.
- [264] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, 1989.
- [265] R. F. Werner. Optimal cloning of pure states. *Phys. Rev. A*, 58:1827–1832, 1998.
- [266] H. Weyl. *The Theory of Groups and Quantum Mechanics*. Methuen & co. ltd., London, 1931. Translated from the second (revised) German edition by H. P. Robertson.
- [267] A. G. White, D. F. V. James, P. H. Eberhard, and P. G. Kwiat. Nonmaximally entangled states: Production, characterization, and utilization. *Phys. Rev. Lett.*, 83:3103–3107, 1999.
- [268] W. K. Wootters. Statistical distance and Hilbert space. *Phys. Rev. D*, 23:357–362, 1981.
- [269] W. K. Wootters. Quantum mechanics without probability amplitudes. *Found. Phys.*, 16(4):391–405, 1986.
- [270] W. K. Wootters. A Wigner-function formulation of finite-state quantum-mechanics. *Ann. Phys.*, 176(1):1–21, 1987.

## Bibliography

---

- [271] W. K. Wootters. Quantum measurements and finite geometry. *Found. Phys.*, 36:112, 2006.
- [272] W. K. Wootters and B. D. Fields. Optimal state-determination by mutually unbiased measurements. *Ann. Phys.*, 191:363, 1989.
- [273] H. P. Yuen and V. W. S. Chan. Noise in homodyne and heterodyne detection. *Opt. Lett.*, 8(3):177–179, 1983.
- [274] H. P. Yuen and M. Lax. Multiple-parameter quantum estimation and measurement of nonselfadjoint observables. *IEEE Trans. Inf. Theory*, 19(6):740–750, 1973.
- [275] G. Zauner. Quantum designs: Foundations of a noncommutative design theory. *Int. J. Quant. Inf.*, 9:445–507, 2011.
- [276] W.-M. Zhang, D. H. Feng, and R. Gilmore. Coherent states: Theory and some applications. *Rev. Mod. Phys.*, 62:867–927, 1990.
- [277] H. Zhu. unpublished.
- [278] H. Zhu. Regrouping phenomena of SIC POVMs covariant with respect to the Heisenberg–Weyl group. Presented at APS March meeting 2011, Dallas, Texas, USA, March 2011.
- [279] H. Zhu. SIC POVMs and Clifford groups in prime dimensions. *J. Phys. A: Math. Theor.*, 43:305305, 2010.
- [280] H. Zhu, L. Chen, and M. Hayashi. Additivity and non-additivity of multipartite entanglement measures. *New J. Phys.*, 12(8):083002, 2010.
- [281] H. Zhu and B.-G. Englert. Quantum state tomography with fully symmetric measurements and product measurements. *Phys. Rev. A*, 84:022327, 2011.
- [282] H. Zhu, Y. S. Teo, and B.-G. Englert. Minimal tomography with entanglement witnesses. *Phys. Rev. A*, 81:052339, 2010.
- [283] H. Zhu, Y. S. Teo, and B.-G. Englert. Two-qubit symmetric informationally complete positive-operator-valued measures. *Phys. Rev. A*, 82:042308, 2010.
- [284] K. Życzkowski and H.-J. Sommers. Induced measures in the space of mixed quantum states. *J. Phys. A: Math. Gen.*, 34:7111, 2001.



# Index

- $\mathcal{D}$ -invariant, 117
- $p$ -group, 246
- $t$ -design, 232
  - weighted, 232
- 2-design
  - mutually unbiased bases (MUB), 232
  - SIC POM, 232
  - weighted, 38
- adaptive measurement, 5, 77
  - comparison with nonadaptive schemes, 107
  - degenerate two-level system, 103
  - optimal quantum state estimation, 92
  - two-step adaptive, 82
- adjacency matrix, 255
- angle matrix, 203, 212
- angle tensor, 203
- Appleby, 153, 183
  - Clifford group, 153, 164
  - equivalence criterion for SIC POMs, 155, 204
  - HW group, 161
  - SIC POMs in dimension three, 176
- automorphism, 256
- automorphism group, 256
- balanced measurement, 67
  - SIC POM and MUB, 67
- Bayesian approach, 17
- Bayesian mean estimation (BME), 25
- Born rule, 18
- bound
  - Cramér–Rao (CR), 27
  - generalized Gill–Massar (GM), 137
  - Gill–Massar (GM), 85
  - joint, 138
  - quantum Cramér–Rao (CR), 80, 115
  - RLD, 115
  - Scott, 38
  - SLD, 81
  - Welch, 152
- Bures distance, 231
- SLD Fisher information, 82, 231
- canonical form of a graph, 255
- canonical order-3 Clifford unitary transformation, 153
- central-limit theorem, quantum, 16, 114
- Clifford group, 153, 160, 247
  - normalizer of, 190, 251
  - structure, 163
- coherent measurement, 6, 127
  - complementarity polynomial, 128, 144
  - covariant, 142, 148
  - GMT, 131, 144
  - optimal, 148
  - quantum state estimation, 122
  - qubit state estimation, 140, 146
- coherent state, 6, 126
  - conjecture, 132
- collective measurement, 6, 20
  - asymptotic state estimation, 115
  - coherent measurement, *see* coherent measurement
  - quantum state estimation, 113
  - qubit state estimation, 136
- collineation group, 156
- commutation superoperator, 117
- complementarity polynomial, 128, 131
  - qubit, 144
- complementarity principle, 1, 85
- conjecture
  - balanced measurement, 67
  - coherent state, 132
  - extremal Fisher information matrix, 97
  - Gill–Massar trace (GMT), 132
  - HW covariant SIC POM, 181
- conjecture of
  - Slater, 145
  - Zauner, 153, 154
- covariant measurement, 68
  - covariant coherent measurement, 142, 148

- Cramér–Rao (CR) bound, 2, 27, *see also* quantum Cramér–Rao (CR) bound
- displacement operator, 165
- entangled measurement, 21
- entangled state, 20
- entanglement, 20
- equiangular lines, 152
  - Welch bound, 152
- equivalence problem, 8, 156, 200
- extended automorphism group, 256
- extended Clifford group, 163
- extended special linear group, 163
- extended symmetry group, 157
- Fano, 11, 22
- fidelity, 1, 17, 230
- fiducial measurement, 31
- fiducial POM, 7
- fiducial state, 7
  - dimension eight, 224
  - dimension four, 184
  - dimension three, 176
  - Hoggar lines, 221
  - known solutions, 152
  - qubit, 168
- figures of merit
  - Bures distance, 231
  - fidelity, 230
  - Hilbert–Schmidt (HS) distance, 229
  - trace distance, 229
- Fisher information, 27
- Fisher information matrix, 2, 28
  - $U_\rho$ -invariant, 93, 103
  - coherent measurement, 128, 140
  - extremal, 95
  - frame superoperator, 63
  - GM inequality, 86
  - mutually unbiased measurements on a qubit, 75
  - RLD bound, 116, 120
  - SLD bound, 81
- frame superoperator, 34, 35
  - covariant measurement, 68
  - Fisher information matrix, 63
  - optimal reconstruction, 60
  - product measurement, 51
- Gaussian
  - approximation, 36
  - unitary ensemble, 39
- general linear group, 123
- generalized Bloch vector (GBV), 192
- generalized Gill–Massar (GM) bound, 137
- generalized measurement, 18
- generalized Pauli group, *see* HW group
- geometric phase, 178, 203
- Gill–Massar (GM), 15
  - bound, 5
    - generalized, 137
    - qubit, 88
    - scaled MSB, 89
    - scaled MSH, 90, 91
    - scaled WMSE, 88
  - inequality, 87
  - theorem, 86
  - trace, *see* GMT
- GMT, 5, 86
  - asymptotic maximum, 121, 135, 144
  - coherent measurement, 131
  - complementarity polynomial, 131
  - maximum in qubit state estimation, 145
  - maximum over collective measurements when  $\rho = 1/d$ , 135
  - Slater’s conjecture, 145
- graph, 205, 255
- graph automorphism problem, 202, 256
- graph isomorphism problem, 202, 256
- graph-theoretic approach, 10, 200
- Grassl, 153–155, 223
- group
  - $p$ -group and Sylow  $p$ -subgroup, 246
  - (extended) Clifford group, *see* Clifford group
  - (extended) automorphism group, 256
  - (extended) special linear group, 163
  - (extended) symmetry group, 157
  - general linear group, 123
  - generalized Pauli group, *see* HW group
  - Heisenberg–Weyl (HW) group, *see* HW group

## Index

---

- index group, 245
- symmetric group, 123
- three-qubit Pauli group, 162
- group covariance, 151, 157
- groups that can generate SIC POMs, 159, 168
  
- Hayashi, 15, 16, 114
- hedged maximum-likelihood estimation (HMLE), 13, 26
- height of a partition, 123
- Heisenberg–Weyl group, *see* HW group
- Helstrom, 11, 15, 81
- highest-weight state, 125
- Hilbert–Schmidt (HS) distance, 1, 229
- Hoggar lines, 9, 10, 155, 162, 196, 226
  - fiducial state, 221
  - group covariance and symmetry, 180, 220
- Holevo, 15
  - bound, 15, 16, 117
  - commutation superoperator, 117
  - lemma, 116
- homogeneous complementarity polynomial, 131
- Hradil, 13, 24, 58
- HW covariant SIC POM, 211
  - beyond prime dimensions, 180
  - dimension three, 171, 212
  - equivalence relation, 171, 175, 177, 181, 215
  - prime dimensions, 8, 167, 169
  - qubit, 168
  - regrouping phenomenon, 189
  - symmetry group, 171, 175, 178, 180, 185, 215
  - two-qubit, 9, 183, 191
- HW group, 7, 160
  - in the Clifford group, 252, 253
  - standard, 176, 253
- hypergraph, 205, 256
  - $k$ -uniform, 256
  
- index group, 245
- individual measurement, 20
- informationally complete (IC), 3, 19
  - minimal, 20
  - tight, 31, 36
- informationally overcomplete measurement, 4, 20, 57, 59
- isomorphism, 255
- isotropic measurement, 40, 41, 45
- Ivanović, 12, 58
  
- joint bound, 138
- joint SIC POM, 4, 32, 50
  
- lemma
  - $N$ -partite symmetric state, 241
  - a matrix inequality, 60, 233
  - bipartite antisymmetric state, 242
  - Gill–Massar trace (GMT), 132
  - HW covariant SIC POM, 169
  - orbits of states in a SIC POM, 158
  - SLD Fisher information matrix, 129
- lemma of Holevo, 116
- likelihood functional, 23
- linear state reconstruction, 22
- linear state tomography, 3, 34
- local asymptotic normality, 16, 114
- log likelihood functional, 23
  
- maximum entropy (ME) principle, 13
- maximum-likelihood (ML)
  - estimation, 13, 23
  - estimator, 24, 28
  - method, 63
  - principle, 23
- mean square Bures distance (MSB), 2, *see also* scaled MSB
- mean square error (MSE), 31, *see also* scaled MSE
  - matrix, 28
  - weighted, 2
- mean square Hilbert Schmidt distance (MSH), 1, *see also* scaled MSH
- measurement, 18
  - adaptive measurement, 77
  - coherent measurement, *see* coherent measurement
  - collective measurement, *see* collective measurement
  - covariant measurement, 68
  - entangled measurement, 21
  - individual measurement, 20

- informationally complete (IC) measurement, 19
- informationally overcomplete measurement, 20, 57
- isotropic measurement, 40
- mutually unbiased measurements, *see* mutually unbiased measurements
- probability operator measurement (POM), 19
- product measurement, 21, 50
- separable measurement, 21
- symmetric informationally complete (SIC) measurement, 3, *see also* SIC POM
- tight informationally complete (IC) measurement, 36
- measurement operator, 18
- monomial, 170
- monotone Riemannian metric, 71, 78
- mutually unbiased bases (MUB), 3, 64
- mutually unbiased measurements, 3, 12, 58, 64
  - balanced measurement, 67
  - MSE matrix, 66
  - optimality in state estimation, 66
  - qubit state estimation, 48, 74, 89
- Nagaoka, 15, 16
- nice error basis, 245
  - SIC POM, 159, 180, 216, 223, 226
- nice graph, 210
- normalizer of the Clifford group, 190, 251
- open questions, *see also* conjecture
  - quantum state estimation, 111, 150
  - SIC POM, 227
- ordered partition, 206
- partition, 123
  - height, 123
  - ordered, 206
- Pauli-exclusion principle, 134
- positive-operator-valued measure (POVM), *see* probability operator measurement (POM)
- probability operator measurement (POM), 1, 19
- product measurement, 21, 50
  - product SIC POM, 4, 32
    - bipartite, 50
    - multipartite, 54
  - projective measurement, 19
  - pure-state limit
    - adaptive measurement, 235, 238
    - asymptotic state estimation, 120, 121
    - coherent measurement, 143, 145, 147
    - covariant measurement, 71
  - pure-state models, 16
- quantum central-limit theorem, 16, 114
- quantum cloning, 17
- quantum Cramér–Rao (CR) bound, 15, 80, *see also* SLD bound *and* RLD bound
- quantum estimation theory, 11, 78, 114
- quantum Fisher information, 80, *see also* SLD Fisher information *and* RLD Fisher information
- quantum process tomography (QPT), 14
- quantum state estimation, 1
  - adaptive measurement, 77
  - collective measurement, 113
    - asymptotic limit, 115
    - coherent measurement, 122
  - nonadaptive measurement
    - linear reconstruction, 31
    - optimal reconstruction, 57
  - open problems, 111, 150
- quantum state tomography, *see* quantum state estimation
- qubit state estimation
  - adaptive measurement, 88
  - collective measurement, 136
  - nonadaptive measurement
    - linear reconstruction, 45
    - optimal reconstruction, 72
- quorum, 11
- random-matrix theory, 3, 38
- reconstruction operator, 35, 201
  - canonical, 35
  - informationally incomplete measurement, 62
  - linear state reconstruction, 22
  - optimal, 60

## Index

---

- product measurement, 51
- tight IC measurement, 37
- recursive ordered partition (ROP), 208
- reference vertex, 206
- regrouping phenomenon of SIC POMs, 9, 189
- Renes, 153, 155
- right logarithmic derivative (RLD), 15, 116
- RLD bound, 116
  - scaled GMT, 121
  - scaled MSB, 120
  - scaled MSE, 120
  - scaled WMSE, 116
- RLD Fisher information, 116, 118, 119
- ROP, 208
  
- scaled mean HS distance, 40, 42, 43
- scaled mean trace distance, 39
  - isotropic measurement, 42, 48
  - product SIC POM, 52, 54
  - qubit tomography, 48
  - SIC POM, 43
- scaled MSB
  - approximate saturation of the GM bound, 100
  - covariant measurement, 70
  - discontinuity at the boundary of the state space, 237
  - GM bound, 89
  - optimal adaptive measurement, 109
  - qubit state estimation with collective measurements, 146
  - RLD bound (saturated in the asymptotic limit), 120
- scaled MSE
  - covariant measurement, 70
  - linear reconstruction, 35
  - mutually unbiased measurements, 66
  - optimal reconstruction, 61
  - product SIC POM, 51, 54
  - qubit, 46, 73, 74
  - RLD bound (saturated in the asymptotic limit), 120
  - tight IC measurement, 37
- scaled MSE matrix, 28, 35, 120
  - qubit, 46, 75
  
- scaled MSH
  - GM bound, 90, 91
  - optimal adaptive measurement, 108
  - qubit state estimation with collective measurements, 146
- Schur symmetric polynomial, 124, 140
- Schur–Weyl duality, 123
- Scott
  - bound, 38
  - SIC POM, 153, 154
  - tight IC measurement, 3, 31, 36
- separable measurement, 21
  - GMT, 86
- separable state, 20
- Sibley’s theorem, 170
- SIC POM, 7, 151
  - equiangular lines, 152
  - equivalence *or* equivalent, 153, 159
  - equivalence problem, 8, 156, 200
  - fiducial state, *see* fiducial state
  - graph-theoretic approach, 200
  - group covariance, 151, 157, 159, 168
  - Hoggar lines, *see* Hoggar lines
  - HW covariant, *see* HW covariant SIC POM
  - known solutions, 152
  - MUB, 49, 67, 74, 151
  - nice error basis, 159, 180, 216, 223, 226
  - numerical search, 153, 154, 225
  - open questions, 67, 227
  - orbit, 153, 159
  - quantum state estimation, 42, 48, 50, 74
  - symmetry group, 157
  - symmetry problem, 8, 156, 200
    - algorithm, 206
  - tight IC measurement, 31, 38
  - unitary symmetry and permutation symmetry, 201
- skew isomorphism, 255
- Slater’s conjecture, 145
- Slater-determinant state, 126, 240, 242
- SLD bound, 81
  - equivalent formulations, 83
  - scaled MSB, 85
  - scaled MSH, 85
  - scaled WMSE, 84

- SLD Fisher information, 81, 84, 118, 119
  - Bures distance, 82
- special linear group, 163
- standard HW group, 176, 253
- state reconstruction, 21
  - Bayesian mean estimation (BME), 25
  - hedged maximum-likelihood estimation (HMLE), 26
  - linear state reconstruction, 22, 34
  - maximum-likelihood estimation (MLE), 23
  - optimal reconstruction in the perspective of frame theory, 60
- strong group covariant, 157
- swap operator, 124, 129
- Sylow  $p$ -subgroup, 246
  - group covariant SIC POM, 169, 171, 180
- Sylow's theorem, 246
- symmetric group, 123
- symmetric informationally complete (SIC), 3, 151
- symmetric informationally complete probability operator measurement, *see* SIC POM
- symmetric logarithmic derivative (SLD), 11, 81, 83
- symmetry group of a SIC POM, 157
- symmetry problem, 8, 156, 200
- symplectic operator, 165
  
- theorem
  - asymptotic Fisher information matrix, 143, 244
  - asymptotic GMT, 135, 243
  - Gill–Massar trace (GMT), 133
  - group covariant SIC POM in a prime dimension, 168
  - groups that can generate SIC POMs, 159
  - normalizer of the Clifford group, 252
  - orbits and equivalence of SIC POMs, 159
  - symmetry of a SIC POM
    - beyond prime dimensions, 180
    - dimension three, 175
    - prime dimensions not equal to three, 171
    - trace of a Clifford operator, 248
- theorem of
  - Gill–Massar, 86
  - Sibley, 170
  - Sylow, 246
  - Zauner, 158
- three-qubit Pauli group, 162
- tight informationally complete (IC), 31, 36
  - quantum state estimation, 42
  - weighted 2-design, 38
- trace distance, 1, 32, 229
- triple product, 178, 185, 203
- two-step adaptive, 82
  
- unimodular, 157, 170
- von Neumann measurement, *see* projective measurement
  
- weighted  $t$ -design, 232
- weighted 2-design, 31, 38
- weighted 3-design, 41
- weighted mean square error (WMSE), 2, 29
  - convex optimization, 99
  - generalized GM bound, 137, 138
  - GM bound, 87
  - joint bound, 138, 139
  - RLD bound, 116
- Welch bound, 152
- Wigner function, 12, 13, 151
- Wigner semicircle law, 39
- Wootters, 12, 17, 58, 64, 113
  
- Zauner, 152, 220
  - conjecture, 153, 154
  - matrix, 165
  - theorem, 157
  - unitary transformation, 153, 165
- zero-eigenvalue problem, 25, 26