Regular Article - Theoretical Physics

Quantum tomography for collider physics: illustrations with lepton-pair production

John C. Martens[a], **John P. Ralston** , **J. D. Tapia Takaki**[b]

Department of Physics and Astronomy, The University of Kansas, Lawrence, KS 66045, USA

Abstract Quantum tomography is a method to experimentally extract all that is observable about a quantum mechanical system. We introduce quantum tomography to collider physics with the illustration of the angular distribution of lepton pairs. The tomographic method bypasses much of the field-theoretic formalism to concentrate on what can be observed with experimental data. We provide a practical, experimentally driven guide to model-independent analysis using density matrices at every step. Comparison with traditional methods of analyzing angular correlations of inclusive reactions finds many advantages in the tomographic method, which include manifest Lorentz covariance, direct incorporation of positivity constraints, exhaustively complete polarization information, and new invariants free from frame conventions. For example, experimental data can determine the entanglement entropy of the production process. We give reproducible numerical examples and provide a supplemental standalone computer code that implements the procedure. We also highlight a property of complex positivity that guarantees in a least-squares type fit that a local minimum of a χ^2 statistic will be a global minimum: There are no isolated local minima. This property with an automated implementation of positivity promises to mitigate issues relating to multiple minima and convention dependence that have been problematic in previous work on angular distributions.

1 Introduction

Tomography builds up higher-dimensional objects from lower-dimensional projections. Quantum tomography [1] is a strategy to reconstruct all that can be observed about a quantum physical system. After becoming a focal point of quantum computing, quantum tomography has recently been applied in a variety of domains [2–9].

The method of quantum tomography uses a known "probe" to explore an unknown system. Data is related directly to matrix elements, with minimal model dependence and optimal efficiency.

Collider physics is conventionally set up in a framework of unobservable and model-dependent scattering amplitudes. In quantum tomography these unobservable features are skipped to deal directly with observables. The unknown system is parameterized by a certain density matrix $\rho(X)$, which is model-independent. The probe is described by a known density matrix ρ(probe). The matrices are represented by numbers generated and fit to experimental data, not abstract operators. Quantum mechanics predicts an experiment will measure $\mathrm{tr}(\rho(\text{probe}) \cdot \rho(X))$, where tr is the trace. In many cases ρ(probe) is extremely simple: A 3×3 matrix, say. What will be observed is strictly limited by the dimension and symmetries of the probe. The powerful efficiency of quantum tomography comes from exploiting the probe's simplicity in the first steps. The description never involves more variables than will actually be measured.

We illustrate the advantages of quantum tomography with inclusive lepton-pair production. It is a relatively mature subject chosen for its pedagogical convenience. Despite the maturity of the subject, we discover new things. For example, the puzzling plethora of plethora of ad hoc invariant quantities is completely cleared up. We also find new ways to assist experimental data analysis. Positivity is a central issue overlooked in the literature, which we show how to control. Moreover, the tomography procedure carries over straightforwardly to many final states, including the inclusive production of charmonium, bottomonium, dijets, including boosted tops, HH, W^+W^-, ZZ [10–31]. Our practical guide to analyzing experimental data uses density matrices at each step and circumvents the more elaborate traditional theoretical formalism. We concentrate on making tools available to experimentalists. We give a step-by-step guide where density matrices stand as definite arrays of numbers, bypassing unnecessary formalism.

[a] e-mail: martens@ku.edu

[b] e-mail: daniel.tapia.takaki@cern.ch

Springer

2 The quantum tomography procedure applied to inclusive lepton-pair production

The tomography procedure reconstructs all that can be observed about a quantum physical system. For inclusive lepton-pair production, what can be observed is the invariant mass distribution, the lepton-pair angular distribution $dN/d\Omega$ and the polarization of the unknown intermediate state of the system, contained in $\rho(X)$.[1] In this section we reconstruct $dN/d\Omega$ and $\rho(X)$ from first principles using tomography. Structure functions and model-dependent assumptions as regards the intermediate state, common to the traditional formalism [32–36], do not appear.

The expert reader, who is accustomed to seeing some of these formulas derived, might note that the method of derivation is particularly simple. The particular steps we do *not* follow are to be noted. That also explains why some of the relations we find seem to have been overlooked in the past.

2.1 Kinematics

Consider inclusive production of a lepton pair with 4-momenta k, k' from the collision of two hadrons with 4-momenta P_A, P_B:

$$P_A P_B \rightarrow \ell^+(k)\ell^-(k') + \mathcal{X},$$

where $\mathcal{X}$ and the final-state lepton spins are unobserved and thus summed over. In the high energy limit $k^2 = k'^2 = 0$.

Let the total pair momentum $Q = k + k'$. The azimuthal distribution of total pair momenta in the lab frame is isotropic. Lepton-pair angular distributions are described in the pair rest frame defined event-by-event. In this frame the pair momenta are back-to-back and equal in magnitude. The frame orientation depends on the beam momenta and the pair total momentum.

Defining momentum[2] observables via a Lorentz-covariant frame convention allows calculations to be done in any frame. In its rest frame the total pair momentum $Q^\mu = (\sqrt{Q^2}, \vec{Q} = 0)$. A set of xyz spatial axes in this frame will be defined by three 4-vectors X^μ, Y^μ, Z^μ, satisfying

$$Q \cdot X = Q \cdot Y = Q \cdot Z = 0. \tag{1}$$

The frame vectors being orthogonal implies

$$X \cdot Y = Y \cdot Z = X \cdot Z = 0. \tag{2}$$

Taking $P_A = (1, 0, 0, 1)$, $P_B = (1, 0, 0, -1)$ (light-cone $\pm$ vectors), a frame satisfying the relations of Eqs. (1) and (2) is given by[3]

$$\tilde{Z}^\mu = P_A^\mu Q \cdot P_B - P_B^\mu Q \cdot P_A;$$

$$\tilde{X}^\mu = Q^\mu - P_A^\mu \frac{Q^2}{2Q \cdot P_A} - P_B^\mu \frac{Q^2}{2Q \cdot P_B};$$

$$\tilde{Y}^\mu = \epsilon^{\mu\nu\alpha\beta} P_{A\nu} P_{B\alpha} Q_\beta.$$

These frame vectors define the Collins–Soper (CS) frame.[4] The normalized frame vectors are

$$\left(X^\mu, Y^\mu, Z^\mu\right) = \left(\frac{\tilde{X}^\mu}{\sqrt{-\tilde{X} \cdot \tilde{X}}}, \frac{\tilde{Y}^\mu}{\sqrt{-\tilde{Y} \cdot \tilde{Y}}}, \frac{\tilde{Z}^\mu}{\sqrt{-\tilde{Z} \cdot \tilde{Z}}}\right).$$

To analyze data for each event labeled J:

$$\text{Compute} \quad Q_J = k_J + k'_J; \; \ell_J = k_J - k'_J; \; (X_J^\mu, Y_J^\mu, Z_J^\mu);$$

$$\vec{\ell}_{XYZ,J} = (X_J \cdot \ell_J, \; Y_J \cdot \ell_J, \; Z_J \cdot \ell_J);$$

$$\hat{\ell}_J = \vec{\ell}_{XYZ, J}/\sqrt{\vec{\ell}_{XYZ, J} \cdot \vec{\ell}_{XYZ, J}}. \tag{3}$$

In fact, $\hat{\ell}_J = (\sin\theta\cos\phi, \; \sin\theta\sin\phi, \; \cos\theta)_J$, where θ, ϕ are the polar and azimuthal angles of one (e.g. plus-charge) lepton in the rest frame of Q. The meaning of a "Lorentz-invariant $\cos\theta$" is a scalar $Z_\mu(k - k')^\mu$ which becomes $\hat{z} \cdot (k \hat{-} k')$ in the rest frame of Q.

2.2 The angular distribution, in terms of the probe and target density matrices

The standard amplitude for inclusive production of a fermion–antifermion pair of spin s, s' has a string of gamma-matrices contracted with final-state spinors $v_a(k's')$, $\bar{u}_b(k, s)$. When the amplitude is squared, these factors appear bi-linearly, as in

$$u_a(ks)\bar{u}_{a'}(ks) = (1/2)\left[(\not{k} + m)(1 + \gamma_5\not{s})\right]_{aa'}.$$

Summing over unobserved s and dropping $m\delta_{aa'}$, a form of density matrix appears:

$$\sum_s u_a(ks)\bar{u}_{a'}(ks) \rightarrow k_\mu \gamma_{aa'}^\mu.$$

The Feynman rules for the density matrix of two relativistic final-state fermions (or antifermions, or any combination) is a factor given by

$$\rho_{aa',\,bb'}(k,\,k') \rightarrow \not{k}_{aa'}\not{k}'_{bb'} \qquad (4)$$

This fundamental equality is not present in pure state quantum systems. There is no spinor corresponding to a fermion averaged over initial spins, nor to a fermion summed over final spins.

As shown in the appendix, the rest of the cross section appears in the target density matrix $\rho(X)$, which must have four indices to contract with the probe indices:

$$d\sigma \sim \sum_{aa'bb'} \rho_{aa',\,bb'}(k,\,k')\rho_{aa',\,bb'}(X)d\text{LIPS}$$
$$= \text{tr}\left(\rho(k,\,k')\rho(X)\right) d\text{LIPS}, \qquad (5)$$

where dLIPS is the Lorentz-invariant phase space.

Note that $u_a(ks)\bar{u}_{a'}(ks)$ is not positive definite since the Dirac adjoint $\bar{u}_{a'}(ks) = (u^\dagger(ks)\gamma_0)_a$ has a factor of γ_0, introduced by convention. Removing it, $\sum_s u_a(ks)u^\dagger_{a'}(ks)$ becomes positive by inspection. (Any matrix of the form $M \cdot M^\dagger$ has positive eigenvalues.) $\rho(k,\,k')$ as written is not normalized, because the Feynman rules shuffle spinor normalizations into overall factors. To make the arrow in Eq. (4) into an equality, multiply on the right by γ_0 twice, and standardize the normalizations. The same steps applied to $\rho_{aa',\,bb'}(X)$ cancels the γ_0 factors. The result is that the probability to find two fermions has the fundamental quantum mechanical form[5] $P(k,\,k') = \text{tr}(\rho(k,\,k')\rho(X))$.

The left side of Eq. (5) is $d\sigma(k,\,k')$, the same as the joint probability $P(Q,\,\ell|\text{init})$ where init are the initial-state variables. The phase space for two leptons converts as

$$k_0 k'_0 \frac{d\sigma}{d^3k\, d^3k'} = \frac{d\sigma}{d^4Q d\Omega}.$$

We can write

$$P(Q,\,\ell|\text{init}) = P(\ell|Q,\,\text{init})P(Q|\text{init}).$$

Here $P(Q|\text{init}) = d\sigma/d^4Q$, and $P(\ell|Q,\,\text{init}) = dN/d\Omega$ is the conditional probability to find ℓ given Q and the initial state. This factorization is general and unrelated to one-boson exchange, parton model, or other considerations. Since $P(\ell|Q,\,\text{init})$ is a probability, quantum mechanics predicts it is a trace:

$$\frac{dN}{d\Omega} = \frac{1}{\sigma}\frac{d\sigma}{d\Omega} = P(\ell|Q,\,\text{init}) = \frac{3}{4\pi}\text{tr}(\rho(\ell)\rho(X)), \qquad (6)$$

where tr indicates the trace, $d\Omega = d\cos\theta \cdot d\phi$, and $\rho(\ell)$, the probe, is a 3×3 matrix to be defined momentarily which depends only on the directions $\hat{\ell}_J$. The target hadronic system is represented by $\rho(X)$. Since the probe $\rho(\ell)$ is a 3×3 matrix, then $\rho(X)$ is a 3×3 matrix of numbers.[6]

The description has just been reduced from $\rho_{aa',\,bb'}(k,\,k')$, a Dirac tensor with 4^4 possible matrix elements, to a 3×3 Hermitian matrix with 8 independent elements, since $\text{tr}(\rho(\ell)) = 1$ is one condition. Equation 6 is the most general angular distribution that can be observed. It is valid for like sign and unlike sign pairs, and assumes no model for how the pairs are produced. The Dirac form (and Dirac traces) is over-complicated, because describing every possible exclusive reaction for every possible in and out state is over-achieved in the formalism.

2.3 The probe matrix

The probe matrix $\rho(\ell)$ is given by

$$\rho_{ij}(\ell) = \frac{1+a}{3}\delta_{ij} - a\hat{\ell}_i\hat{\ell}_j - \imath b\epsilon_{ijk}\hat{\ell}_k, \qquad (7)$$

which is derived in the appendix. The Standard Model predicts only two parameters, a and b. If on-shell lepton helicity is conserved (as in lowest-order production by a minimally coupled vector boson) then $a = 1/2$ and $b = c_A c_V$. The latter is not a prediction but a definition. If the production is parity-symmetric then $c_A = 0$. The only non-trivial prediction of the Standard Model is the value of $c_A c_V$. Lowest-order production by Z bosons predicts $b = \sin^2\theta_W \sim 0.22$.

More generally, the probe matrix itself represents a reduced system that is unknown a priori. It should be determined experimentally. Consider the angular distribution of $e^+e^- \rightarrow \mu^+\mu^-$. Let $\rho(e;\,\hat{z})$ describe electrons with parameters a_e, b_e colliding along the z axis. Let $\rho(\mu;\,\hat{\ell})$ describe muons with parameters a_μ, b_μ emerging along direction $\hat{\ell}$. A short calculation using Eq. (7) twice gives[7]

$$\frac{3}{4\pi}\text{tr}\left(\rho(e;\,\hat{z})\rho(\mu;\,\hat{\ell})\right)$$
$$= \frac{3}{4\pi}\left(\frac{1}{3} + 2b_e b_\mu \hat{\ell} \cdot \hat{z} + a_e a_\mu((\hat{z} \cdot \hat{\ell})^2 - 1/3)\right),$$
$$= \frac{3}{4\pi}\left(\frac{1}{3} + 2b_e b_\mu \cos\theta + a_e a_\mu(\cos^2\theta - 1/3)\right). \qquad (8)$$

---

[5] We remind the reader that the phase space factors dLIPS originate in further organizational steps computing the quantum mechanical transition probability per volume per time, which afterwards restore the phase space factors.

[6] This is a more general statement than enumerating "structure functions".

[7] This may be a new result, which goes beyond what is known from one-boson exchange with or without radiative corrections. The production details can only renormalize the parameters.

Fitting experimental data will give $a_e a_\mu$ and $b_e b_\mu$. If lepton universality is assumed the probe $\rho(\mu; \hat{\ell})$ has measured the probe $\rho(e; \hat{z})$.

2.4 How tomography works: $dN/d\Omega$ as a function of $\rho(\ell)$, $\rho(X)$

Let $\hat{G}_\ell$ be a set of probe operators, with expectation values $\langle G_\ell \rangle = \mathrm{tr}(G_\ell \rho(X))$. The trace defines the Hilbert–Schmidt inner product of operators. The condition for operators (matrices) to be orthonormal is

$$\mathrm{tr}(G_\ell G_k) = \delta_{\ell k} \quad \text{orthonormal matrices.} \tag{9}$$

There are $N^2 - 1$ orthonormal $N \times N$ Hermitian operators, not including the identity. When a complete set of probe operators has been measured, the density matrix is tomographically reconstructed from observables as

$$\rho(X) = \sum_\ell G_\ell \mathrm{tr}(G_\ell \rho) = \sum_\ell G_\ell \langle G_\ell \rangle .$$

For a pure state density matrix, there exists a basis $\{G_\ell\}$ such that only one term appears in the sum over ℓ. Then $\rho_{\mathrm{pure}} = |\psi\rangle\langle\psi|$, and $|\psi\rangle$ is reconstructed as the eigenvector of ρ_{pure}.

Each orthogonal probe operator measures the corresponding component of the unknown system, and is classified by its transformation properties. For angular distributions the transformations of interest are rotations. $\rho(\ell)$ contains tensors transforming like spin-0, spin-1 and spin-2. Each tensor of a given type is orthogonal to the others.

Organizing transformation properties simplifies things significantly. Recall the general form of $\rho(\ell)$, from Eq. (7). The most general form for $\rho(X)$ that is observable will have the same general expansion, with new parameters:

Probe: $\quad \rho_{ij}(\ell) = \dfrac{1}{3}\delta_{ij} + b\hat{\ell} \cdot \vec{J}_{ij} + aU_{ij}(\hat{\ell});$

where $\quad U_{ij}(\hat{\ell}) = \dfrac{\delta_{ij}}{3} - \hat{\ell}_i \hat{\ell}_j = U_{ji}(\ell); \ tr(U(\ell)) = 0;$
$$\tag{10}$$

System: $\quad \rho_{ij}(X) = \dfrac{1}{3}\delta_{ij} + \dfrac{1}{2}\vec{S} \cdot \vec{J}_{ij} + U_{ij}(X);$

where $\quad U(X) = U^T(X); \quad tr(U(X)) = 0. \tag{11}$

These formulas reiterate Eq. (7) while identifying $(J_k)_{ij} = -\iota\epsilon_{ijk}$ as the generator of the rotation group in the 3×3 representation.[8] Upon taking the trace as an inner product, orthogonality selects each term in $\rho(X)$ that matches its counterpart

---

[8] The real Cartesian basis for $\vec{J}$ is being used because it is more transparent than the $J_z \to m$ basis that is an alternative. It would have complex parameters.

in $\rho(\ell)$. For example $\vec{J}$ is orthogonal to all the other terms except the same component of $\vec{J}$:

$$\frac{1}{2}\mathrm{tr}(J_i J_j) = \delta_{ij}; \quad \text{hence} \quad \frac{1}{2}\mathrm{tr}(\hat{\ell} \cdot \vec{J}\, \vec{S} \cdot \vec{J}) = \hat{\ell} \cdot \vec{S}.$$

Orthogonality makes it trivial to predict which density matrix terms can be measured by probe matrix terms. We call the matching of terms "the mirror trick".

We now make several relevant comments about Eqs. (10) and (11):

- All density matrices can be written as $1_{N \times N}/N$ to take care of the normalization, plus a traceless Hermitian part. The unit matrix is the spin-0 part and invariant under rotations. The only contribution of the 1 terms is $\mathrm{tr}(1 \times 1)/N^2 = 1/N$.
- The textbook density matrix spin vector $\vec{S}$ consists of those parameters coupled to the angular momentum operator. This is also called the spin-1 contribution. The quantum mechanical average angular momentum of the system is

$$\left\langle \vec{J} \right\rangle = \mathrm{tr}(\rho_X \vec{J}) = \vec{S}.$$

When the coordinates are rotated, the $\vec{J}$ matrices transform exactly so that $\vec{S}$ rotates like a vector under proper rotations, and a pseudovector under a change of parity.

- The last term of Eq. (10), the spin-2 part, is real, symmetric and traceless. By the mirror trick it can only communicate with a corresponding spin-2 term in $\rho(X)$ denoted $U_{ij}(X)$, which is real, symmetric and traceless. It can be considered a measure of angular momentum fluctuations:

$$\left\langle \frac{1}{2}\left(J_i J_j + J_j J_i\right) - \frac{1}{3}\vec{J}^2 \delta_{ij} \right\rangle = U_{ij}(X).$$

A common mistake assumes the quadrupole U should be zero in a pure "spin state". Actually a pure state with $|\vec{S}| = 1$ has a density matrix

$$\rho_{\mathrm{pure}, ij}(\vec{S}) = \frac{1}{2}(\delta_{ij} - \hat{S}_i \hat{S}_j) - \frac{\iota}{2}\epsilon_{ijk}\hat{S}_k. \tag{12}$$

For example, when $\vec{S} = \hat{z}$ the density matrix has one circular polarization eigenstate with eigenvalue unity, and two zero eigenvalues. Pure states exist with $\vec{S} = 0$: They have real eigenvectors corresponding to linear polarization. From the spectral resolution $\rho(X) = \sum_\alpha \lambda_\alpha |e_a\rangle\langle e_\alpha|$, there is no observable distinction between a density matrix and the occurrence of pure states $|e_a\rangle$ with probabilities λ_α, which are the density matrix eigenvalues.

- As it stands the U_{ij} matrices in Eqs. (10) and (11) have not been expanded in a complete set of symmetric, orthonormal 3×3 matrices. Regardless $\rho(X)$ can be fit

to data whether or not an expansion is done. The purpose of such work is to complete the classification process to assist with interpreting data. We sketch the steps here. Details are provided in an appendix. Let E_M be a basis of traceless orthonormal matrices where $U(\ell) = \sum_M \text{tr}(U(\ell)E_M)E_M$. This is the tomographic expansion of the probe. Choose E_M so the outputs are normalized real-valued spin-2 spherical harmonics $Y_M(\theta, \phi)$. The expansion of the unknown system will be $U(X) = \sum_M \text{tr}(\rho(X)E_M)E_M = \sum_M \rho_M(X)E_M$. By orthogonality the spin-2 contribution to the angular distribution will be

$$\frac{dN}{d\Omega} \sim \text{tr}(\rho(\ell)\rho(X))_{spin-2} \sim \sum_M \rho_M(X)Y_M(\theta, \phi).$$

Writing out the terms gives

$$\begin{aligned}
\frac{dN}{d\Omega} =& \frac{1}{4\pi} + \frac{3}{4\pi}S_x \sin\theta \cos\phi \\
&+ \frac{3}{4\pi}S_y \sin\theta \sin\phi + \frac{3}{4\pi}S_z \cos\theta \\
&+ c\rho_0 \left(\frac{1}{\sqrt{3}} - \sqrt{3}\cos^2\theta\right) - c\rho_1 \sin(2\theta)\cos\phi \\
&+ c\rho_2 \sin^2\theta \cos(2\phi) \\
&+ c\rho_3 \sin^2\theta \sin(2\phi) - c\rho_4 \sin(2\theta)\sin\phi. \quad (13)
\end{aligned}$$

The label X has been dropped in ρ_M and $c = 3/(8\sqrt{2}\pi)$. Since E_M transform like Y_M, the coefficients ρ_M transform under rotations like spin-2. That means $\rho_M \rightarrow$

not the same concept as yielding a positive cross section, and generally is a more restrictive set of relations.[9] If density matrices are not used it is quite straightforward to fit data yielding a positive cross section while violating positivity.

Fortunately positivity can be implemented by the Cholesky decomposition of ρ_X [38], which is discussed in the appendix. For the 3×3 case it is

$$\rho(X)(m) = M(m) \cdot M^\dagger(m);$$

$$M(m) = \frac{1}{\sqrt{\sum_k m_k^2}} \begin{pmatrix} m_1 & m_4 + im_5 & m_6 + im_7 \\ 0 & m_2 & m_8 + im_9 \\ 0 & 0 & m_3 \end{pmatrix}, \quad (14)$$

where the parameters $-1 \leq m_\alpha \leq 1$.

Event-by-event $\rho(\ell)$ is an array of numbers, and $\rho(X)(m)$ is an array of parameters. The results are combined to make the Jth instance of $tr(\rho_J(\ell)\rho(X)(m))$, where $\rho(X)(m)$ has been parameterized in Eq. (14). Fit the m_α parameters to the data set. For example, the log likelihood $\mathcal{L}$ of the set $J = 1 \cdots J_{\max}$ is

$$\begin{aligned}
\mathcal{L}(m) =& \sum_J^{J_{\max}} \log\left(\text{tr}\left(\rho^{(J)}(\ell) \cdot \rho(X)(m)\right)\right) \\
&+ J_{\max} log(3/4\pi). \quad (15)
\end{aligned}$$

Sample code available online[10] carries out these steps, returning the parameters m_α. The details of cuts and acceptance appear in fitting the numbers m_α using numbers for the lepton matrix $\rho(\text{lep})$ (not angles, nor trigonometric functions.) In one example with simulated Z-boson data we found

$$\rho_{\text{fit}}(X) = \begin{pmatrix} 0.5574 & 0.01399 - 0.07144i & -0.004026 + 0.013487i \\ 0.01399 + 0.07144i & 0.4422 & 0.003138 - 0.002670i \\ -0.004026 - 0.013487i & 0.003138 + 0.002670i & 0.0004268 \end{pmatrix}.$$

$R^{(2)}_{MM'}\rho_{M'}(X)$, where $R^{(2)}_{MM'}$ is a matrix available from textbooks [37]. The traditional A_k, λ_k conventions do not use orthogonal functions. Transformations from the traditional conventions to the ρ_M convention are given in an appendix.

Note the transformation properties listed are exact. The systematic and statistical errors of a measurement appear in fitting $\rho(X)$.

2.5 Fitting $\rho(X)$, $dN/d\Omega$

Quantum mechanics requires $\rho(X)$ must be positive, which means it has positive eigenvalues. Positivity produces subtle nonlinear constraints, similar to unitarity. In the 3×3 case the relations are generally cubic polynomials. Positivity is

---

[9] When $\text{tr}(\rho(X)) = 1$ is maintained, positivity is violated when one or more eigenvalues of $\rho(X)$ exceeds unity, and one or more goes negative. Then for some vector $|e\rangle$ the quadratic form $\langle e|\rho_X|e\rangle < 0$, which would appear to provide a signal. Yet no such signal might be found in the angular distribution, because $\text{tr}(\rho(\ell)\rho(X)) > 0$ is a much weaker condition. Thus, positivity cannot generally be reduced to bounds on angular distribution coefficients, unless the bounds are so intricately constructed to be equivalent to positivity of the density matrix eigenvalues.

[10] To help readers appreciate the practical value of these advantages, we constructed standalone analysis code in both ROOT and Mathematica [39]. We expect the code to provide useful cross-checks on code users might write for themselves.

Using the Standard Model parameters for $\rho(\ell)$, Eqs. (3) and (7), the trace yields

$$\frac{dN}{d\Omega_{\text{fit}}} \sim \text{tr}(\rho(\ell)\rho(X)) = 0.5000 + 0.0007739 \sin(\phi) \sin(\theta)$$
$$+ 0.3090 \cos^2(\phi) \cos^2(\theta) + 0.1904 \sin^2(\phi) \cos^2(\theta) + \cdots$$

where $\cdots$ indicates several terms there is no need to write out. Integrated over ϕ, this expression becomes

$$\frac{dN_{\text{fit}}}{d\cos\theta} = \frac{3}{4\pi}\left(1.57 + 0.137 \cos\theta + 1.56 \cos^2\theta\right).$$

A $1 + \cos^2\theta$ distribution is the leading-order Drell–Yan prediction for virtual spin-1 boson annihilation, while $0.137 \cos\theta$ represents a charge asymmetry.

It is trivial to go from $\text{tr}(\rho(\ell)\rho(X))$ to a conventional parameterization of an angular distribution by taking inner products of orthogonal functions. It is also easy to expand $\rho(X)$ in a basis of orthonormal matrices with the same results. Note these steps are exact, and very different from fitting data to trigonometric functions in some convention, which tends to yield multiple solutions, along with violations of positivity, which can introduce pathological convention dependence. Perhaps struggles with convention dependence of quarkonium data [40,41] are related to this. It would be interesting to investigate.

2.6 Summary of quantum tomography procedure

To analyze data for each event labeled J:

- Compute $Q_J = k_J + k'_J$; $\ell_J = k_J - k'_J$; $(X^\mu_J, Y^\mu_J, Z^\mu_J)$;
$$\vec{\ell}_{XYZ,J} = (X_J \cdot \ell_J, Y_J \cdot \ell_J, Z_J \cdot \ell_J);$$
$$\hat{\ell}_J = \vec{\ell}_{XYZ,J}/\sqrt{\vec{\ell}_{XYZ,J} \cdot \vec{\ell}_{XYZ,J}}.$$

- Make the lepton density matrix. For Z bosons in the Standard Model it is
$$\rho_{ij}(\ell) = \frac{1}{2}(\delta_{ij} - \hat{\ell}_i\hat{\ell}_j) - 0.22 \, \iota\epsilon_{ijk}\hat{\ell}_k. \tag{16}$$

- For Beyond Standard Model (BSM) phenomenology, one can readily assign values for a and b according to model specifics.[11] See footnote 11 for details. The results are

---

[11] $\rho(\ell)$ for Z bosons in the Standard Model can be constructed from the $Z \to \ell^+\ell^-$ matrix element: $i\mathcal{M} \sim \bar{v}(k)\gamma_\mu(1-\gamma_5)u(k)\epsilon_\mu$. Stripping off the polarization vector, taking the modulus squared of the resulting $\mathcal{M}_\mu$, and summing over spins gives $\rho^{\mu\nu} = \sum_{\text{spin}} |\mathcal{M}_\mu|^2 \sim \text{tr}(\slashed{k}\gamma_\mu\slashed{k}\gamma_\nu)$, which reduces to $\rho(\ell)$ after transforming to sum and difference variables q and ℓ and moving to the rest frame of the Z. A similar procedure can be used to obtain $\rho(\ell)$, as well as more general density matrices, for an arbitrary BSM process.

combined to make the Jth instance of $\text{tr}(\rho_J(\ell)\rho(X)(m))$, where $\rho(X)(m)$ has been parameterized in Eq. (14). Fit the m_α parameters to the data set. For example, the log likelihood $\mathcal{L}$ of the set $J = 1\cdots J_{\text{max}}$ is

$$\mathcal{L}(m) = \sum_{J}^{J_{\text{max}}} \log\left(\text{tr}\left(\rho^{(J)}(\ell) \cdot \rho(X)(m)\right)\right)$$
$$+ J_{\text{max}}\log(3/4\pi). \tag{17}$$

Sample code available online (see footnote 10) carries out these steps, returning the parameters m_α.

2.7 Comments

1. The possible symmetries of $\rho(\ell)$ enter here. Suppose $c_A = 0$. Then $\rho(\ell)$ is even under parity, real and symmetric. The imaginary antisymmetric elements of $\rho(X)$ are orthogonal, and contribute nothing to the angular distribution. When known in advance, the redundant parameters of $\rho(X)$ can be set to zero while making the fit. (That does not mean unmeasured parameters can be forgotten when dealing with positivity.) In general a fitting routine will either report a degeneracy for redundant parameters, or converge to values generated by round-off errors. Degeneracy will always be detected in the Hessian matrix computed to evaluate uncertainties.

2. The normalization condition $\sum_k m^2(k) = 1$ can be postponed by removing $1/\sqrt{m_\alpha^2}$ from Eq. (14), and subtracting $J_{\text{max}}\log(\sum_k^{J_{\text{max}}} m^2(k))$ from the log likelihood (Eq. (15)). When that is done the fitted density matrix will not be automatically normalized, due to the symmetry $\rho(X) \to \lambda\rho(X)$ of the modified likelihood. The density matrix becomes normalized by dividing by its trace. Incorporating such tricks improved the speed of the code available online (see footnote 10) by a factor of about 100.

3. Algorithms are said to compute a "unique" Cholesky decomposition, which would seem to predict m_α given $\rho(X)$. The algorithms choose certain signs of m_α by a convention making the diagonals of M positive. However, that is not quite enough to ensure a numerical fit finds a unique solution.
   The fundamental issue is that $MM^\dagger = \rho(X)$ is solved by $M = \sqrt{\rho(X)}$, and the square root is not unique. There are 2^N arbitrary sign choices possible among N eigenvalues of $\sqrt{\rho(X)}$. Forcing the diagonals of M to be positive reduces the possibilities greatly, and an algorithm exists to force a unique, canonical form of m_α in a data fitting routine. We did not make use of such a routine, since fitting $\rho(X)$ is the objective. Depending upon the data fitting method, increasing the number of ways for $M(m_\alpha)$ to make a fit sometimes makes convergence faster.

Table 1 Terms in the angular distribution with their properties under discrete transformations C_ℓ, P, and T. Here ℓ stands for $\hat{\ell}$, $X\ell$ stands for $\hat{X} \cdot \hat{\ell} = -X_\mu \ell^\mu$, and so on with scalar normalization factors removed. T-odd scattering observables from the imaginary parts of amplitudes generally exist without violating fundamental T symmetry. See the text for more explanation

Term	Origin	$dN/d\Omega$	C_ℓ	P	T	$C_\ell P$	PT
.	ℓ	.	−	−	−	+	+
.	X	.	.	−	−	+	+
.	Y	.	.	+	+	+	+
.	Z	.	.	−	−	+	+
S_x	$X\ell$	$\sin\theta\cos\phi$	−	+	+	−	+
S_y	$Y\ell$	$\sin\theta\sin\phi$	−	−	−	+	+
S_z	$Z\ell$	$\cos\theta$	−	+	+	−	+
ρ_2	$XX\ell\ell$	$\sin^2\theta\cos 2\phi$	+	+	+	+	+
ρ_3	$XY\ell\ell$	$\sin^2\theta\sin 2\phi$	+	−	−	+	+
ρ_1	$XZ\ell\ell$	$\sin 2\theta\cos\phi$	+	+	+	+	+
ρ_4	$YZ\ell\ell$	$\sin 2\theta\sin\phi$	+	−	−	+	+
ρ_0	$ZZ\ell\ell$	$1/\sqrt{3} - \sqrt{3}\cos^2\theta$	+	+	+	+	+

4. Let $\langle\rangle_{\mathrm{exp}}$ stand for the expectation value of a quantity in the experimental distribution of events. By symmetry $\left\langle\hat{\ell}\right\rangle_{\mathrm{exp}}$ and $\left\langle\hat{\ell}_i\hat{\ell}_j\right\rangle_{\mathrm{exp}}$ are vector and tensor estimators, respectively, which must depend on the vector and tensor parameters $\vec{S}$, $U_{ij}(X)$ in the underlying density matrix. A calculation finds

$$\left\langle\hat{\ell}\right\rangle_{\mathrm{exp}} = \frac{1}{J_{\mathrm{max}}}\sum_J \hat{\ell}_J = -\frac{1}{4}\vec{S};$$

$$\left\langle\hat{\ell}_i\hat{\ell}_j\right\rangle_{\mathrm{exp}} = \frac{1}{J_{\mathrm{max}}}\sum_J \hat{\ell}_{Ji}\hat{\ell}_{Ji} = \frac{1}{3}\delta_{ij} - \frac{1}{5}\mathrm{Re}[U_{ij}].$$

An estimate of $\rho(X)$ not needing a parameter search then exists directly from data. However, positivity of $\rho(X)$ is more demanding, and it is not automatically maintained by such estimates.

3 Results

3.1 Analysis bonuses of the quantum tomography procedure

3.1.1 Convex optimization

The issue of multiple solutions for $\rho(X)$ is different. Multiple minima of χ^2 statistics affects fits to cross sections parameterized by trigonometric functions. However, quantum tomography using maximum likelihood happens to be a problem of convex optimization. In brief, when ρ is positive then $\langle e|\rho|e\rangle$ is a positive convex function of $|e\rangle$. Then $\mathrm{tr}(\rho(\ell)\rho(X))$ is convex, being equivalent to a positively weighted sum of such terms. The logarithm is a concave func-

tion, leading to a convex optimization problem. That means that when $\rho(X)$ is a local maximum of likelihood it is the global maximum. Exceptions can only come from degeneracies due to symmetry or an inadequate number of data points [42]. Convex optimization is important because without such a property the evaluation of high-dimensional fits by trial and error can be exponentially difficult.

3.1.2 Discrete transformation properties

Table 1 lists discrete transformation properties of all terms under parity P, time reversal T, and lepton charge conjugation C_ℓ. If leptons have different flavors (as in like or unlike sign $e\mu$) the C_ℓ operation swaps the particle defining $\hat{\ell}$.

When coordinates XYZ are defined the direction of $\hat{Y} = \hat{Z}\times\hat{X}$ is even under time reversal and parity, which is exactly the opposite of X and Z. Then $\vec{S}\cdot\hat{Y}$ is T-odd, contributing the $\sin\theta\sin\phi$ term.[12] The XY and ZY matrix elements of $\rho(X)$ are also odd under T, contributing the terms shown. T-odd terms come from imaginary parts of amplitudes, which are generated by loop corrections in perturbative QCD. An excellent example of T-odd effects in QCD is given in Ref. [44] by Hagiwara et al.

Notice that every term in the lepton density matrix (Eq. (7)) is automatically symmetric under $C_\ell P$. This is a kinematic fact of the lepton-pair probe which does not originate in the Standard Model. As a result the $C_\ell P$ transformations of the angular distribution depend on the coupling to the unknown system. If overall CP symmetry exists the target density matrix will have CP odd terms where $C_\ell P$ odd terms are found. In the Standard Model these $\cos\theta$ and $\sin\theta\cos\phi$

---

[12] In a forthcoming study [43] of inclusive lepton-pair production near the Z pole, we find interesting, new features in the S_y data of Ref. [11].

terms correspond to charge asymmetries of leptons correlated with charge asymmetries of the system, namely the beam quark and antiquark distributions.

While weak CP violation is a mainstream topic, P and CP symmetry of the strong interactions at high energies has not been tested [45]. The gauge sector of QCD is kinematically CP symmetric, because the non-Abelian $\mathrm{tr}(\vec{E} \cdot \vec{B})$ term is a pure divergence.[13] Higher-order terms in a gauge-covariant derivative expansion are expected to exist, and can violate CP symmetry [45].

However, measuring violation of CP or fundamental T symmetry in scattering experiments is invariably frustrated by the experimental impossibility of preparing time-reversed counterparts. Some ingenuity is needed to devise a signal. It appears that any signal will involve four independent 4-momenta p_J and a quantity of the form $\Omega_4 = \epsilon_{\alpha\beta\lambda\sigma} p_1^\alpha p_2^\beta p_3^\lambda p_4^\sigma$. For example a term going like $\ell \cdot Y \sim \epsilon_{\alpha\beta\lambda\sigma} \ell_\alpha Q_\beta P_{A\lambda} P_{B\sigma}$ might possibly originate in fundamental T symmetry violation and might be mistaken for perturbative loop effects. A more creative road to finding CP violation involves two pairs with sum and difference vectors Q, ℓ; Q', ℓ', and the scalar $\epsilon_{\alpha\beta\lambda\sigma} \ell_\alpha Q_\beta Q'_\lambda \ell'_\sigma$, which is even under C and odd under P. The pairs need not be leptons (although "double Drell Yan" has long been discussed) but might be (say) $\mu^+\mu^-\pi^+\pi^-$. It would be interesting to explore further what a tomographic approach to such observables might uncover.

3.1.3 Density matrix invariants

We mentioned that scattering planes, trig functions, boosts and rotations could be avoided, and the examples show how. Once a frame convention is defined the lepton "coordinates" $(X_J \cdot \ell_J, Y_J \cdot \ell_J, Z_J \cdot \ell_J)$ are actually Lorentz scalars. However, they depend on the convention for XYZ, which is arbitrary. At least four different conventions compete for attention. Moreover, once a frame is chosen, at least two naming schemes (the "A_k" and "λ_k" schemes) exist to describe the angular distribution in terms of trigonometric polynomials..

Well-constructed invariants can reduce the confusion associated with convention-dependent quantities [43,46–48]. Since $\vec{S}$ transforms like a vector its magnitude-squared $\vec{S}^2$ is rotationally invariant. The spin-1 part of $\rho(X)$ does not mix with the real symmetric part under rotations. Since it is traceless, the real symmetric (spin-2) part has two independent eigenvalues, which are rotationally invariant.[14] Finally the dot-products of three eigenvectors $\hat{e}_J$ of the spin-2 part

with $\vec{S}$ are rotationally invariant. Then $(\hat{e}_j \cdot \vec{S})^2$ are three invariants not depending on the sign of eigenvectors. That suggests six possible invariants, but $\sum_j (\hat{e}_j \cdot \vec{S})^2 = \vec{S}^2$ makes the $\vec{S}$ invariants dependent, leaving five independent rotational invariants. That is consistent with counting eight real parameters in a 3×3 Hermitian matrix, subject to three free parameters of the rotation group, leaving $8 - 3 = 5$ rotational invariants. The same counting for unitary transformations would leave only the two independent eigenvalues of the matrix.

Any function of invariants is invariant. The combinations below have useful physical interpretations:

- The degree of polarization d is a standard measure of the deviation from the unpolarized case. It comes from the sum of the squares of the eigenvalues of ρ minus 1/3, normalized to the maximum possible:

$$d = \sqrt{(3\mathrm{tr}(\rho_X^2) - 1)/2},$$
$$\text{where} \quad 0 \le d \le 1.$$

When $d = 0$ the system is unpolarized, and when $d = 1$ the system is a pure state.

- The entanglement entropy $\mathcal{S}$ is the quantum mechanical measure of order. The formula is

$$\mathcal{S} = -\mathrm{tr}(\rho_X \log(\rho_X)).$$

In terms of eigenvalues ρ_α, $\mathcal{S} = -\sum_\alpha \rho_\alpha \log(\rho_\alpha)$. When $\rho \to 1_{N \times N}/N$ the system is unpolarized, and $\mathcal{S} = \log(N)$. That is the maximum possible entropy, and minimum possible information. When $\mathcal{S} = 0$ the entropy is the minimum possible, providing the maximum possible information, and the system is a pure state.

It is instructive to interpret $e^{\mathcal{S}}$ as the "effective dimension" of the system. For example the eigenvalues $(1/2 + b, 1/2 - b, 0)$ occur in the density matrix of on-shell fermion annihilation with helicity conservation. One zero-eigenvalue describes an elliptical disk-shaped object. The entropy ranges from $\mathcal{S} = 0$, ($e^{\mathcal{S}} = 1$ for $b = 1/2$, a one-dimensional stick shape) to $\mathcal{S} = \log(2)$, ($e^{\mathcal{S}} = 2$ for a disk-shaped object with maximum symmetry.) As expected, an unpolarized three-dimensional system has three equal eigenvalues, is shaped like a sphere, and $e^{\mathcal{S}} \to 3$.

---

[13] Non-perturbative strong CP violation in QCD by a surface term has been proposed. Tests have been dominated by the neutron dipole moment, while calculations of non-perturbative effects are problematic.

[14] Work by Faccioli et al. [49,50] attempted to construct invariants by inspecting the transformation properties of ratios of sums of angular distribution coefficients upon making rotation about the conventional

Footnote 14 continued

Y axis. The method cannot identify a true invariant unless Y happens to be an eigenvector of the matrix. By the same method the group also identified $\vec{S}^2$ as a "parity violating invariant", while $\vec{S}^2$ is actually even under parity. Parity violation is not required to measure $\vec{S}$ with polarized beams.

Figure 1 shows the entropy of the lepton density matrix $\rho(\ell)$ (Eq. (7)) in the plane of parameters (a, b). The matrix eigenvalues are[15] $(1/3 - 2a/3, 1/3 + a/3 - b, 1/3 + a/3 + b)$. The triangular boundaries are the positivity bounds on these parameters, outside of which the entropy has an imaginary part. The corners of the triangle are pure states. The left corner represents a purely longitudinal polarization, $\rho_L = |L\rangle\langle L|$ where $|L> = (0, 0, 1)$ in a coordinate system where $\hat{\ell} = \hat{z}$. The two right corners are purely circular polarizations, $\rho_{\pm} = |\epsilon_{\pm}\rangle\langle\epsilon_{\pm}|$, where in the same coordinates $\epsilon_{\pm} = (1, \pm\iota, 0)/\sqrt{2}$. The interior lines $a = \pm b$, $b = 0$ represent maximal symmetry matrices having two equal eigenvalues. The figure also indicates the constraints of a positive distribution for the example of Eq. (8) assuming lepton universality. The values of a and b are actually unrestricted in all directions, so long as they lie within the bounding curves.

The Standard Model leptons from lowest-order s-channel Z production have $a = 1/2$, $b = \sin^2\theta_W$, which is shown in Fig. 1 as a dot. The edge $a = 1/2$ corresponds to on-shell helicity conservation, with eigenvalues 0, $1/2 - \sin^2\theta_W$, $1/2 + \sin^2\theta_W$. The a, b parameters of leptons from a different production process, or subject to radiative corrections, must still lie inside the triangle. Maximal symmetry with eigenvalues $(1/2, 1/2, 0)$ occurs where the line of $b = \sin^2\theta_W$ just touches the $b = -a$ line, which happens at $\sin^2\theta_W = 1/4$. That is not far from the Standard Model value, which is very interesting. Since no established theory predicts $\sin^2\theta_W$ one cannot rule out a deeper connection.

It is tempting but incorrect to assume the bounds discussed would apply to the same terms of a more general density matrix. For example, add $-c\hat{n}_i\hat{n}_j$ to the expression in Eq. (7), where $\hat{n}\cdot\hat{\ell} = 0$ and update the normalization condition. The resulting positivity region of a, b, c is shown in Fig. 2, which also shows the plane $c = 0$ equivalent to Fig. 1. At the extrema $c = \pm 1$ the region of consistent (a, b) parameters shrinks to single points. The matrix for $\rho(X)$ computed earlier is an example where all terms in any standard convention happen to occur. By inspection this system (mostly quark–antiquark annihilation) is superficially much like the lepton one. The entropy of is 0.68 and $e^S = 1.96$, and one eigenvalue is close to zero. Of course there is much more information in the other parameters, the orientation of eigenvectors, and in $\vec{S}$ and its magnitude.
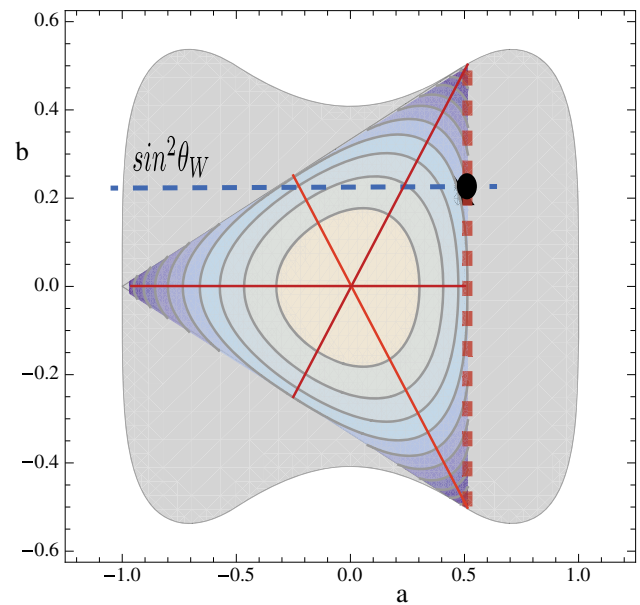

Fig. 1 Contours of constant entropy S of the lepton density matrix $\rho(\ell)$ (Eq. (7)) in the plane of parameters (a, b). Contours are separated by $1/10$ unit with $S = 0$ at the central intersection. The horizontal dashed line shows the lowest-order Standard Model prediction $b = \sin^2\theta_W$. Annihilation with on-shell helicity conservation is indicated by the vertical dashed line $a = 1/2$. The left corner of the triangle is a pure state with longitudinal polarization, while the two right corners are pure states of circular polarization. The interior lines represent matrices with maximal symmetry, where two eigenvalues are equal. They cross at the unpolarized limit. The curved gray region represents the much less restrictive constraints of a positive distribution using Eq. (8) and lepton universality
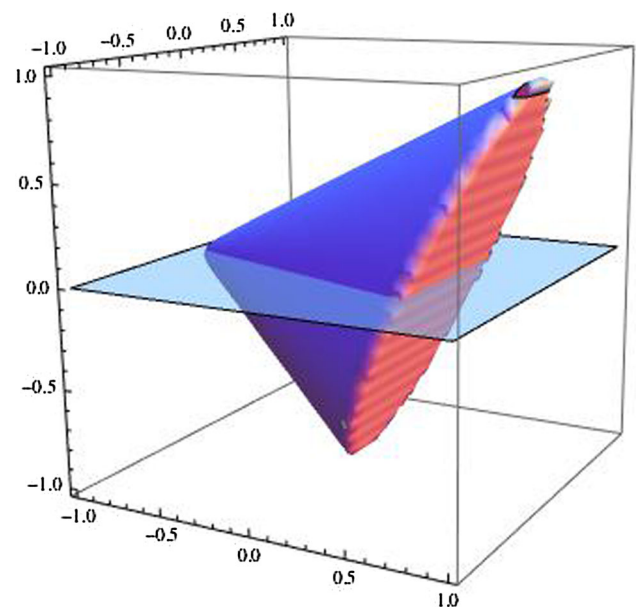

Fig. 2 Boundary of the positivity region of a density matrix depending on three parameters a, b, c described in the text. The two-dimensional region cut by the plane $c = 0$ corresponds to Fig. 1

---

[15] It can be shown that the eigenvalues $\lambda_k = 1/3 + (2d/3)\cos(\theta_k)$, where d is the degree of polarization and $\theta_k = \cos^{-1}(det((3\rho(X) - 1_{3\times3})/d)/2 + 2\pi k/3)$.

4 Discussion

The quantum tomography procedure offers at least seven significant advantages over standard methods of analyzing the angular correlations of inclusive reactions:

- *Simplicity and Efficiency* Tomography exploits a structured order of analysis. By construction, unobservable elements never appear.
- *Covariance* Physical quantities are expressed covariantly every step of the way. That is not always the case with quantities like angular distributions.
- *Complete polarization information* The unknown density matrix $\rho(X)$ contains all possible information, ready for classification under symmetry groups.
- *Model-independence* No theoretical planning, nor processing, nor assumptions are made about the unknown state. The process of defining general structure functions has been completely bypassed. It is not even necessary to assume anything about the spin of s- or t-channel intermediates. The observable target structures is always a mirror of the probe structure. The "mirror trick" is universal as described in Sect. 2.4.
- *Manifest positivity* A pattern of misconceptions in the literature misidentifies positivity as being equivalent to positive cross sections. It is not difficult to fit data to an angular distribution and violate positivity. In fact, an angular distribution expressed in terms of expansion coefficients actually lacks the quantum mechanical information to enforce positivity.
- *Convex optimization* The positive character of the density matrix leads to convex optimization procedures to fit experimental data. This provides a powerful analysis tool that ensures convergence.
- *Frame independence* Once the unknown density matrix has been reconstructed, rotationally invariant quantities can be made by straightforward methods. This is illustrated in Sect. 3.1.3, which includes a discussion of the entanglement entropy.

Quantum tomography has already yielded significant results. Our tomographic analysis [43] of a recent ATLAS study of Drell–Yan lepton pairs with invariant mass near the Z pole [11] discovered surprising features in the density matrix eigenvalues and entanglement entropy. By way of advertising, we have also gained insight into the mysterious Lam–Tung relation [51,52], including why it holds at NLO but fails at NNLO. These topics will be presented in separate papers.

Appendix A: Cross section in terms of density matrices

The density matrix approach stands on its own as an efficient tool kof quantum mechanics. However, we see the need to relate it to more traditional scattering formalism.

Consider the cross section for the inclusive production of two final-state particles of spin s, s' from particle intermediates:

$$d\sigma \sim \sum_{s,s'} \left| \sum_J \mathcal{M}(\chi_J \rightarrow f_{s,s'}) \right|^2 \cdot d\Pi_{\text{LIPS}},$$

where $\sum_J \mathcal{M}(\chi_J \rightarrow f) = \sum_J \langle f|T|\chi_J \rangle \delta^4(\sum p_f - \sum p_i)$, T is a transfer matrix, and $d\Pi_{LIPS}$ is the Lorentz-invariant phase space.

Then

$$d\sigma \sim \sum_{s,s'} \left[\left(\sum_J \langle f_{s,s'}|T|\chi_J \rangle \right) \right.$$
$$\left. \cdot \left(\sum_K \langle \chi_K|T^\dagger|f_{s,s'} \rangle \right) \right] \cdot d\Pi_{\text{LIPS}},$$
$$= \text{tr} \left[\left(\sum_{s,s'} T^\dagger|f_{s,s'}\rangle\langle f_{s,s'}|T \right) \right.$$
$$\left. \cdot \left(\sum_{J,K} |\chi_J\rangle\langle\chi_K| \right) \right] \cdot d\Pi_{\text{LIPS}}, \tag{18}$$

where $\sum_{J,K} |\chi_J\rangle\langle\chi_K|$ accounts for any interference between intermediate states. We identify the quantity in the first set of parentheses on the last line of Eq. (18) with the probe density matrix ρ_{probe} and the quantity in the second set of parentheses with the density matrix for the unknown particle intermediates, ρ_X. The $T, T^\dagger$ in ρ_{probe} ensure the overlap of ρ_X with ρ_{probe}, inside the trace, is taken at 'equal times'.

Rewriting $d\sigma$ in terms of the density matrices, we find

$$d\sigma \sim \text{tr}(\rho_{\text{probe}} \cdot \rho_X) \cdot d\Pi_{\text{LIPS}},$$

where $d\Pi_{\text{LIPS}} \sim d\Omega \cdot d^4q$ and q is the sum of the final-state pair momentum. Then, for given pair momentum q, the angular distribution is,

$$\frac{d\sigma}{d\Omega}(\theta, \phi | q) \sim \mathrm{tr}(\rho_{\mathrm{probe}} \cdot \rho_X),$$

where the proportionality suppresses an overall normalization. The conditional probability given q captures the event-by-event character of angular correlations. The explicit q dependence might suggest we assumed an s-channel boson intermediate state, but we have not.

If we know $d\sigma/d\Omega$ and ρ_{probe} for a given event, we can reconstruct ρ_X for that event. This is the essence of the quantum tomography procedure. The probe is what is known, and it determines what can be discovered.

Appendix B: Probe matrix

Figure 3 shows the diagram for the simplest lepton-pair probe. It is completely non-specific about the process creating a lepton with momentum k_1 and Dirac density matrix polarization $\alpha\alpha'$, and an antilepton with momentum k_2 and polarization $\beta\beta.'$ From the Feynman rules

$$\rho(lep)^{\beta\beta'}_{\alpha\alpha'} \sim (\not k_1)_{\alpha\alpha'}(\not k_2)_{\beta\beta'}. \tag{19}$$

The symbol $\sim$ indicates the high energy limit and ignoring a trivial overall normalization.

Continuing, $\rho(X)$ is something of vast complexity, which can only couple to $\rho(\mathrm{lep})$ via the indices shown. The Dirac structure of $\rho(X)$ can be expanded over several complete sets. However, the relevant (observable) part of $\rho(X)^{\beta\beta'}_{\alpha\alpha'}$ will be its projection onto the subspace coupled to this particular probe, just as the general analysis prescribed. It is ideal to classify only what will be observed. That sector is predetermined by the very limited Dirac structure of the probe. Thus

$$d\sigma \sim k_1^\mu k_2^\nu \rho_{\mu\nu}(X),$$
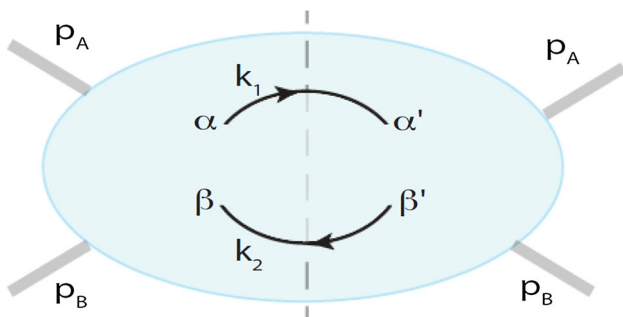where $\quad \rho_{\mu\nu}(X) = \mathrm{tr}\left((\gamma_\mu \otimes \gamma_\nu)\,\rho(X)\right).$

Fig. 3 The lepton-pair density matrix $\rho^{\alpha\alpha}_{\beta\beta'}(lep)$, in black, coupled to the colliding system density matrix $\rho(X)$. The matrix labels on legs are diagonal in momenta k_1, k_2. Off-diagonal polarization (Dirac) indices are explicitly shown. The Feynman rules are the same as for ordinary diagrams

With this probe massless fermion pairs produce no combinations of $k_1^\mu k_1^\nu$ or $k_2^\mu k_2^\nu$ or anything else. The operations show how the tr symbol comes to be used repeatedly with different meanings implied by the context.

We now use Hermiticity, which makes two predictions:

$$\rho(\mathrm{lep}) = \rho^{\mu\nu}(\mathrm{lep})\gamma_\mu\gamma_\mu; \quad \rho^{\mu\nu}(\mathrm{lep}) = \rho^{\nu\mu}(\mathrm{lep})^*,$$

plus the same relation for $\rho(X)$. All $k_1^\mu k_2^\nu$ factors must be strictly bilinear, and occur in a real symmetric plus $\iota \times$ anti-symmetric combinations. The most general possibility is

$$\rho^{\mu\nu}(\mathrm{lep}) = \alpha k_1 \cdot k_2 \eta^{\mu\nu} + \beta(k_1^\mu k_2^\nu + k_1^\nu k_2^\mu)$$
$$+ \iota\gamma(k_1^\mu k_2^\nu - k_1^\nu k_2^\mu) + \iota\delta\epsilon^{\mu\nu\alpha\beta}k_{1\alpha}k_{2\beta}, \tag{20}$$

where $\eta_{\mu\nu}$ is the Minkowski metric. By algebra

$$\rho^{\mu\nu}(\mathrm{lep}) = \frac{\alpha}{2}q^2\eta^{\mu\nu} + \frac{\beta}{2}q^\mu q^\nu - \frac{\beta}{2}\ell^\mu \ell^\nu$$
$$- \iota\frac{\gamma}{2}(q^\mu \ell^\nu - q^\nu \ell^\mu) - \iota\frac{\delta}{2}\epsilon^{\mu\nu\alpha\beta}q_\alpha \ell_\beta. \tag{21}$$

Event-by-event there exists a preferred, oriented pair rest frame where $q^\mu = (q, \vec{0})$ and $\ell^\mu = (0, q\hat{\ell})$. In that frame Eq. (21) predicts a normalized 3×3 tensor of spatial components which is

$$\rho^{jk}(\mathrm{lep}) = \frac{1}{3}\delta^{jk} + a\,J_p^{jk}\ell^p - b\,U^{jk}; \tag{22}$$

$$U^{jk}(\hat{\ell}) = \hat{\ell}^j \hat{\ell}^k - \frac{\delta^{jk}}{3}. \tag{23}$$

Here a and b are real, and J_p are the spin-1 rotation generators in Cartesian coordinates:

$$J_p^{jk} = -\iota\epsilon_{pjk}.$$

Equation 22 has been decomposed into tensors transforming under rotations like spin-0, spin-1 and spin-2. The expansion above is kinematic and not a consequence of any special theory.

The approach has the virtue of maintaining strict control of how outputs depend on assumptions. We have made few assumptions, yet we have a result, which is that $\rho^{jk}(\mathrm{lep})$ only depends on two scalars a, b. The only scalar available from $\rho(\mathrm{lep}; \ell, q)$ is q^2, hence $a = a(q^2)$, $b = b(q^2)$. The enormous body of field theory and the Standard Model only predicts only two parameters of Eq. (22). If and when the lepton pair originates from an intermediate boson with vertex $c_V\gamma^\mu + c_A\gamma_\mu\gamma_5$, then

$$a = c_A c_V; \quad b = 1/2.$$

The general possibility these parameters might be functions of q^2, namely vertex form factors, has emerged on its own. Notice that in the rest frame oriented naturally along the lepton momentum ρ (lep) is not diagonal. The diagonal elements are interpretable as probabilities, even classical probabilities. The off-diagonal elements convey the information as regards entanglement.

Appendix C: Positivity

There is a positivity issue in fitting angular distribution data. Represent $\rho = MM^\dagger$, and then $\rho > 0$. Any $M = HU$, where $H = H^\dagger$ and $UU^\dagger = 1$. We can make M self-adjoint since U cancels out. To parameterize $N \times N$ matrices M use $SU(N)$ representations G_a, normalized to $G_aG_b = (1/2)\delta_{ab}$. For $N = 3$ those are the Gell-Mann matrices. We define parameters with

$$M = m_0 1_{3\times 3}/\sqrt{3} + \sqrt{2}m_aG_a. \tag{24}$$

Compute

$$MM^\dagger = m_0^2/3 + 2\sqrt{\frac{2}{3}}m_0m_aG_a + 2m_am_bG_aG_b.$$

The symmetric product is

$$G_aG_b + G_bG_a = \delta_{ab}/3 + f_{abg}G_g.$$

Check the trace of both sides for the normalization of δ_{ab}. Everything else must be traceless and must be spanned by G_g. Then

$$MM^\dagger = \frac{1}{3}\left(m_0^2 + \sum_a m_a^2\right) + 2\sqrt{\frac{2}{3}}m_0m_gG_g$$
$$+ f_{abg}m_am_bG_g.$$

The normalization $tr(\rho) = 1$ needs $\sum_{\mu=0}^{9} m_\mu m_\mu = 1$. This requires each $0 < m_\mu^2 < 1$, while it is more restrictive.

There is a degeneracy issue in the nonlinear relation of m_μ to a straight expansion $\rho = 1/3 + c_gG_g$,

$$c_g = 2\sqrt{\frac{2}{3}}\sqrt{1 - m \cdot m}\, m_g + f_{abg}m_am_b. \tag{25}$$

Notice

$$[\rho, M] = [M^2, M] = 0.$$

Then M and ρ have the same eigenvectors. The eigenvalues of ρ are those of M, squared. For N eigenvalues of ρ there are 2^N possible M's. If ρ is positive definite, however, there is only one M with strictly positive diagonal entries. In that sense, the M satisfying $\rho = MM^\dagger$ can be said to be unique.

The positivity problem is often solved with the Cholesky decomposition: $\rho = LL^\dagger$, where L is a lower-triangular matrix with real entries on the diagonal. L is related to M by a similarity transform. There are $N + 2N(N-1)/2 = N^2$ free real parameters in a lower-triangular matrix with real diagonals, which is just right for Hermiticity. As before, the condition $\text{tr}(\rho) = 1$ requires $\sum m_\mu m_\mu = 1$. The Cholesky decomposition is unique, in the sense above, when ρ is positive definite.

Appendix D: Collected conventions

As a consequence of consistent definitions, our ρ_M and Y_M transform under rotations like real representations of spin-2. Other conventions have long existed. Table 2 shows the relations of the ρ_M parameters compared to the ad hoc conventions known as A_k and λ_k. The ρ_M are self-explanatory because they correspond to orthonormal harmonics and transform like spin-2 representations. The arbitrary normalizations and conventions relating different basis functions have been a barrier to interpretation, needlessly complicated transformations between angular frame conventions.

Our self-explanatory conventions for the spin-1 parameters are given in Table 3. For example, it is quite easy to

Table 2 Three ways of parameterizing the monopole (left of double vertical line) relative to quadrupole (right) part of the angular distribution. The bottom row (right) represents our spin-2 basis functions that are both orthogonal and uniformly normalized, compared to ad hoc conventions of the other rows. The spin-2 coefficient combinations listed in each row are the ones that mix linearly under rotations of the frame coordinates. To find the parameterization of a given row, multiply the

coefficient in each column by the function at the bottom, and add. To absolutely normalize the A_j form, multiply the entire sum by $3/(16\pi)$, and the normalized series will begin at $1/(4\pi)$. To absolutely normalize the λ_j form multiply the entire sum by $3/(4\pi)$. The ρ_M form is absolutely normalized by definition. The constant $c = 3/(8\pi\sqrt{2})$ has been divided out to match the other conventions. The absolutely normalized form uses the sum of the ρ_M row multiplied by c

$16\pi/3$	$(\sqrt{3}A_0/2 - 1/\sqrt{3})$	A_1	$A_2/2$	A_5	A_6
$4\pi/(3+\lambda_\theta)$	$-\lambda_\theta/(3\sqrt{3}+\sqrt{3}\lambda_\theta)$	$\lambda_{\theta\phi}/(3+\lambda_\theta)$	$\lambda_\phi/(3+\lambda_\theta)$	$\lambda_\phi^\perp/(3+\lambda_\theta)$	$\lambda_{\theta\phi}^\perp/(3+\lambda_\theta)$
$1/c$	ρ_0	ρ_1	ρ_2	ρ_3	ρ_4
$1/(4\pi)$	$1/\sqrt{3} - \sqrt{3}\cos^2(\theta)$	$\sin(2\theta)\cos(\phi)$	$\sin^2(\theta)\cos(2\phi)$	$\sin^2(\theta)\sin(2\phi)$	$\sin(2\theta)\sin(\phi)$

Table 3 Parameterizing the spin-1 part of the angular distribution. To form the angular distribution coefficients from each row are multiplied by functions on the bottom row and added to those from Table 2

A_3	A_7	A_4
$2A_\phi$	$2A_\phi^\perp$	$2A_\theta$
S_x	S_y	S_z
$\sin(\theta)\cos(\phi)$	$\sin(\theta)\sin(\phi)$	$\cos(\theta)$

remember that $S_x \to \sin\theta\cos\phi$, $S_y \to \sin\theta\sin\phi$ and S_z leads to $\cos\theta$ angular dependence.

References

1. U. Fano, Phys. Rev. **93**, 121 (1954)
2. H.F. Hofmann, S. Takeuchi, Phys. Rev. A **69**, 042108 (2004). arXiv:quant-ph/0310003
3. G.L. Long, H.Y. Yan, Y. Sun, J. Opt. B Quantum Semiclass. Opt. **3**, 376 (2001). arXiv:quant-ph/0012047
4. R.C. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A.D. O'Connell, D. Sank, H. Wang, J. Wenner, M. Steffen, A.N. Cleland, J.M. Martinis, Nat. Phys. **6**, 409 (2010). arXiv:0910.1118 [quant-ph]
5. S.S. Straupe, Sov. J. Exp. Theor. Phys. Lett. **104**, 510 (2016). arXiv:1610.02840 [quant-ph]
6. D. Gross, Y.-K. Liu, S.T. Flammia, S. Becker, J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010). arXiv:0909.3304 [quant-ph]
7. A. Bisio, G. Chiribella, G. Mauro D'Ariano, S. Facchini, P. Perinotti, (2017). arXiv:1702.08751 [quant-ph]
8. I.A. Fedorov, A.K. Fedorov, Y.V. Kurochkin, A.I. Lvovsky, New J. Phys. **17**, 043063 (2015). arXiv:1403.0432 [quant-ph]
9. A. Anis, A. I. Lvovsky, (2012). arXiv:1204.5936 [quant-ph]
10. B. Abelev et al. (ALICE), Phys. Rev. Lett. **108**, 082001 (2012). arXiv:1111.1630 [hep-ex]
11. G. Aad et al. (ATLAS), JHEP **08**, 159 (2016). arXiv:1606.00689 [hep-ex]
12. A. Melnitchouk (ATLAS, CDF, LHCb, CMS, D0), in *Proceedings, 31st International Conference on Physics in collisions (PIC 2011)*, Vancouver, Canada, August 28–September 1, 2011 (2011). arXiv:1111.2840 [hep-ex]
13. N. Brambilla et al., Eur. Phys. J. C **71**, 1534 (2011). arXiv:1010.5827 [hep-ph]
14. T. Aaltonen et al. (CDF), Phys. Rev. Lett. **106**, 241801 (2011). arXiv:1103.5699 [hep-ex]
15. S. Chatrchyan et al. (CMS), Phys. Lett. B **710**, 91 (2012). arXiv:1202.1489 [hep-ex]
16. S. Chatrchyan et al. (CMS), Phys. Lett. B **727**, 101 (2013). arXiv:1303.5900 [hep-ex]
17. S. Chatrchyan et al. (CMS), Phys. Lett. B **727**, 381 (2013). arXiv:1307.6070 [hep-ex]
18. S. Chatrchyan et al. (CMS), Phys. Rev. Lett. **110**, 081802 (2013). arXiv:1209.2922 [hep-ex]
19. R. Aaij et al. (LHCb), Eur. Phys. J. C **73**, 2631 (2013). arXiv:1307.6379 [hep-ex]
20. R. Aaij et al. (LHCb), Eur. Phys. J. C **74**, 2872 (2014). arXiv:1403.1339 [hep-ex]
21. R.E. McClellan, Angular Distributions of High-Mass Dilepton Production in Hadron Collisions. Ph.D. thesis, Illinois U., Urbana (2016)
22. J.-C. Peng, J.-W. Qiu, Prog. Part. Nucl. Phys. **76**, 43 (2014). arXiv:1401.0934 [hep-ph]
23. J.-C. Peng, W.-C. Chang, R.E. McClellan, O. Teryaev, Phys. Lett. B **758**, 384 (2016). arXiv:1511.08932 [hep-ph]
24. W.J. Stirling, E. Vryonidou, JHEP **07**, 124 (2012). arXiv:1204.6427 [hep-ph]
25. J. Han (CDF), in *Particles and fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2011*, Providence, USA, August 9–13, 2011 (2011). arXiv:1110.0153 [hep-ex]
26. CMS Collaboration, Search for the standard model Higgs boson decaying to a W pair in the fully leptonic final state in pp collisions at sqrt(s) = 8 TeV (2012). Report Number CMS-PAS-HIG-12-017
27. V. Khachatryan et al. (CMS), Phys. Lett. B **749**, 14 (2015). arXiv:1501.07750 [hep-ex]
28. V. Cheung, R. Vogt, Phys. Rev. D **95**, 074021 (2017). arXiv:1702.07809 [hep-ph]
29. V. Cheung, R. Vogt, (2017). arXiv:1706.07686 [hep-ph]
30. Z.-B. Kang, Y.-Q. Ma, J.-W. Qiu, G. Sterman, Phys. Rev. D **91**, 014030 (2015). arXiv:1411.2456 [hep-ph]
31. A. Cervera-Lierta, J.I. Latorre, J. Rojo, L. Rottoli, (2017). arXiv:1703.02989 [hep-th]
32. H. Terazawa, Phys. Rev. D **8**, 3026 (1973)
33. K.V. Vasavada, Phys. Rev. D **16**, 146 (1977). [Erratum: Phys. Rev. D **17**, 2515 (1978)]
34. E. Mirkes, Nucl. Phys. B **387**, 3 (1992)
35. E. Mirkes, J. Ohnemus, Phys. Rev. D **50**, 5692 (1994). arXiv:hep-ph/9406381
36. E. Mirkes, J. Ohnemus, Phys. Rev. D **51**, 4891 (1995). arXiv:hep-ph/9412289
37. T. Sochi, (2016). arXiv:1603.01660 [math.HO]
38. N.J. Higham, M. Eprint, N.J. Higham, *Reliable Numerical Computation* (Oxford University Press, New York, 1990), pp. 161–185
39. J.C. Martens, J.P. Ralston, J.D. Tapia Takaki, Quantum tomography sample code. http://cern.ch/quantum-tomography
40. P. Faccioli, C. Lourenco, J. Seixas, Phys. Rev. D **81**, 111502 (2010). arXiv:1005.2855 [hep-ph]
41. P. Faccioli, C. Lourenco, J. Seixas, H.K. Wohri, Eur. Phys. J. C **69**, 657 (2010). arXiv:1006.2738 [hep-ph]
42. S. Burer, R.D.C. Monteiro, Math. Program. **103**, 427 (2005)
43. J.P. Ralston, in *Proceedings, Probing QCD in Photon-Nucleus Interactions at RHIC and LHC: the Path to EIC (INT-17-65W)*, Seattle, Washington, February 12–17, 2017 (2017)
44. K. Hagiwara, T. Kuruma, Y. Yamada, Nucl. Phys. B **369**, 171 (1992)
45. M. Backovic, J.P. Ralston, Phys. Rev. D **84**, 016010 (2011). arXiv:1101.1965 [hep-ph]
46. S. Palestini, Phys. Rev. D **83**, 031503 (2011). arXiv:1012.2485 [hep-ph]
47. H.-S. Shao, K.-T. Chao, Phys. Rev. D **90**, 014002 (2014). arXiv:1209.4610 [hep-ph]
48. Y.-Q. Ma, J.-W. Qiu, H. Zhang, (2017). arXiv:1703.04752 [hep-ph]
49. P. Faccioli, C. Lourenco, J. Seixas, H.K. Wohri, Phys. Rev. D **82**, 096002 (2010). arXiv:1010.1552 [hep-ph]
50. P. Faccioli, C. Lourenco, J. Seixas, Phys. Rev. Lett. **105**, 061601 (2010). arXiv:1005.2601 [hep-ph]
51. C.S. Lam, W.-K. Tung, Phys. Rev. D **18**, 2447 (1978); D **21**, 2712 (1980)
52. C.S. Lam, W.-K. Tung, Phys. Rev. D **21**, 2712 (1980)