Whither All the Scope and Generality of Bell's Theorem?

Joy Christian*

Wolfson College, University of Oxford, Oxford OX2 6UD, United Kingdom and
Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, England, UK

In a recent preprint James Owen Weatherall has attempted a simple local-deterministic model for the EPR-Bohm correlation and speculated about why his model fails when my counterexample to Bell's theorem succeeds. Here I bring out the physical, mathematical, and conceptual reasons why his model fails. In particular, I demonstrate why no model based on a tensor representation of the rotation group SU(2) can reproduce the EPR-Bohm correlation. I demonstrate this by calculating the correlation explicitly between measurement results $\mathscr{A} = \pm 1$ and $\mathscr{B} = \pm 1$ in a local and deterministic model respecting the spinor representation of SU(2). I conclude by showing how Weatherall's reading of my model is misguided, and bring out a number of misconceptions and unwarranted assumptions in his imitation of my model as it relates to the Bell-CHSH inequalities.

I. INTRODUCTION

In a recent preprint [1] James Owen Weatherall has attempted a local and deterministic model for the EPR-Bohm experiment, which he affirms to have been inspired by my work on Bell's theorem [2][3][4][5][6]. Although his analysis mainly concerns his attempted reconstruction of my model and why the attempt fails, his preprint has been worded in a manner that has allowed some readers to embrace his discussion as a criticism of my work on Bell's theorem [2][3][6]. Here I demonstrate that (as he himself stresses to some extent) the analysis Weatherall presents in his preprint has nothing to do with my model, or with the physics and mathematics of the EPR-Bohm correlation [6][7]. In fact his analysis exhibits no understanding of how my local-realistic framework works, nor of the reasons why it explains the origins of all quantum correlations [3][8]. I show this by first calculating the EPR-Bohm correlation in a successful local-deterministic model based on the spinor representation of SU(2), and then revealing a number of misconceptions and unwarranted assumptions in Weatherall's reconstruction of my model as it relates to the Bell-CHSH inequalities. I conclude that, contrary to first impressions, Weatherall's thinly veiled criticism of my work is entirely vacuous.

II. AN EXACT LOCAL-DETERMINISTIC MODEL FOR THE EPR-BOHM CORRELATION

In order to bring out the erroneous assumptions in Weatherall's analysis, it would be convenient to assess it in the light of a successful model for the EPR-Bohm correlation. This will allow us to unveil his assumptions more easily.

A. A Complete Specification of the Singlet State

To this end, let us recall Bell's definition of a *complete* theory [9]. He considered a physical theory to be complete just in case its predictions for the EPR-Bohm experiment are dictated by local-deterministic functions of the form

$$\mathscr{A}(\mathbf{r},\,\lambda) : \mathrm{I\!R}^3 \times \Lambda \longrightarrow S^0 \equiv \{-1,\,+1\}, \tag{1}$$

where $\mathrm{I\!R}^3$ is the space of 3-vectors $\mathbf{r}$, Λ is a space of the complete states λ, and $S^0 \equiv \{-1,\,+1\}$ is a unit 0-sphere. He then proved a mathematical theorem concluding that no pair of functions of this form can reproduce the correlation as strong as that predicted by quantum mechanics for the rotationally invariant singlet state:

$$\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathscr{A}(\mathbf{a},\,\lambda^k)\,\mathscr{B}(\mathbf{b},\,\lambda^k)\right] \,\neq\, -\,\mathbf{a}\cdot\mathbf{b}\,. \tag{2}$$

*Electronic address: joy.christian@wolfson.ox.ac.uk

As it stands, this conclusion of Bell is entirely correct and beyond dispute, provided we accept prescription (1) as codifying a complete specification of the singlet state. In his preprint, following Bell and CHSH [9][10], Weatherall accepts (1) as codifying a complete specification of the singlet state, whereas my work begins by recognizing that (1) *does not*, and *cannot*, codify a complete specification of the singlet state [3][8]. I have argued that Bell's prescription is based on an incorrect underpinning of both the EPR argument [11] and the actual topological configurations involved in the EPR-Bohm experiments [3], even if we leave the physics and mathematics underlying the correlation aside. My argument is rather subtle and requires a clear understanding of what is meant by both a *function* in mathematics and the geometry and topology of a parallelized 3-sphere. But the bottom line of the argument is that, for any two-level system, the EPR criterion of completeness demands the correct measurement functions to be necessarily of the form

$$\pm\, 1 \;=\; \mathscr{A}(\mathbf{r},\,\lambda) : \mathrm{I\!R}^3\times\Lambda \longrightarrow S^3 \sim \mathrm{SU}(2), \tag{3}$$

with the *simply-connected* codomain S^3 of $\mathscr{A}(\mathbf{r},\,\lambda)$ replacing the *totally disconnected* codomain S^0 assumed by Bell. Thus $\mathscr{A}(\mathbf{r},\,\lambda)=\pm\,1$ now represents a point of a parallelized 3-sphere, S^3. As a function, it takes values from the domain $\mathrm{I\!R}^3\times\Lambda$ and ends up *belonging* to the codomain S^3. Consequently, any correlation between a pair of such results is a correlation between points of a parallelized 3-sphere. Unless based on a prescription of this precise form, any Bell-type analysis simply does not get off the ground, because without completeness there can be no theorem [3].

Here S^3—which can be thought of as the configuration space of all possible rotations of a rotating body (including spinorial sign changes)—is defined as the set of all unit quaternions isomorphic to a unit parallelized 3-sphere:

$$S^3 := \left\{\,\mathbf{q}(\psi,\,\mathbf{r}) := \exp\left[\boldsymbol{\beta}(\mathbf{r})\,\frac{\psi}{2}\right]\,\Bigg|\;||\,\mathbf{q}(\psi,\,\mathbf{r})\,||^2 = 1\right\}, \tag{4}$$

where $\boldsymbol{\beta}(\mathbf{r})$ is a bivector rotating about $\mathbf{r}\in\mathrm{I\!R}^3$ with the rotation angle ψ in the range $0\leq\psi<4\pi$. Throughout this paper I will follow the concepts, notations, and terminology of geometric algebra [12][13]. Accordingly, $\boldsymbol{\beta}(\mathbf{r})\in S^2\subset S^3$ can be parameterized by a unit vector $\mathbf{r}=r_1\,\mathbf{e}_1+r_2\,\mathbf{e}_2+r_3\,\mathbf{e}_3\in\mathrm{I\!R}^3$ as

$$\begin{aligned}\boldsymbol{\beta}(\mathbf{r}) \;&:=\; (\,I\cdot\mathbf{r}\,)\\ &=\; r_1\,(\,I\cdot\mathbf{e}_1\,)\,+\,r_2\,(\,I\cdot\mathbf{e}_2\,)\,+\,r_3\,(\,I\cdot\mathbf{e}_3\,)\\ &=\; r_1\;\mathbf{e}_2\,\wedge\,\mathbf{e}_3\,+\,r_2\;\mathbf{e}_3\,\wedge\,\mathbf{e}_1\,+\,r_3\;\mathbf{e}_1\,\wedge\,\mathbf{e}_2\,,\end{aligned} \tag{5}$$

with $\boldsymbol{\beta}^2(\mathbf{r})=-1$. Here the trivector $I:=\mathbf{e}_1\wedge\mathbf{e}_2\wedge\mathbf{e}_3$ (which also squares to -1) represents a volume form of the physical space [12][13]. Each configuration of the rotating body can thus be represented by a quaternion of the form

$$\mathbf{q}(\psi,\,\mathbf{r}) \;=\; \cos\frac{\psi}{2}\,+\,\boldsymbol{\beta}(\mathbf{r})\,\sin\frac{\psi}{2}\,, \tag{6}$$

which in turn can always be decomposed as a product of two bivectors belonging to an $S^2\subset S^3\sim\mathrm{SU}(2)$,

$$\boldsymbol{\beta}(\mathbf{r}'')\,\boldsymbol{\beta}(\mathbf{r}') \;=\; \cos\frac{\psi}{2}\,+\,\boldsymbol{\beta}(\mathbf{r})\,\sin\frac{\psi}{2}\,, \tag{7}$$

with ψ being its rotation angle from $\mathbf{q}(0,\,\mathbf{r})=1$. Note also that $\mathbf{q}(\psi,\,\mathbf{r})$ reduces to $\pm\,1$ as $\psi\to\,2\kappa\pi$ for $\kappa\,=\,0,\,1,\,\text{or}\,\,2.$

It is of paramount importance to note here that our topologically corrected prescription (3) does not alter the actual measurement results. For a given vector $\mathbf{r}$ and an initial state λ, both operationally and mathematically we still have

$$\mathscr{A}(\mathbf{r},\,\lambda) \;=\; +\,1\;\;\text{or}\;\;-\,1 \tag{8}$$

as the *image points* of the function $\mathscr{A}(\mathbf{r},\,\lambda)$ as demanded by Bell, but now the topology of its *codomain* has changed from a 0-sphere to a 3-sphere, with the latter embedded in $\mathrm{I\!R}^4$ in such a manner that the prescriptions (1) and (3) are *operationally identical* [3][8]. On the other hand, without this topological correction it is impossible to provide a complete account of all possible measurement results in the sense specified by EPR. Thus the selection of the codomain $S^3\hookrightarrow\mathrm{I\!R}^4$ in prescription (3) is not a matter of choice but necessity. What is responsible for the EPR correlation is *the topology of the set of all possible measurement results.* But once the codomain of $\mathscr{A}(\mathbf{r},\,\lambda)$ is so corrected, the proof of Bell's theorem (as given in Refs. [9]) simply falls apart. Moreover, it turns out that the strength of the correlation for *any* physical system is entirely determined by the torsion within the codomain of the local functions $\mathscr{A}(\mathbf{r},\,\lambda)$.

Returning to Weatherall's analysis, it should be clear now that, because it is based on a prescription other than (3), it too is a *non-starter* [3]. More importantly, once we recognize that the *only* way of providing a complete account of all possible measurement results for the singlet state is by means of prescription (3), the statistical procedure for analyzing the correlation must be consistently customized for the set S^3 of all unit quaternions, *which Weatherall fails to do*. Since this procedure can be appreciated more readily by studying the explicit construction of my model from my book and elsewhere, I now proceed to reproduce the model in some detail in the following subsections. Doing so will also dispel a persistent but gravely disingenuous charge that my model encounters "certain technical complications" [4][5].

B. Construction of the Measurement Functions

Once the measurement results are represented by functions of the form (3), it is easy to reproduce the EPR-Bohm correlation in a manifestly local, realistic, and deterministic manner. This is because a parallelized 3-sphere has quite a unique and distinctive topological structure [6][7]. It is one of the only two parallelizable spheres with non-vanishing torsion—the other one, with variable torsion, being the 7-sphere. Once parallelized by a constant torsion, the 3-sphere remains closed under multiplication, and forms one of the only four possible normed division algebras [7][8]. These are profound concepts underlying the very existence and strength of quantum correlations [3]. By ignoring them and dismissing them as irrelevant, Weatherall is ignoring the physics and mathematics of the quantum correlations. More importantly, because of the unique and distinctive topological characteristics of the 3-sphere [7], the measurement functions such as $\mathscr{A}(\mathbf{r},\,\lambda)$ for the EPR-Bohm correlation have to be constructed in a very specific manner for any model to be successful [3]. To this end, let us begin with the following definition of the orientation of a vector space:

Definition II.1 *An orientation of a finite dimensional vector space $\mathcal{V}_d$ is an equivalence class of ordered basis, say $\{f_1,\,\dots,\,f_d\}$, which determines the same orientation of $\mathcal{V}_d$ as the basis $\{f'_1,\,\dots,\,f'_d\}$ if $f'_i=\omega_{ij}f_j$ holds with $det(\omega_{ij})>0$, and the opposite orientation of $\mathcal{V}_d$ as the basis $\{f'_1,\,\dots,\,f'_d\}$ if $f'_i=\omega_{ij}f_j$ holds with $det(\omega_{ij})<0$.*

(Here repeated indices are summed over.) Thus each positive dimensional real vector space has precisely two possible orientations, which (rather suggestively) can be denoted as $\lambda=+1$ or $\lambda=-1$. More generally an oriented smooth manifold such as S^3 consists of that manifold together with a choice of orientation for each of its tangent spaces.

It is important to note that orientation of a manifold is a *relative* concept [14]. In particular, the orientation of a tangent space $\mathcal{V}_d$ of a manifold defined by the equivalence class of ordered basis such as $\{f_1,\,\dots,\,f_d\}$ is meaningful only with respect to that defined by the equivalence class of ordered basis $\{f'_1,\,\dots,\,f'_d\}$, and vice versa. To be sure, we can certainly orient a manifold absolutely by choosing a set of ordered bases for all of its tangent spaces, but the resulting manifold can be said to be left or right oriented only with respect of another such set of ordered basis [14].

Now the natural configuration space for an EPR-Bohm type experiment is a unit parallelized 3-sphere, which can be embedded in $\mathrm{I\!R}^4$ with a choice of orientation, say $\lambda=+1$ or -1. This choice of orientation can be identified with the initial state of the particle pair in the singlet state with respect to the orientation of the detector basis as follows. We first characterize the embedding space $\mathrm{I\!R}^4$ by the graded basis

$$\{\,1,\;L_1(\lambda),\;L_2(\lambda),\;L_3(\lambda)\,\}\,, \tag{9}$$

with $\lambda=\pm\,1$ representing the two possible orientations of S^3 and the basis elements $L_\mu(\lambda)$ satisfying the algebra

$$L_\mu(\lambda)\,L_\nu(\lambda)\;=\;-\,g_{\mu\nu}\;-\;\epsilon_{\mu\nu\rho}\,L_\rho(\lambda)\,, \tag{10}$$

with an arbitrary metric $g_{\mu\nu}$ on S^3. Here the bivectors $\{\,a_\mu\,L_\mu(\lambda)\,\}$ will represent the spin angular momenta of the particles, with $\mu=1,\,2,\,3$ and the repeated indices summed over. These momenta can be assumed to be detected by the detector bivectors, say $\{\,a_\mu\,D_\mu\,\}$, with the corresponding detector basis $\{\,1,\,D_1,\,D_2,\,D_3\,\}$ satisfying the algebra

$$D_\mu\,D_\nu\;=\;-\,g_{\mu\nu}\;-\;\epsilon_{\mu\nu\rho}\,D_\rho \tag{11}$$

and related to the spin basis $\{\,1,\,L_1(\lambda),\,L_2(\lambda),\,L_3(\lambda)\,\}$ as

$$\begin{pmatrix}1\\ L_1(\lambda)\\ L_2(\lambda)\\ L_3(\lambda)\end{pmatrix}\;=\;\begin{pmatrix}1&0&0&0\\ 0&\lambda&0&0\\ 0&0&\lambda&0\\ 0&0&0&\lambda\end{pmatrix}\begin{pmatrix}1\\ D_1\\ D_2\\ D_3\end{pmatrix}. \tag{12}$$

Evidently, the determinant of this matrix works out to be $det(\omega_{ij})=\lambda$. Since $\lambda^2=+1$ and ω^2 is a 4×4 identity matrix, this relation can be more succinctly written as

$$L_\mu(\lambda)\;=\;\lambda\,D_\mu\quad\text{and}\quad D_\mu\;=\;\lambda\,L_\mu(\lambda)\,, \tag{13}$$

or equivalently as

$$\{\,1,\; L_1(\lambda),\; L_2(\lambda),\; L_3(\lambda)\,\} \;=\; \{1,\, \lambda\, D_1,\, \lambda\, D_2,\, \lambda\, D_3\} \tag{14}$$

and

$$\{\,1,\; D_1,\; D_2,\; D_3\,\} \;=\; \{1,\, \lambda\, L_1(\lambda),\, \lambda\, L_2(\lambda),\, \lambda\, L_3(\lambda)\}\,. \tag{15}$$

These relations reiterate the fact that orientation of any manifold is a *relative* concept. In particular, orientation of S^3 defined by the spin basis $\{\,1,\, L_\mu(\lambda)\,\}$ is meaningful only with respect to that defined by the detector basis $\{\,1,\, D_\mu\,\}$ with the orientation $\lambda = +1$, and vice versa. Thus the spin basis are said to define the *same* orientation of S^3 as the detector basis if $L_\mu(\lambda = +1) = +D_\mu$, and the spin basis are said to define the *opposite* orientation of S^3 as the detector basis if $L_\mu(\lambda = -1) = -\,D_\mu$. Note also that the numbers 1 and $\mathbf{L}(\mathbf{r},\, \lambda)$ are treated here on equal footing.

We are now in a position to define the functions $\mathscr{A}(\mathbf{a},\, \lambda)$ and $\mathscr{B}(\mathbf{b},\, \lambda)$ as results of measurement interactions (or Clifford products) between detector bivectors $-\,\mathbf{D}(\mathbf{a})$ and $+\,\mathbf{D}(\mathbf{b})$ and spin bivectors $\mathbf{L}(\mathbf{a},\, \lambda)$ and $\mathbf{L}(\mathbf{b},\, \lambda)$ as follows:

$$\mathrm{SU}(2) \sim S^3 \ni \pm\,1 \;=\; \mathscr{A}(\mathbf{a},\, \lambda^k) \;=\; -\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\, \lambda^k) \;=\; \{-\,a_\mu\, D_\mu\,\}\,\{\,a_\nu\, L_\nu(\lambda^k)\,\} \;=\; \begin{cases} +\,1 & \text{if} \quad \lambda^k \;=\; +\,1 \\ -\,1 & \text{if} \quad \lambda^k \;=\; -\,1 \end{cases} \tag{16}$$

and

$$\mathrm{SU}(2) \sim S^3 \ni \pm\,1 \;=\; \mathscr{B}(\mathbf{b},\, \lambda^k) \;=\; +\,\mathbf{D}(\mathbf{b})\,\mathbf{L}(\mathbf{b},\, \lambda^k) \;=\; \{+\,b_\mu\, D_\mu\,\}\,\{\,b_\nu\, L_\nu(\lambda^k)\,\} \;=\; \begin{cases} -\,1 & \text{if} \quad \lambda^k \;=\; +\,1 \\ +\,1 & \text{if} \quad \lambda^k \;=\; -\,1\,, \end{cases} \tag{17}$$

where the relative orientation λ is now assumed to be a random variable, with 50/50 chance of being $+1$ or $-\,1$ at the moment of creation of the singlet pair of spinning particles. In what follows, I will assume that the orientation of S^3 defined by the detector basis $\{\,1,\, D_\nu\,\}$ has been fixed before hand [3][6]. Thus the spin bivector $\{\,a_\mu\, L_\mu(\lambda)\,\}$ is a random bivector with its handedness determined *relative* to the detector bivector $\{\,a_\nu\, D_\nu\,\}$, by the relation

$$\mathbf{L}(\mathbf{a},\, \lambda) \;\equiv\; \{\,a_\mu\, L_\mu(\lambda)\,\} \;=\; \lambda\,\{\,a_\nu\, D_\nu\,\} \;\equiv\; \lambda\,\mathbf{D}(\mathbf{a}), \tag{18}$$

where, as a direct consequence of the algebra (10) with $g_{\mu\nu} = \delta_{\mu\nu}$, the bivectors $\mathbf{L}(\mathbf{a},\, \lambda)$ satisfy the following identity:

$$\mathbf{L}(\mathbf{a},\, \lambda)\,\mathbf{L}(\mathbf{a}',\, \lambda) \;=\; -\,\mathbf{a}\cdot\mathbf{a}' \;-\; \mathbf{L}(\mathbf{a}\times\mathbf{a}',\, \lambda). \tag{19}$$

Using these relations the spin detection events (16) and (17) follow at once from the algebras defined in (10) and (11).

Evidently, the measurement results $\mathscr{A}(\mathbf{a},\, \lambda)$ and $\mathscr{B}(\mathbf{b},\, \lambda)$ as defined above, in addition to being manifestly realistic, are strictly local and deterministically determined numbers. In fact, they are not even contextual. Alice's measurement result $\mathscr{A}(\mathbf{a},\, \lambda)$—although it refers to a freely chosen direction $\mathbf{a}$—depends only on the initial state λ; and likewise, Bob's measurement result $\mathscr{B}(\mathbf{b},\, \lambda)$—although it refers to a freely chosen direction $\mathbf{b}$—depends only on the initial state λ. Let us also not overlook the fact that, as binary numbers, $\mathscr{A}(\mathbf{a},\, \lambda) = \pm\,1$ and $\mathscr{B}(\mathbf{b},\, \lambda) = \pm\,1$ are still points of a parallelized 3-sphere. To confirm this, recall that a parallelized 3-sphere is a set of unit quaternions of the form

$$\mathbf{q}^k(\psi,\, \mathbf{r},\, \lambda) \;:=\; \left\{\lambda^k\, \cos\frac{\psi}{2} \,+\, \mathbf{L}\big(\mathbf{r},\, \lambda^k\big)\, \sin\frac{\psi}{2}\right\}, \tag{20}$$

and a measurement result such as $\mathscr{A}(\mathbf{a},\, \lambda) = \pm 1$ is a limiting case of such a quaternion constituting the 3-sphere:

$$\begin{aligned} S^3 \ni \pm\,1 \;=\; \mathscr{A}(\mathbf{a},\, \lambda) \;&=\; \lim_{\mathbf{a}'\to\,\mathbf{a}} \; \mathscr{A}(\mathbf{a},\, \mathbf{a}',\, \lambda) \\ &=\; \lim_{\mathbf{a}'\to\,\mathbf{a}} \; \{\,-\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a}',\, \lambda)\,\} \\ &=\; \lim_{\mathbf{a}'\to\,\mathbf{a}} \; \{\,(\,-\,I\cdot\mathbf{a})(\,\lambda\, I\cdot\mathbf{a}')\,\} \\ &=\; \lim_{\mathbf{a}'\to\,\mathbf{a}} \; \{\,\lambda\,\mathbf{a}\cdot\mathbf{a}' \,+\, \lambda\, I\cdot(\mathbf{a}\times\mathbf{a}')\,\} \\ &=\; \lim_{\psi\to\,2\kappa\pi} \; \left\{\lambda\, \cos\frac{\psi}{2} \,+\, \mathbf{L}(\mathbf{c},\, \lambda)\, \sin\frac{\psi}{2}\right\} \\ &=\; \lim_{\psi\to\,2\kappa\pi} \; \{\,\mathbf{q}(\psi,\, \mathbf{c},\, \lambda)\,\}\,. \end{aligned} \tag{21}$$

Here $I = \mathbf{e}_1 \wedge \mathbf{e}_2 \wedge \mathbf{e}_3$ is the volume form, limit $\mathbf{a}' \to \mathbf{a}$ is equivalent to the limit $\psi \to 2\kappa\pi$ for $\kappa = 0,\, 1,$ or 2, $\psi = 2\,\eta_{\mathbf{a}\,\mathbf{a}'}$ is the rotation angle about the axis $\mathbf{c} := \mathbf{a}\times\mathbf{a}'/|\mathbf{a}\times\mathbf{a}'|,$ and $\eta_{\mathbf{a}\,\mathbf{a}'}$ is the angle between $\mathbf{a}$ and $\mathbf{a}'$ [6].

C. A Crucial Lesson from Basic Statistics

It is important to note that the variables $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$ defined in equations (16) and (17) are generated with *different* bivectorial scales of dispersion (or different standard deviations) for each measurement direction **a** and **b**. Consequently, in statistical terms these variables are raw scores, as opposed to standard scores [15]. Recall that a standard score, z, indicates how many standard deviations an observation or datum is above or below the mean. If x is a raw (or unnormalized) score and $\overline{\mathrm{x}}$ is its mean value, then the standard (or normalized) score, z(x), is defined by

$$\mathrm{z}(\mathrm{x}) \,=\, \frac{\mathrm{x}\,-\,\overline{\mathrm{x}}}{\sigma(\mathrm{x})}\,, \tag{22}$$

where $\sigma(\mathrm{x})$ is the standard deviation of x. A standard score thus represents the distance between a raw score and population mean in the units of standard deviation, and allows us to make comparisons of raw scores that come from very different sources [3][15]. In other words, the mean value of the standard score itself is always zero, with standard deviation unity. In terms of these concepts the correlation between raw scores x and y is defined as

$$\mathcal{E}(\mathrm{x},\,\mathrm{y}) \,=\, \frac{\lim\limits_{n\,\gg\,1}\left[\dfrac{1}{n}\sum\limits_{k\,=\,1}^{n}\,(\mathrm{x}^k\,-\,\overline{\mathrm{x}}\,)\;(\mathrm{y}^k\,-\,\overline{\mathrm{y}}\,)\right]}{\sigma(\mathrm{x})\;\sigma(\mathrm{y})} \tag{23}$$

$$=\,\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathrm{z}(\mathrm{x}^k)\;\mathrm{z}(\mathrm{y}^k)\right]. \tag{24}$$

It is vital to appreciate that covariance by itself—*i.e.*, the numerator of equation (23) by itself—does not provide the correct measure of association between the raw scores, not the least because it depends on different units and scales (or different scales of dispersion) that may have been used in the measurements of such scores. Therefore, to arrive at the correct measure of association between the raw scores one must either use equation (23), with the product of standard deviations in the denominator, or use covariance of the standardized variables, as in equation (24).

Now, as discussed above, the random variables $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$ are products of two factors—one random and another non-random. Within the variable $\mathscr{A}(\mathbf{a},\,\lambda)$ the bivector $\mathbf{L}(\mathbf{a},\,\lambda)$ is a random factor—a function of the orientation λ, whereas the bivector $-\,\mathbf{D}(\mathbf{a})$ is a non-random factor, independent of the orientation λ:

$$\mathscr{A}(\mathbf{a},\,\lambda) \,=\, -\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\,\lambda) \tag{25}$$

$$\text{and}\quad \mathscr{B}(\mathbf{b},\,\lambda) \,=\, +\,\mathbf{D}(\mathbf{b})\,\mathbf{L}(\mathbf{b},\,\lambda) \tag{26}$$

Thus, as random variables, $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$ are generated with *different* standard deviations—*i.e.*, *different* sizes of the typical error. More specifically, $\mathscr{A}(\mathbf{a},\,\lambda)$ is generated with the standard deviation $-\,\mathbf{D}(\mathbf{a})$, whereas $\mathscr{B}(\mathbf{b},\,\lambda)$ is generated with the standard deviation $+\,\mathbf{D}(\mathbf{b})$. These two deviations can be calculated as follows. Since errors in the linear relations propagate linearly, the standard deviation $\sigma(\mathscr{A}\,)$ of $\mathscr{A}(\mathbf{a},\,\lambda)$ is equal to $-\,\mathbf{D}(\mathbf{a})$ times the standard deviation of $\mathbf{L}(\mathbf{a},\,\lambda)$ [which I will denote as $\sigma(A) = \sigma(\mathbf{L_a})$], whereas the standard deviation $\sigma(\mathscr{B}\,)$ of $\mathscr{B}(\mathbf{b},\,\lambda)$ is equal to $+\,\mathbf{D}(\mathbf{b})$ times the standard deviation of $\mathbf{L}(\mathbf{b},\,\lambda)$ [which I will denote as $\sigma(B) = \sigma(\mathbf{L_b})$]:

$$\sigma(\mathscr{A}\,) \,=\, -\,\mathbf{D}(\mathbf{a})\,\sigma(A) \tag{27}$$

$$\text{and}\quad \sigma(\mathscr{B}\,) \,=\, +\,\mathbf{D}(\mathbf{b})\,\sigma(B). \tag{28}$$

But since the bivector $\mathbf{L}(\mathbf{a},\,\lambda)$ is normalized to unity, and since its mean value $m(\mathbf{L_a})$ vanishes on the account of λ being a fair coin, its standard deviation is easy to calculate, and it turns out to be equal to unity:

$$\sigma(A) \,=\, \sqrt{\frac{1}{n}\sum_{k\,=\,1}^{n}\,\Big|\Big|\,A(\mathbf{a},\,\lambda^k)\,-\,\overline{A(\mathbf{a},\,\lambda^k)}\,\Big|\Big|^2} \,=\, \sqrt{\frac{1}{n}\sum_{k\,=\,1}^{n}\,||\,\mathbf{L}(\mathbf{a},\,\lambda^k)\,-\,0\,||^2} \,=\, 1, \tag{29}$$

where the last equality follows from the normalization of $\mathbf{L}(\mathbf{a},\,\lambda)$. Similarly, it is easy to see that $\sigma(B)$ is also equal to 1. Consequently, the standard deviation of $\mathscr{A}(\mathbf{a},\,\lambda) = \pm\,1$ works out to be $-\,\mathbf{D}(\mathbf{a})$, and the standard deviation of $\mathscr{B}(\mathbf{b},\,\lambda) = \pm\,1$ works out to be $+\,\mathbf{D}(\mathbf{b})$. Putting these two results together, we arrive at the following standard scores corresponding to the raw scores $\mathscr{A}$ and $\mathscr{B}$:

$$A(\mathbf{a},\,\lambda) = \frac{\mathscr{A}(\mathbf{a},\,\lambda)\,-\,\overline{\mathscr{A}(\mathbf{a},\,\lambda)}}{\sigma(\mathscr{A})} \,=\, \frac{-\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\,\lambda)\,-\,0}{-\,\mathbf{D}(\mathbf{a})} \,=\, \mathbf{L}(\mathbf{a},\,\lambda) \tag{30}$$

and

$$B(\mathbf{b},\,\lambda) = \frac{\mathscr{B}(\mathbf{b},\,\lambda)\,-\,\overline{\mathscr{B}(\mathbf{b},\,\lambda)}}{\sigma(\mathscr{B})} \;=\; \frac{+\,\mathbf{D}(\mathbf{b})\,\mathbf{L}(\mathbf{b},\,\lambda)\,-\,0}{+\,\mathbf{D}(\mathbf{b})} \;=\; \mathbf{L}(\mathbf{b},\,\lambda), \tag{31}$$

where I have used identities such as $-\,\mathbf{D}(\mathbf{a})\mathbf{D}(\mathbf{a}) = +1$. Needless to say, these standard scores are pure bivectors:

$$\mathrm{SU}(2) \sim S^3 \supset S^2 \ni \mathbf{L}(\mathbf{a},\,\lambda) \,=\, \pm\,1 \;\text{ about } \mathbf{a} \in \mathrm{I\!R}^3, \tag{32}$$

$$\text{and}\quad \mathrm{SU}(2) \sim S^3 \supset S^2 \ni \mathbf{L}(\mathbf{b},\,\lambda) \,=\, \pm\,1 \;\text{ about } \mathbf{b} \in \mathrm{I\!R}^3. \tag{33}$$

D. How Errors Propagate within a Parallelized 3-sphere

As noted towards the end of subsection II A, one of several oversights in Weatherall's reading of my model concerns his failure to recognize the necessity of applying the correct statistical procedure for analyzing the correlation between the measurement results defined by (16) and (17). This is surprising, because I pointed out this oversight to him in a private correspondence more than eighteen months ago. As we noted above, a parallelized 3-sphere is a set of unit quaternions. Each point of a parallelized 3-sphere is thus represented by a unit quaternion. As a result, the correct statistical procedure within my model must take into account how errors propagate in a 3-sphere of unit quaternions.

Accordingly, let a probability density function $P(\mathbf{q}) : S^3 \rightarrow [0,\,1]$ of random quaternions over S^3 be defined as:

$$P(\mathbf{q}) \;=\; \frac{1}{\sqrt{2\pi\,||\,\sigma(\mathbf{q})||^2}}\;\exp\left\{-\,\frac{||\,\mathbf{q} - m(\mathbf{q})||^2}{2\;||\,\sigma(\mathbf{q})||^2}\right\}, \tag{34}$$

where the square root of $\mathbf{q} = \mathbf{p}\,\mathbf{p}$, $\mathbf{p} \in S^3$, is defined as

$$\sqrt{\mathbf{q}} \;=\; \sqrt{\mathbf{p}\,\mathbf{p}} \;:=\; \pm\,\mathbf{p}^\dagger(\,\mathbf{p}\,\mathbf{p}\,) \;=\; \pm(\,\mathbf{p}^\dagger\mathbf{p}\,)\,\mathbf{p} \;=\; \pm\,\mathbf{p}\,. \tag{35}$$

It is a matter of indifference whether the distribution of $\mathbf{q} \in S^3$ so chosen happens to be normal or not. Here $\mathbf{q}$ is an arbitrary quaternion within $S^3(\lambda)$ of the form (20), which is a sum of a scalar and a bivector (treated on equal footing), with $0 \leq \psi \leq 4\pi$ being the double-covering rotation angle about $\mathbf{r}$-axis. The mean value of $\mathbf{q}$ is defined as

$$m(\mathbf{q}) \;=\; \frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathbf{q}^k\,, \tag{36}$$

and the standard deviation of $\mathbf{q}$ is defined as

$$\sigma[\mathbf{q}(\psi,\,\mathbf{r},\,\lambda)] \;:=\; \sqrt{\frac{1}{n}\sum_{k\,=\,1}^{n}\;\{\,\mathbf{q}^k(\psi)\;-\;m(\mathbf{q})\}\;\{\,\mathbf{q}^k(2\pi-\psi)\;-\;m(\mathbf{q})\}^\dagger}\;. \tag{37}$$

Note that in this definition $\mathbf{q}(\psi)$ is coordinated by ψ to rotate from 0 to 2π, whereas the conjugate $\mathbf{q}^\dagger(2\pi-\psi)$ is coordinated by ψ to rotate from 2π to 0. Thus, for a given value of λ, both $\mathbf{q}(\psi)$ and $\mathbf{q}^\dagger(2\pi-\psi)$ represent the same sense of rotation about $\mathbf{r}$ (either both represent clockwise rotations or both represent counterclockwise rotations). This is crucial for the calculation of standard deviation, for it is supposed to give the average rotational distance within S^3 from its mean, with the average being taken, not over rotational distances within a fixed orientation of S^3, but over the changes in the orientation λ of S^3 itself. Note also that, according to the definition (20), $\mathbf{q}(\psi)$ and its conjugate $\mathbf{q}^\dagger(\psi)$ satisfy the following relation:

$$\mathbf{q}^\dagger(2\pi-\psi) \;=\; -\,\mathbf{q}(\psi)\,. \tag{38}$$

Consequently, the standard deviation of both $\mathbf{q}^\dagger(2\pi-\psi)$ and $-\,\mathbf{q}(\psi)$ must, with certainty, give the same number:

$$\sigma[\,\mathbf{q}^\dagger(2\pi-\psi)] \;\equiv\; \sigma[\,-\,\mathbf{q}(\psi)]\,. \tag{39}$$

It is easy to verify that definition (37) for the standard deviation of $\mathbf{q}(\psi)$ does indeed satisfy this requirement, at least when $m(\mathbf{q}) = 0$. What is more, from equation (38) we note that the quantity being averaged in the definition (37) is essentially $-\,\mathbf{q}\,\mathbf{q}$. This quantity is insensitive to spinorial sign changes such as $\mathbf{q} \to -\,\mathbf{q}$, but transforms into the quantity $-\,\mathbf{q}^\dagger\mathbf{q}^\dagger$ under orientation changes such as $\lambda \to -\lambda$. By contrast, the quantity $-\,\mathbf{q}\,\mathbf{q}^\dagger$ would be insensitive to both spinorial sign changes as well as orientation changes. Thus $\sigma[\mathbf{q}(\lambda)]$, as defined in (37), is designed to remain sensitive to orientation changes for correctly computing its averaging function on $\mathbf{q}(\lambda)$ in the present context.

Now, in order to evaluate $\sigma(\mathscr{A})$ and $\sigma(\mathbf{L_a})$, we can rewrite the quaternion (20) rotating about $\mathbf{r} = \mathbf{a}$ as a product

$$\mathbf{q}(\psi,\,\mathbf{a},\,\lambda) \,=\, \mathbf{p}(\psi,\,\mathbf{a})\,\mathbf{L}(\mathbf{a},\,\lambda) \tag{40}$$

of a non-random, non-pure quaternion

$$\mathbf{p}(\psi,\,\mathbf{a}) \,:=\, \sin\left(\frac{\psi}{2}\right) \,-\, \mathbf{D}(\mathbf{a})\,\cos\left(\frac{\psi}{2}\right) \,=\, \exp\left\{-\,\mathbf{D}(\mathbf{a})\left(\frac{\pi-\psi}{2}\right)\right\} \tag{41}$$

and a random, unit bivector $\mathbf{L}(\mathbf{a},\,\lambda)$ satisfying

$$\frac{1}{n}\sum_{k\,=\,1}^{n}\mathbf{L}(\mathbf{a},\,\lambda^k)\,\mathbf{L}^\dagger(\mathbf{a},\,\lambda^k) \,=\, 1\,. \tag{42}$$

Note that $\mathbf{p}(\psi,\,\mathbf{a})$ reduces to the unit bivector $\pm\,\mathbf{D}(\mathbf{a})$ for rotation angles $\psi = 0$, $\psi = 2\pi$, and $\psi = 4\pi$. Moreover, using the relations $\mathbf{L}(\mathbf{a},\,\lambda) = \lambda\,\mathbf{D}(\mathbf{a})$ and $\mathbf{D}^2(\mathbf{a}) = -\,1$ it can be easily checked that the product in (40) is indeed equivalent to the quaternion defined in (20) for $\mathbf{r} = \mathbf{a}$. It is also easy to check that the non-random quaternion $\mathbf{p}(\psi,\,\mathbf{a})$ satisfies the following relation with its conjugate:

$$\mathbf{p}^\dagger(2\pi-\psi,\,\mathbf{a}) \,=\, \mathbf{p}(\psi,\,\mathbf{a})\,. \tag{43}$$

Consequently we have

$$\mathbf{q}^\dagger(2\pi-\psi,\,\mathbf{a},\,\lambda) \,=\, \{\mathbf{p}(2\pi-\psi,\,\mathbf{a})\,\mathbf{L}(\mathbf{a},\,\lambda)\}^\dagger \,=\, \mathbf{L}^\dagger(\mathbf{a},\,\lambda)\,\mathbf{p}^\dagger(2\pi-\psi,\,\mathbf{a}) \,=\, \mathbf{L}^\dagger(\mathbf{a},\,\lambda)\,\mathbf{p}(\psi,\,\mathbf{a})\,. \tag{44}$$

Thus, substituting for $\mathbf{q}(\psi,\,\mathbf{a},\,\lambda)$ and $\mathbf{q}^\dagger(2\pi-\psi,\,\mathbf{a},\,\lambda)$ from Eqs. (40) and (44) into Eq. (37), together with

$$m(\mathbf{q_a}) \,=\, \frac{1}{n}\sum_{k\,=\,1}^{n}\mathbf{q}^k_{\mathbf{a}} \,=\, \mathbf{p}(\psi,\,\mathbf{a})\,\left\{\frac{1}{n}\sum_{k\,=\,1}^{n}\mathbf{L}(\mathbf{a},\,\lambda^k)\right\} \,=\, \mathbf{p}(\psi,\,\mathbf{a})\,\left\{\frac{1}{n}\sum_{k\,=\,1}^{n}\lambda^k\right\}\mathbf{D}(\mathbf{a}) \,=\, 0\,, \tag{45}$$

we have

$$\begin{aligned}
\sigma[\mathbf{q}(\psi,\,\mathbf{a},\,\lambda)] \,&=\, \sqrt{\frac{1}{n}\sum_{k\,=\,1}^{n}\{\,\mathbf{p}(\psi,\,\mathbf{a})\,\mathbf{L}(\mathbf{a},\,\lambda^k)\,\}\;\{\,\mathbf{L}^\dagger(\mathbf{a},\,\lambda^k)\,\mathbf{p}(\psi,\,\mathbf{a})\,\}} \\
&=\, \sqrt{\mathbf{p}(\psi,\,\mathbf{a})\,\left\{\frac{1}{n}\sum_{k\,=\,1}^{n}\mathbf{L}(\mathbf{a},\,\lambda^k)\,\mathbf{L}^\dagger(\mathbf{a},\,\lambda^k)\right\}\,\mathbf{p}(\psi,\,\mathbf{a})} \\
&=\, \sqrt{\mathbf{p}(\psi,\,\mathbf{a})\,\mathbf{p}(\psi,\,\mathbf{a})} \\
&=\, \pm\,\mathbf{p}(\psi,\,\mathbf{a})\,.
\end{aligned} \tag{46}$$

Here I have used the normalization of $\mathbf{L}(\mathbf{a},\,\lambda)$ as in (42), and the last equality follows from the definition (35). It can also be deduced from the polar form of the product

$$\mathbf{p}(\psi,\,\mathbf{a})\,\mathbf{p}(\psi,\,\mathbf{a}) \,=\, \cos(\pi-\psi) \,-\, \mathbf{D}(\mathbf{a})\,\sin(\pi-\psi) = \exp\{\,-\,\mathbf{D}(\mathbf{a})\,(\pi-\psi)\}\,. \tag{47}$$

The result for the standard deviation we have arrived at, namely

$$\sigma[\,\mathbf{q}(\psi,\,\mathbf{a},\,\lambda)] \,=\, \pm\,\mathbf{p}(\psi,\,\mathbf{a})\,, \tag{48}$$

is valid for all possible rotation angles ψ between the detector bivector $-\,\mathbf{D}(\mathbf{a})$ and the spin bivector $\mathbf{L}(\mathbf{a},\,\lambda)$. For the special cases when $\psi = 0,\ \pi,\ 2\pi,\ 3\pi$, and 4π, it reduces to the following set of standard deviations:

$$\begin{aligned}
\sigma[\,\mathbf{q}(\psi=0,\,\mathbf{a},\,\lambda)] \,&=\, \sigma(\mathscr{A}) \,=\, \pm\,\mathbf{D}(\mathbf{a}) \\
\sigma[\,\mathbf{q}(\psi=\pi,\,\mathbf{a},\,\lambda)] \,&=\, \sigma(\mathbf{L_a}) \,=\, \pm\,1 \\
\sigma[\,\mathbf{q}(\psi=2\pi,\,\mathbf{a},\,\lambda)] \,&=\, \sigma(\mathscr{A}) \,=\, \pm\,\mathbf{D}(\mathbf{a}) \\
\sigma[\,\mathbf{q}(\psi=3\pi,\,\mathbf{a},\,\lambda)] \,&=\, \sigma(\mathbf{L_a}) \,=\, \pm\,1 \\
\text{and}\quad \sigma[\,\mathbf{q}(\psi=4\pi,\,\mathbf{a},\,\lambda)] \,&=\, \sigma(\mathscr{A}) \,=\, \pm\,\mathbf{D}(\mathbf{a})\,.
\end{aligned} \tag{49}$$

To understand the physical significance of these results, let us first consider the special case when $\psi = \pi$. Then

$$\mathbf{q}(\psi = \pi,\, \mathbf{a},\, \lambda) \,=\, +\,\mathbf{L}(\mathbf{a},\, \lambda)\,, \tag{50}$$

which can be seen as such from the definition (20) above. Similarly, for the conjugate of $\mathbf{q}(\psi = \pi,\, \mathbf{a},\, \lambda)$ we have

$$\mathbf{q}^\dagger(\psi = \pi,\, \mathbf{a},\, \lambda) \,=\, -\,\mathbf{L}(\mathbf{a},\, \lambda) \,=\, +\,\mathbf{L}^\dagger(\mathbf{a},\, \lambda)\,. \tag{51}$$

Moreover, we have $m(\mathbf{L_a}) = 0$, since $\mathbf{L}(\mathbf{a},\, \lambda) = +\,\lambda\,\mathbf{D}(\mathbf{a})$ with $\lambda = \pm\,1$ being a fair coin. Substituting these results into definition (37)—together with $\psi = \pi$—we arrive at

$$\sigma(\mathbf{L_a}) \,=\, \sqrt{\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathbf{L}(\mathbf{a},\, \lambda^k)\,\mathbf{L}^\dagger(\mathbf{a},\, \lambda^k)\,} \,=\, \pm\,1\,, \tag{52}$$

since $\mathbf{L}(\mathbf{a},\, \lambda)\,\mathbf{L}^\dagger(\mathbf{a},\, \lambda) = 1$. Similarly, we can consider the case when $\psi = 3\pi$ and again arrive at $\sigma(\mathbf{L_a}) = \pm\,1$.

Next, we consider the three remaining special cases, namely $\psi = 0,\; 2\pi,$ or 4π. These cases correspond to the measurement results, as defined, for example, in equation (16). To confirm this, recall from equation (21) that a measurement result such as $\mathscr{A}(\mathbf{a},\, \lambda) = \pm 1$ is a limiting case of the quaternion (20). If we now rotate the bivector $\mathbf{L}(\mathbf{c},\, \lambda)$ to $\mathbf{L}(\mathbf{a},\, \lambda)$ as

$$\mathbf{D}(\mathbf{r})\,\mathbf{L}(\mathbf{c},\, \lambda)\,\mathbf{D}^\dagger(\mathbf{r}) \,=\, \mathbf{L}(\mathbf{a},\, \lambda) \tag{53}$$

using some $\mathbf{D}(\mathbf{r})$, and multiply Eq. (21) from the left by $\mathbf{D}(\mathbf{r})$ and from the right by $\mathbf{D}^\dagger(\mathbf{r})$, then we arrive at

$$\begin{aligned}\mathscr{A}(\mathbf{a},\, \lambda) \,&=\, \lim_{\mathbf{a'}\to\,\mathbf{a}}\,\left\{\mathbf{D}(\mathbf{r})\,\mathbf{q}(\psi,\, \mathbf{c},\, \lambda)\,\mathbf{D}^\dagger(\mathbf{r})\right\} \\ &=\, \lim_{\psi\to\,2\kappa\pi}\,\left\{\lambda\,\cos\frac{\psi}{2} \,+\, \mathbf{L}(\mathbf{a},\, \lambda)\,\sin\frac{\psi}{2}\right\} \\ &=\, \lim_{\psi\to\,2\kappa\pi}\,\left\{\,\mathbf{q}(\psi,\, \mathbf{a},\, \lambda)\,\right\} \\ &=\, \lim_{\psi\to\,2\kappa\pi}\,\left\{\,\mathbf{p}(\psi,\, \mathbf{a})\,\mathbf{L}(\mathbf{a},\, \lambda)\,\right\},\end{aligned} \tag{54}$$

where $\mathbf{p}(\psi,\, \mathbf{a})$ is defined in equation (41). The limit $\mathbf{a'} \to \mathbf{a}$ is thus physically equivalent to the limit $\psi \to 2\kappa\pi$ for $\kappa \,=\, 0,\; 1,$ or 2. We therefore have the following relation between $\mathscr{A}(\mathbf{a},\, \lambda)$ and $\mathbf{q}(\psi,\, \mathbf{a},\, \lambda)$:

$$\mathbf{q}(\psi = 2\kappa\pi,\, \mathbf{a},\, \lambda) \,=\, \pm\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\, \lambda) \,=\, \pm\,\mathscr{A}(\mathbf{a},\, \lambda)\,, \tag{55}$$

and similarly between $\mathscr{A}^\dagger(\mathbf{a},\, \lambda)$ and $\mathbf{q}^\dagger(\psi,\, \mathbf{a},\, \lambda)$:

$$\mathbf{q}^\dagger(\psi = 2\kappa\pi,\, \mathbf{a},\, \lambda) \,=\, \left\{\,\pm\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\, \lambda)\,\right\}^\dagger \,=\, \pm\,\mathscr{A}^\dagger(\mathbf{a},\, \lambda)\,. \tag{56}$$

For example, for $\psi = 0$ the definition (20) leads to

$$\mathbf{q}(\psi = 0,\, \mathbf{a},\, \lambda) \,=\, -\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\, \lambda) \,=\, +\,\mathscr{A}(\mathbf{a},\, \lambda)\,. \tag{57}$$

This tells us that in the $\psi \to 0$ limit the quaternion $\mathbf{q}(\psi,\, \mathbf{a},\, \lambda)$ reduces to the scalar point $-\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\, \lambda)$ of S^3. Moreover, we have $m(\mathscr{A}) = 0$, since $m(\mathbf{L_a}) = 0$ as we saw above. On the other hand, from the definition (20) of $\mathbf{q}(\psi,\, \mathbf{a},\, \lambda)$ we also have the following relation between the conjugate variables $\mathscr{A}^\dagger(\mathbf{a},\, \lambda)$ and $\mathbf{q}^\dagger(\psi = 2\pi,\, \mathbf{a},\, \lambda)$:

$$\mathbf{q}^\dagger(\psi = 2\pi,\, \mathbf{a},\, \lambda) \,=\, +\,\left\{\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\, \lambda)\,\right\}^\dagger \,=\, +\,\mathbf{L}^\dagger(\mathbf{a},\, \lambda)\,\mathbf{D}^\dagger(\mathbf{a}) \,=\, -\,\mathbf{L}^\dagger(\mathbf{a},\, \lambda)\,\mathbf{D}(\mathbf{a}) \,=\, -\,\mathscr{A}^\dagger(\mathbf{a},\, \lambda)\,. \tag{58}$$

This tells us that in the $\psi \to 2\,\pi$ limit the quaternion $\mathbf{q}^\dagger(\psi,\, \mathbf{a},\, \lambda)$ reduces to the scalar point $-\,\mathbf{L}^\dagger(\mathbf{a},\, \lambda)\,\mathbf{D}(\mathbf{a})$ of S^3. Thus the case $\psi = 0$ does indeed correspond to the measurement events. The physical significance of the two remaining cases, namely $\psi = 2\pi$ and 4π, can be verified similarly, confirming the set of results listed in (49):

$$\sigma(\mathscr{A}) \,=\, \pm\,\mathbf{D}(\mathbf{a})\,. \tag{59}$$

Substituting this and $\sigma(\mathscr{B}) \,=\, \pm\,\mathbf{D}(\mathbf{b})$ into Eqs. (30) and (31) then immediately leads to the standard scores:

$$A(\mathbf{a},\, \lambda) \,=\, \frac{\pm\,\mathscr{A}(\mathbf{a},\, \lambda) \,-\, \overline{\mathscr{A}(\mathbf{a},\, \lambda)}}{\sigma(\mathscr{A})} \,=\, \frac{\pm\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\, \lambda) \,-\, 0}{\sigma(\mathscr{A})} \,=\, \left\{\frac{\pm\,\mathbf{D}(\mathbf{a})}{\sigma(\mathscr{A})}\right\}\mathbf{L}(\mathbf{a},\, \lambda) \,=\, \mathbf{L}(\mathbf{a},\, \lambda) \tag{60}$$

$$\text{and}\quad B(\mathbf{b},\, \lambda) \,=\, \frac{\pm\,\mathscr{B}(\mathbf{b},\, \lambda) \,-\, \overline{\mathscr{B}(\mathbf{b},\, \lambda)}}{\sigma(\mathscr{B})} \,=\, \frac{\pm\,\mathbf{D}(\mathbf{b})\,\mathbf{L}(\mathbf{b},\, \lambda) \,-\, 0}{\sigma(\mathscr{B})} \,=\, \left\{\frac{\pm\,\mathbf{D}(\mathbf{b})}{\sigma(\mathscr{B})}\right\}\mathbf{L}(\mathbf{b},\, \lambda) \,=\, \mathbf{L}(\mathbf{b},\, \lambda)\,. \tag{61}$$

This confirms the standard scores derived in the equations (30) and (31) of the previous subsection.

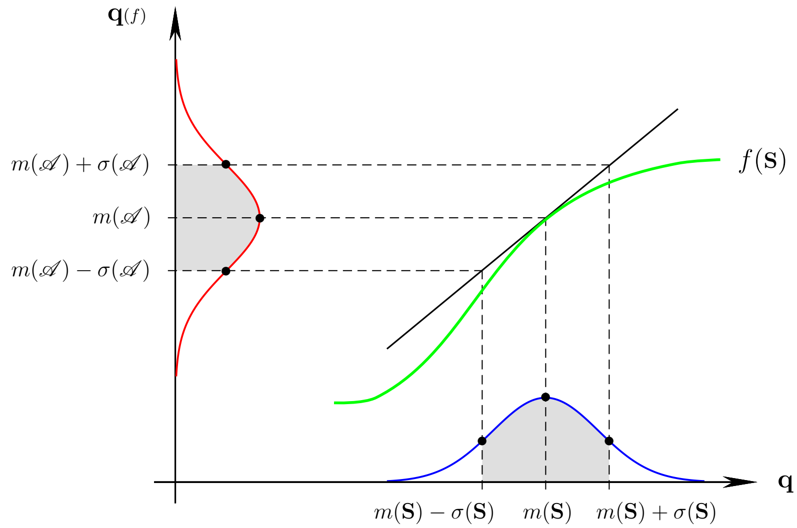

FIG. 1 Propagation of the 68% probability interval from a random bivector $\mathbf{S}$ to a scalar $\mathscr{A}$ within the parallelized 3-sphere.

So far I have assumed that randomness in the measurement results $\mathscr{A}(\mathbf{a}, \lambda)$ and $\mathscr{B}(\mathbf{b}, \lambda)$ originates entirely from the initial state λ representing the orientation of the 3-sphere. In other words, I have assumed that the local interaction of the detector $\mathbf{D}(\mathbf{a})$ with the random spin $\mathbf{L}(\mathbf{a}, \lambda)$ does not introduce additional randomness in the measurement result $\mathscr{A}(\mathbf{a}, \lambda)$. Any realistic interaction between $\mathbf{D}(\mathbf{a})$ and $\mathbf{L}(\mathbf{a}, \lambda)$, however, would inevitably introduce such a randomness, of purely local, experimental origin. We can model this randomness by introducing an additional random variable, say $r_{\mathbf{a}} \in [0,\, 1]$, not dependent on λ. Physically we can think of $r_{\mathbf{a}}$ as an alignment parameter between the detector bivector $\mathbf{D}(\mathbf{a})$ and the spin bivector $\mathbf{L}(\mathbf{a}, \lambda)$, with $r_{\mathbf{a}} = 1$ representing the perfect alignment. Clearly, introduction of this additional random parameter will make all the bivectors and quaternions unnormalized, and the corresponding probability density function (34) would then represent a Gaussian distribution—provided we also assume that the orientation $\lambda = \pm 1$ of S^3 itself is distributed non-uniformly between its values $+1$ and -1. Moreover, although the measurement results would then fall within the range $-1 \leq \mathscr{A}(\mathbf{a},\, \lambda) \leq +1$, their mean value would be zero for a uniformly distributed λ, since the mean value of the product of the independent random variables $r_{\mathbf{a}}$ and λ would then be the product of their mean values:

$$m(r_{\mathbf{a}}\, \lambda) \;=\; m(r_{\mathbf{a}})\, m(\lambda). \tag{62}$$

More importantly, the standard scores computed above in equations (60) and (61) would not be affected by this more realistic random process $r_{\mathbf{a}}\, \lambda$—at least for the special case of uniformly distributed λ—because they involve the ratios of the corresponding raw scores and standard deviations centered about the mean values $m(r_{\mathbf{a}}\, \lambda) = 0 = m(r_{\mathbf{b}}\, \lambda)$:

$$A(\mathbf{a},\, \lambda) \;=\; \frac{\pm\, \mathscr{A}(\mathbf{a},\, \lambda) \;-\; \overline{\mathscr{A}(\mathbf{a},\, \lambda)}}{\sigma(\mathscr{A})} \;=\; \frac{\pm\, r_{\mathbf{a}}\, \mathbf{D}(\mathbf{a})\, \mathbf{L}(\mathbf{a},\, \lambda) \;-\; 0}{\sigma(\mathscr{A})} \;=\; \left\{ \frac{\pm\, r_{\mathbf{a}}\, \mathbf{D}(\mathbf{a})}{\sigma(\mathscr{A})} \right\} \mathbf{L}(\mathbf{a},\, \lambda) \;=\; \mathbf{L}(\mathbf{a},\, \lambda) \tag{63}$$

$$\text{and} \quad B(\mathbf{b},\, \lambda) \;=\; \frac{\pm\, \mathscr{B}(\mathbf{b},\, \lambda) \;-\; \overline{\mathscr{B}(\mathbf{b},\, \lambda)}}{\sigma(\mathscr{B})} \;=\; \frac{\pm\, r_{\mathbf{b}}\, \mathbf{D}(\mathbf{b})\, \mathbf{L}(\mathbf{b},\, \lambda) \;-\; 0}{\sigma(\mathscr{B})} \;=\; \left\{ \frac{\pm\, r_{\mathbf{b}}\, \mathbf{D}(\mathbf{b})}{\sigma(\mathscr{B})} \right\} \mathbf{L}(\mathbf{b},\, \lambda) \;=\; \mathbf{L}(\mathbf{b},\, \lambda)\,. \tag{64}$$

Let us now try to understand the propagation of error within S^3 from this physically more realistic perspective. To this end, let the random variable $\mathbf{q}(\psi,\, \mathbf{a},\, \lambda) \in S^3(\lambda)$ be such that the measurement results $\mathscr{A}(\mathbf{a},\, \lambda) \in [-1,\, +1]$ remain as before, but the bivectors $\mathbf{L}(\mathbf{a},\, \lambda)$ are subject to a random process $r_{\mathbf{a}}\, \lambda$ such that $\mathbf{S}(\mathbf{a},\, \lambda,\, r_{\mathbf{a}}) = r_{\mathbf{a}}\, \mathbf{L}(\mathbf{a},\, \lambda)$ with $r_{\mathbf{a}} \in [0,\, 1]$. Then the mean value $m(\mathbf{S})$ and standard deviation $\sigma(\mathbf{S})$ of $\mathbf{S}$ would be a bivector and a scalar:

$$\begin{aligned} m(\mathbf{S}) \;&=\; \text{a bivector} \\ \text{and} \quad \sigma(\mathbf{S}) \;&=\; \text{a scalar.} \end{aligned} \tag{65}$$

If we now take the detector bivector to be $\mathbf{D}(\mathbf{a}) = I \cdot \mathbf{a}$ as before, then the measurement results can be identified as $-1 \leq \mathscr{A} = \mathbf{D}\, \mathbf{S} \leq +1$ so that $m(\mathscr{A}) \geq 0$. Since $\mathbf{D}$ is a non-random bivector, errors generated within $\mathscr{A}$ by the random process $r_{\mathbf{a}}\,\lambda$ would stem entirely from the random bivector $\mathbf{S}$, and propagate linearly. In other words, the standard deviations within the random number $\mathscr{A}$ due to the random process $r_{\mathbf{a}}\,\lambda$ would be given by

$$\sigma(\mathscr{A}) = \mathbf{D}\,\sigma(\mathbf{S}). \tag{66}$$

But since $\sigma(\mathbf{S})$, as we noted, is a scalar, the typical error $\sigma(\mathscr{A})$ generated within $\mathscr{A}$ due to the random process $r_{\mathbf{a}}\,\lambda$ is a bivector. The standardized variable (which must be used to compare the raw scores $\mathscr{A}$ with other raw scores $\mathscr{B}$) is thus also a bivector: $A := \mathscr{A}/\sigma(\mathscr{A}) = \text{scalar} \times \mathbf{S}$.

As straightforward as it is, the above conclusion may seem unusual. It is important to recall, however, that in geometric algebra both scalars and bivectors are treated on equal footing [12][13]. They both behave as real-valued c-numbers, albeit of different grades. To appreciate the consistency and naturalness of the above conclusion, let

$$\mathscr{A} \,=\, f(\mathbf{S}) \,=\, \mathbf{D}\,\mathbf{S} \tag{67}$$

be a continuous function generated by the geometric product of two bivectors $\mathbf{D}(\mathbf{a})$ and $\mathbf{S}(\mathbf{a},\,\lambda,\,r_{\mathbf{a}})$ as before. The natural question then is: How does a typical error in $\mathbf{S}$ governed by the probability density (34)—which can be represented by the 68% probability interval

$$[\,m(\mathbf{S}) - \sigma(\mathbf{S}),\; m(\mathbf{S}) + \sigma(\mathbf{S})\,] \tag{68}$$

as shown in the Fig. 1—propagate from the random bivector $\mathbf{S}$ to the random scalar $\mathscr{A}$, through the function $f(\mathbf{S}) = \mathbf{D}\,\mathbf{S}$? To answer this question we note that the two end points of the interval (68) represent two points, say $\mathbf{q}^-$ and $\mathbf{q}^+$, of the 3-sphere, which is a Riemannian manifold. The geometro-algebraic distance between the points $\mathbf{q}^-$ and $\mathbf{q}^+$ can therefore be defined, say, as

$$d\big(\mathbf{q}^-,\,\mathbf{q}^+\big) \,=\, \big(\mathbf{q}^- \,-\, \mathbf{q}^+\big) \times \operatorname{sign}\big(\mathbf{q}^- \,-\, \mathbf{q}^+\big)\,. \tag{69}$$

Moreover, from definition (67) of $\mathscr{A}$ and a first-order Taylor expansion of the function $f(\mathbf{S})$ about the point $\mathbf{S} = m(\mathbf{S})$ we obtain

$$\mathscr{A} \,=\, f(m(\mathbf{S})) \,+\, \frac{\partial f}{\partial \mathbf{S}}\bigg|_{\mathbf{S}\,=\,m(\mathbf{S})} (\mathbf{S} \,-\, m(\mathbf{S})) \,+\, \dots \tag{70}$$

Now it is evident that the slope $\partial f/\partial \mathbf{S} = \mathbf{D}$ of this line is a constant. Therefore the mean $m(\mathscr{A})$ and the standard deviation $\sigma(\mathscr{A})$ of the distribution of $\mathscr{A}$ can be obtained by setting $\mathbf{S} = m(\mathbf{S})$ and $\mathbf{S} = \sigma(\mathbf{S})$:

$$m(\mathscr{A}) \,=\, f(m(\mathbf{S})) \,=\, \mathbf{D}\,m(\mathbf{S}) \,=\, \text{a scalar} \tag{71}$$

$$\text{and}\;\; \sigma(\mathscr{A}) \,=\, \frac{\partial f}{\partial \mathbf{S}}\,\sigma(\mathbf{S}) \,=\, \mathbf{D}\,\sigma(\mathbf{S}) \,=\, \text{a bivector.} \tag{72}$$

The probability distribution of $\mathscr{A}$ is thus represented by the 68% interval

$$[\,m(\mathscr{A}) - \sigma(\mathscr{A}),\; m(\mathscr{A}) + \sigma(\mathscr{A})\,]\,. \tag{73}$$

If we now set $r_{\mathbf{a}} = 1$ and thereby assume that $\mathbf{S}$ is in fact the unit bivector $\mathbf{L}$ with a vanishing mean, then we have $m(\mathscr{A}) = 0$ and $\sigma(\mathscr{A}) = \pm\,\mathbf{D}$, as in equation (59) above.

Finally, it is instructive to note that, geometrically, the propagation of error within S^3 is equivalent to a simple change in the perspective (cf. Fig. 1):

$$S^3 \ni \underbrace{\overbrace{m(\mathbf{S})}^{\text{bivector}} \;\pm\; \overbrace{\sigma(\mathbf{S})}^{\text{scalar}}}_{\text{quaternion}} \quad \xrightarrow{\;f(\mathbf{S})\;} \quad \underbrace{\overbrace{m(\mathscr{A})}^{\text{scalar}} \;\pm\; \overbrace{\sigma(\mathscr{A})}^{\text{bivector}}}_{\text{quaternion}}\,. \tag{74}$$

In particular, the probability density of the scalar $\mathscr{A}$ over S^3 corresponding to interval (73) is equivalent to that of the bivector $\mathbf{S}$ over S^3 corresponding to interval (68). With this, we are now ready to derive the EPR-Bohm correlations.

E. Derivation of Pair Correlations Among the Points of S^3

We begin by noting that, according to my model, EPR-Bohm correlations are correlations among the points of a parallelized 3-sphere [3]. Now, since we have assumed that initially there was 50/50 chance between the right-handed and left-handed orientations of the 3-sphere (*i.e.*, equal chance between the initial states $\lambda = +1$ and $\lambda = -1$), the expectation values of the raw scores $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$ vanish identically. On the other hand, as discussed above, the correlation between these raw scores (or their first product moment coefficient *à la* Pearson [15]) can be obtained only by computing the covariance between the corresponding standardized variables $A(\mathbf{a},\,\lambda)$ and $B(\mathbf{b},\,\lambda)$, which gives

$$\begin{aligned}
\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,&=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A(\mathbf{a},\,\lambda^k)\,B(\mathbf{b},\,\lambda^k)\right] \\
&=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\left\{\,a_\mu\,L_\mu(\lambda^k)\,\right\}\,\left\{\,b_\nu\,L_\nu(\lambda^k)\,\right\}\right] \\
&=\, -\,g_{\mu\nu}\,a_\mu\,b_\nu\,-\,\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\left\{\,\epsilon_{\mu\nu\rho}\,a_\mu\,b_\nu\,L_\rho(\lambda^k)\,\right\}\right] \\
&=\, -\,g_{\mu\nu}\,a_\mu\,b_\nu\,-\,\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\lambda^k\right]\left\{\,\epsilon_{\mu\nu\rho}\,a_\mu\,b_\nu\,D_\rho\,\right\} \\
&=\, -\,g_{\mu\nu}\,a_\mu\,b_\nu\,-\,0\,,
\end{aligned} \tag{75}$$

where I have used algebra defined in (10) and the relation (18). Consequently, as explained in the paragraph just below Eq. (24), when the raw scores $\mathscr{A} = \pm\,1$ and $\mathscr{B} = \pm\,1$ are compared, their product moment will inevitably yield

$$\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathscr{A}(\mathbf{a},\,\lambda^k)\,\mathscr{B}(\mathbf{b},\,\lambda^k)\right] \,=\, -\,g_{\mu\nu}\,a_\mu\,b_\nu\,, \tag{76}$$

since the correlation between the raw scores $\mathscr{A}$ and $\mathscr{B}$ is equal to covariance between the standard scores A and B.

So far in this section we have put no restrictions on the metric tensor, which, in the normal coordinates centered at a point of S^3 would be of the form

$$g_{\mu\nu}(x) \,=\, \delta_{\mu\nu}\,-\,\frac{1}{3}\,\mathscr{R}_{\alpha\,\mu\,\nu\,\gamma}\,x^\alpha\,x^\gamma\,+\,O\left(|x|^3\right). \tag{77}$$

In other words, the algebra (10) we have used in the derivation of correlation (76) is a general Clifford algebra, with no restrictions placed on the quadratic form [16]. On the other hand, if the codomain of the measurement functions $\mathscr{A}(\mathbf{a},\,\lambda)$ is taken to be a parallelized 3-sphere, then the above metric tensor specializes to the Euclidean metric $\delta_{\mu\nu}$, because the Riemann curvature tensor of a parallelized 3-sphere vanishes, inducing a non-vanishing torsion [6]. This case corresponds to the geometry of the group SU(2) and specializes the correlation (76) to exhibit maximum strength:

$$\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,=\, -\,g_{\mu\nu}\,a_\mu\,b_\nu\,\longrightarrow\,-\,\delta_{\mu\nu}\,a_\mu\,b_\nu\,=\,-\,\cos\eta_{\mathbf{ab}}\,, \tag{78}$$

which in turn manifests the sensitivity of $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$ to spinorial sign changes. To appreciate the significance of these changes [6], recall from subsection II A that a parallelized 3-sphere is a set of unit quaternions of the form

$$\mathbf{q}(\psi,\,\mathbf{r}) \,=\, \cos\frac{\psi}{2}\,+\,\boldsymbol{\beta}(\mathbf{r})\,\sin\frac{\psi}{2}\,, \tag{79}$$

with ψ being the rotation angle. It is easy to check that $\mathbf{q}(\psi,\,\mathbf{r})$ respects the following rotational symmetries:

$$\mathbf{q}(\psi + 2\kappa\pi,\,\mathbf{r}) \,=\, -\,\mathbf{q}(\psi,\,\mathbf{r}) \;\; \text{for}\;\; \kappa = 1, 3, 5, 7, \dots \tag{80}$$

$$\mathbf{q}(\psi + 4\kappa\pi,\,\mathbf{r}) \,=\, +\,\mathbf{q}(\psi,\,\mathbf{r}) \;\; \text{for}\;\; \kappa = 0, 1, 2, 3, \dots \tag{81}$$

Thus $\mathbf{q}(\psi,\,\mathbf{r})$ correctly represents the state of a body that returns to itself only after even multiples of a 2π rotation.

It is very important to appreciate that the strong correlation derived in (78) are correlation among the points of a parallelized 3-sphere, S^3, taken as the codomain of the measurement functions $\mathscr{A}(\mathbf{r},\,\lambda)$. Thus the strength and the very existence of the EPR-Bohm correlation (or of *any* correlation for that matter) stem entirely from the topological properties of the codomain of the measurement functions $\mathscr{A}(\mathbf{r},\,\lambda)$. Had we chosen the codomain of $\mathscr{A}(\mathbf{r},\,\lambda)$ to be any manifold other than a parallelized 3-sphere, the resulting correlation would not have been as strong as $-\cos\eta_{\mathbf{ab}}$.

F. Derivation of Upper Bound Exceeding the Bell-CHSH Bound

Returning to the expectation value (76) in its most general form we can now proceed to derive the Bell-CHSH-type bound on possible correlations [3][17]. To this end, consider four observation axes, $\mathbf{a}$, $\mathbf{a}'$, $\mathbf{b}$, and $\mathbf{b}'$, for the standard EPR-Bohm experiment. Then the corresponding CHSH string of expectation values [3], namely the coefficient

$$\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}') \,:=\, \mathcal{E}(\mathbf{a},\,\mathbf{b}) \,+\, \mathcal{E}(\mathbf{a},\,\mathbf{b}') \,+\, \mathcal{E}(\mathbf{a}',\,\mathbf{b}) \,-\, \mathcal{E}(\mathbf{a}',\,\mathbf{b}')\,, \tag{82}$$

would be bounded by the constant $2\sqrt{2}$, as discovered by Tsirel'son within the setting of Clifford algebra applied to quantum mechanics in general [3][17]. Here each of the joint expectation values of the raw scores $\mathscr{A}(\mathbf{a},\,\lambda) = \pm\,1$ and $\mathscr{B}(\mathbf{b},\,\lambda) = \pm\,1$ are defined as

$$\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathscr{A}(\mathbf{a},\,\lambda^k)\,\mathscr{B}(\mathbf{b},\,\lambda^k)\right], \tag{83}$$

with the binary numbers such as $\mathscr{A}(\mathbf{a},\,\lambda)$ defined by the limit

$$S^3 \ni \pm\,1 \,=\, \mathscr{A}(\mathbf{a},\,\lambda) \,=\, \lim_{\psi\,\rightarrow\,2\kappa\pi}\,\{\,\mathbf{q}(\psi,\,\mathbf{a},\,\lambda)\,\} \,=\, -\,\mathbf{D}(\mathbf{a})\,\mathbf{L}(\mathbf{a},\,\lambda). \tag{84}$$

Thus $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$ are points of a parallelized 3-sphere and $\mathcal{E}(\mathbf{a},\,\mathbf{b})$ evaluated in (83) gives correlation between such points of the 3-sphere [3]. The correct value of the correlation, however, cannot be obtained without appreciating the fact that the number $\mathscr{A}(\mathbf{a},\,\lambda) = \pm\,1$ is defined as a product of a λ-*in*-dependent constant, namely $-\,\mathbf{D}(\mathbf{a})$, and a λ-dependent random variable, namely $\mathbf{L}(\mathbf{a},\,\lambda)$. Thus the correct value of the correlation is obtained by calculating the covariance of the corresponding standardized variables

$$A_{\mathbf{a}}(\lambda) \,\equiv\, A(\mathbf{a},\,\lambda) \,=\, \mathbf{L}(\mathbf{a},\,\lambda) \tag{85}$$

$$\text{and}\quad B_{\mathbf{b}}(\lambda) \,\equiv\, B(\mathbf{b},\,\lambda) \,=\, \mathbf{L}(\mathbf{b},\,\lambda)\,, \tag{86}$$

as we discussed just below equation (24). In other words, correlation between the raw scores $\mathscr{A}(\mathbf{a},\,\lambda)$ and $\mathscr{B}(\mathbf{b},\,\lambda)$ is obtained by calculated the covariance between the standard scores $A(\mathbf{a},\,\lambda)$ and $B(\mathbf{b},\,\lambda)$, as in equation (75) above:

$$\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A(\mathbf{a},\,\lambda^k)\,\,B(\mathbf{b},\,\lambda^k)\right] \,=\, -\,g_{\mu\nu}\,a_\mu\,b_\nu\,. \tag{87}$$

The correlation between the raw scores is thus necessarily equal to the covariance between the standard scores:

$$\boxed{\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathscr{A}(\mathbf{a},\,\lambda^k)\,\mathscr{B}(\mathbf{b},\,\lambda^k)\right] \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A(\mathbf{a},\,\lambda^k)\,\,B(\mathbf{b},\,\lambda^k)\right] \,=\, -\,g_{\mu\nu}\,a_\mu\,b_\nu\,.} \tag{88}$$

Using this identity we can now rewrite the CHSH string of expectation values (82) in two equivalent expressions,

$$\begin{aligned}\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}') \,=\, &\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathscr{A}_{\mathbf{a}}(\lambda^k)\,\mathscr{B}_{\mathbf{b}}(\lambda^k)\right] \,+\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathscr{A}_{\mathbf{a}}(\lambda^k)\,\mathscr{B}_{\mathbf{b}'}(\lambda^k)\right] \\ &+\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathscr{A}_{\mathbf{a}'}(\lambda^k)\,\mathscr{B}_{\mathbf{b}}(\lambda^k)\right] \,-\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\mathscr{A}_{\mathbf{a}'}(\lambda^k)\,\mathscr{B}_{\mathbf{b}'}(\lambda^k)\right]\end{aligned} \tag{89}$$

and

$$\begin{aligned}\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}') \,=\, &\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A_{\mathbf{a}}(\lambda^k)\,B_{\mathbf{b}}(\lambda^k)\right] \,+\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A_{\mathbf{a}}(\lambda^k)\,B_{\mathbf{b}'}(\lambda^k)\right] \\ &+\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A_{\mathbf{a}'}(\lambda^k)\,B_{\mathbf{b}}(\lambda^k)\}\right] \,-\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A_{\mathbf{a}'}(\lambda^k)\,B_{\mathbf{b}'}(\lambda^k)\right].\end{aligned} \tag{90}$$

Our goal now is to find the upper bound on these strings of expectation values. To this end, we first note that the four pairs of measurement results occurring in the above expressions do not all occur at the same time. Let us, however, conform to the usual assumption of counterfactual definiteness and pretend that they do occur at the same time, at least counterfactually, with equal distribution. This assumption allows us to simplify the above expressions as

$$\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}') \;=\; \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\left\{\mathscr{A}_{\mathbf{a}}(\lambda^k)\,\mathscr{B}_{\mathbf{b}}(\lambda^k)\,+\,\mathscr{A}_{\mathbf{a}}(\lambda^k)\,\mathscr{B}_{\mathbf{b}'}(\lambda^k)\,+\,\mathscr{A}_{\mathbf{a}'}(\lambda^k)\,\mathscr{B}_{\mathbf{b}}(\lambda^k)\,-\,\mathscr{A}_{\mathbf{a}'}(\lambda^k)\,\mathscr{B}_{\mathbf{b}'}(\lambda^k)\right\}\right] \tag{91}$$

and

$$\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}') \;=\; \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\left\{A_{\mathbf{a}}(\lambda^k)\,B_{\mathbf{b}}(\lambda^k)\,+\,A_{\mathbf{a}}(\lambda^k)\,B_{\mathbf{b}'}(\lambda^k)\,+\,A_{\mathbf{a}'}(\lambda^k)\,B_{\mathbf{b}}(\lambda^k)\,-\,A_{\mathbf{a}'}(\lambda^k)\,B_{\mathbf{b}'}(\lambda^k)\right\}\right]. \tag{92}$$

The obvious question now is: Which of these two expressions should we evaluate to obtain the correct bound on $\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}')$? Clearly, in view of the identity (88) both expressions would give one and the same answer [3]. Thus it should not matter which of the two expressions we use to evaluate the bound. But it is also clear from the discussion in subsections II C and II D that the correct bound on the expression (91) involving the raw scores $\mathscr{A}$ and $\mathscr{B}$ can only be obtained by evaluating the expression (92) involving the standard scores A and B. Stated differently, if we tried to obtain the bound on $\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}')$ by disregarding how the measurement results have been generated in the model statistically, then we would end up getting a wrong answer. By following the Bell-CHSH reasoning blindly Weatherall ends up making such a mistake. In the end $\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}')$ is a functional of a *random* variable, and as such proper statistical procedure tailored to my model must be employed for its correct evaluation. This is an important point and the reader is urged to review the discussions in subsections II C and II D once again to appreciate its full significance.

With these remarks in mind we proceed to obtain the upper bound on $\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}')$ by evaluating the expression (92) as follows. Since the standard scores $A_{\mathbf{a}}(\lambda)=\mathbf{L}(\mathbf{a},\,\lambda)$ and $B_{\mathbf{b}}(\lambda)=\mathbf{L}(\mathbf{b},\,\lambda)$ appearing in this expression represent two independent equatorial points of the 3-sphere, we can take them to belong to two disconnected "sections" of S^3 (*i.e.*, two disconnected 2-spheres within S^3), satisfying

$$[\,A_{\mathbf{r}}(\lambda),\,B_{\mathbf{r}'}(\lambda)\,]\;=\;0\quad\forall\;\mathbf{r}\;\text{and}\;\mathbf{r}'\,\in\,\mathrm{I\!R}^3, \tag{93}$$

which is equivalent to anticipating a null outcome along the direction $\mathbf{r}\times\mathbf{r}'$ exclusive to both $\mathbf{r}$ and $\mathbf{r}'$. If we now square the integrand of equation (92), use the above commutation relations, and use the fact that all bivectors square to -1, then the absolute value of $\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}')$ leads to the following variance inequality [3]:

$$\begin{aligned}|\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}')| \;&=\; |\mathcal{E}(\mathbf{a},\,\mathbf{b})\,+\,\mathcal{E}(\mathbf{a},\,\mathbf{b}')\,+\,\mathcal{E}(\mathbf{a}',\,\mathbf{b})\,-\,\mathcal{E}(\mathbf{a}',\,\mathbf{b}')| \\ &\leqslant\;\sqrt{\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,\left\{\,4\,+\,4\,\mathcal{T}_{\mathbf{a}\,\mathbf{a}'}(\lambda^k)\,\mathcal{T}_{\mathbf{b}'\,\mathbf{b}}(\lambda^k)\,\right\}\right]}\,,\end{aligned} \tag{94}$$

where the classical commutators

$$\mathcal{T}_{\mathbf{a}\,\mathbf{a}'}(\lambda) := \frac{1}{2}\,[\,A_{\mathbf{a}}(\lambda),\,A_{\mathbf{a}'}(\lambda)]\;=\;-\,A_{\mathbf{a}\times\mathbf{a}'}(\lambda) \tag{95}$$

and

$$\mathcal{T}_{\mathbf{b}'\,\mathbf{b}}(\lambda) := \frac{1}{2}\,[\,B_{\mathbf{b}'}(\lambda),\,B_{\mathbf{b}}(\lambda)]\;=\;-\,B_{\mathbf{b}'\times\mathbf{b}}(\lambda) \tag{96}$$

are the geometric measures of the torsion within S^3. Thus, it is the non-vanishing torsion $\mathcal{T}$ within the parallelized 3-sphere—the parallelizing torsion which makes its Riemann curvature tensor vanish—that is ultimately responsible for the strong quantum correlation [3][7]. We can see this at once from Eq. (94) by setting $\mathcal{T}=0$, and in more detail as follows: Using definitions (85) and (86) for $A_{\mathbf{a}}(\lambda)$ and $B_{\mathbf{b}}(\lambda)$ and making a repeated use of the bivector identity

$$\mathbf{L}(\mathbf{a},\,\lambda)\,\mathbf{L}(\mathbf{a}',\,\lambda)\;=\;-\,\mathbf{a}\cdot\mathbf{a}'\,-\,\mathbf{L}(\mathbf{a}\times\mathbf{a}',\,\lambda) \tag{97}$$

specialized for the metric $g_{\mu\nu}=\delta_{\mu\nu}$ on S^3, the above inequality for $\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}')$ can be further simplified to

$$\begin{aligned}|\mathcal{S}(\mathbf{a},\,\mathbf{a}',\,\mathbf{b},\,\mathbf{b}')|\;&\leqslant\;\sqrt{4-4\,(\mathbf{a}\times\mathbf{a}')\cdot(\mathbf{b}'\times\mathbf{b})-4\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\mathbf{L}(\mathbf{z},\,\lambda^k)\right]} \\ &\leqslant\;\sqrt{4-4\,(\mathbf{a}\times\mathbf{a}')\cdot(\mathbf{b}'\times\mathbf{b})-4\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\lambda^k\right]\mathbf{D}(\mathbf{z})} \\ &\leqslant\;2\sqrt{1-(\mathbf{a}\times\mathbf{a}')\cdot(\mathbf{b}'\times\mathbf{b})\,-\,0}\;,\end{aligned} \tag{98}$$

where $\mathbf{z} = (\mathbf{a} \times \mathbf{a}') \times (\mathbf{b}' \times \mathbf{b})$, and—as before—I have used the relation (18) between $\mathbf{L}(\mathbf{z},\, \lambda)$ and $\mathbf{D}(\mathbf{z})$ from subsection II B. Finally, by noticing that $(\mathbf{a} \times \mathbf{a}') \cdot (\mathbf{b}' \times \mathbf{b})$ is bounded by trigonometry as

$$-\,1 \leqslant\, (\mathbf{a} \times \mathbf{a}') \cdot (\mathbf{b}' \times \mathbf{b}) \,\leqslant +1\,, \tag{99}$$

the above inequality can be reduced to the form

$$-\,2\sqrt{2} \,\leqslant\, \mathcal{E}(\mathbf{a},\,\mathbf{b}) \,+\, \mathcal{E}(\mathbf{a},\,\mathbf{b}') \,+\, \mathcal{E}(\mathbf{a}',\,\mathbf{b}) \,-\, \mathcal{E}(\mathbf{a}',\,\mathbf{b}') \,\leqslant +\,2\sqrt{2}\,, \tag{100}$$

which exhibits an extended upper bound on possible correlations. Thus, when in an EPR-Bohm experiment raw scores $\mathscr{A} = \pm\,1$ and $\mathscr{B} = \pm\,1$ are compared by coincidence counts [18], the normalized expectation value of their product

$$\mathcal{E}(\mathbf{a},\,\mathbf{b}) \,=\, \frac{\Big[C_{++}(\mathbf{a},\,\mathbf{b}) \,+\, C_{--}(\mathbf{a},\,\mathbf{b}) \,-\, C_{+-}(\mathbf{a},\,\mathbf{b}) \,-\, C_{-+}(\mathbf{a},\,\mathbf{b})\Big]}{\Big[C_{++}(\mathbf{a},\,\mathbf{b}) \,+\, C_{--}(\mathbf{a},\,\mathbf{b}) \,+\, C_{+-}(\mathbf{a},\,\mathbf{b}) \,+\, C_{-+}(\mathbf{a},\,\mathbf{b})\Big]} \tag{101}$$

is predicted by my model to respect, not the Bell-CHSH upper bound 2, but the Tsirel'son upper bound $2\sqrt{2}$, where $C_{+-}(\mathbf{a},\,\mathbf{b})$ *etc.* represent the number of joint occurrences of detections $+\,1$ along $\mathbf{a}$ and $-\,1$ along $\mathbf{b}$ *etc.*

This completes the presentation of my local, realistic, and deterministic model for the EPR-Bohm correlation.

III. PHYSICAL, MATHEMATICAL, AND CONCEPTUAL FALLACIES OF WEATHERALL'S MODEL

1. What is wrong with Weatherall's measurement ansatz?

With the successful model firmly in place, we are now in a position to understand why Weatherall's model fails. To begin with, his model is based on a different representation of the rotation group. It is in fact based, not on the spinorial rotation group SU(2), but something akin to its tensorial cousin SO(3), which is a group of all proper rotations in $\mathrm{I\!R}^3$, *insensitive* to spinorial sign changes. In fact, Weatherall takes a rather odd space, namely $\mathrm{I\!R}^3 \wedge \mathrm{I\!R}^3$, for the codomain of his measurement functions $\mathscr{A}(\mathbf{r},\,\lambda)$, and then introduces another projection map to arrive at the measurement results $\{-1,\,+1\}$. Compared to my measurement functions (16) obtained through a smooth limiting process (21), his two-step measurement process is rather artificial. Moreover, since his is not a simply-connected, parallelized codomain such as S^3 that remains closed under multiplication, it cannot possibly satisfy the completeness criterion of EPR [3][7]. It is therefore not surprising why Weatherall is unable to find strong correlation among its points. Moreover, at the end of his two-step process his measurement results $\{-1,\,+1\}$ are no longer the image points within the codomain $\mathrm{I\!R}^3 \wedge \mathrm{I\!R}^3$ of his measurement functions. This is in sharp contrast with the situation in my model, where my measurement results $\{-1,\,+1\}$ remain very much a part of the codomain S^3 of my measurement functions. The reason why this comes about naturally within my model is because S^3, the set of unit quaternions, is a simply-connected surface embedded in $\mathrm{I\!R}^4$ that is equipped with a *graded* basis made of both scalars and bivectors:

$$\{1,\; \mathbf{e}_2 \wedge \mathbf{e}_3,\; \mathbf{e}_3 \wedge \mathbf{e}_1,\; \mathbf{e}_1 \wedge \mathbf{e}_2\}\,. \tag{102}$$

Thus the scalars $\{-1,\,+1\}$ are as much a part of S^3 as the bivectors $\mathbf{L}(\mathbf{r},\,\lambda)$ are, regulated by these unified basis [3]. I am tempted to quip: *What Nature has joined together, let no man put asunder*. By contrast Weatherall's codomain is disconnected between the space $\mathrm{I\!R}^3 \wedge \mathrm{I\!R}^3$ of "bivectors" and the set $\{-1,\,+1\}$ of scalars. His image points can thus be at best either bivectors or scalars, but not both. It is a disjoint world, more like the world of quantum mechanics.

Let us, however, be more charitable to Weatherall. Let us grant him the codomain of his measurement functions to be the connected real projective space $\mathrm{I\!RP}^3$, which is homeomorphic to the rotation group SO(3). After all, he does mention the Lie algebra $so(3)$ in one of his footnotes. So let us grant him the smooth one-step measurement process

$$\pm\;1 \,=\, \mathscr{A}(\mathbf{r},\,\lambda) : \mathrm{I\!R}^3 \times \Lambda \longrightarrow \mathrm{I\!RP}^3 \sim \mathrm{SO}(3) \tag{103}$$

to reach the image subset $\{-1,\,+1\}$. This smooth map is well-defined within my model, since $\mathrm{I\!RP}^3$ is simply the set S^3 of unit quaternions (cf. Eq. (4)) with each point identified with its antipodal point [6]. The measurement results $\pm\,1 \in \mathrm{I\!RP}^3 \sim \mathrm{SO}(3)$ are thus limiting points of a quaternion, just as in equation (21). The map that takes us from S^3 to $\mathrm{I\!RP}^3$ can now be used to project the metric $\delta_{\mu\nu}$ on S^3 onto $\mathrm{I\!RP}^3$ to obtain the following induced metric on $\mathrm{I\!RP}^3$:

$$-\,J_{\mu\nu}\,a_\mu\,b_\nu \,=\, \begin{cases} -\,1 \,+\, \frac{2}{\pi}\,\eta_{\mathbf{ab}} & \text{if} \quad 0 \leq \eta_{\mathbf{ab}} \leq \pi \\[2em] +\,3 \,-\, \frac{2}{\pi}\,\eta_{\mathbf{ab}} & \text{if} \quad \pi \leq \eta_{\mathbf{ab}} \leq 2\pi\,. \end{cases} \tag{104}$$

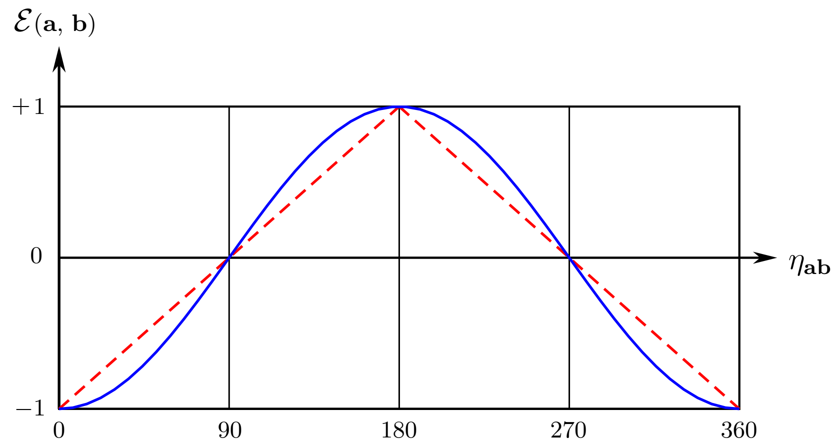

FIG. 2 Local-realistic correlations among the points of a parallelized 3-sphere can be stronger-than-classical but not quantum.

Further details of how this metric is obtained from the metric on S^3 can be found in section III of Ref. [6]. The two metrics $\delta_{\mu\nu}$ and $J_{\mu\nu}$ thus provide relative measures of geodesic distances on the manifolds S^3 and $\mathrm{I\!RP}^3$, respectively. Substituting the metric on $\mathrm{I\!RP}^3$ into equation (75) the correlation between the points of SO(3) then works out to be

$$\mathcal{E}(\mathbf{a},\,\mathbf{b})\;=\;-\,J_{\mu\nu}\,a_\mu\,b_\nu\;=\;\begin{cases}-\,1\,+\,\frac{2}{\pi}\,\eta_{\mathbf{ab}} & \text{if}\;\;0\leq\eta_{\mathbf{ab}}\leq\pi\\ \\ +\,3\,-\,\frac{2}{\pi}\,\eta_{\mathbf{ab}} & \text{if}\;\;\pi\leq\eta_{\mathbf{ab}}\leq 2\pi\,.\end{cases}\tag{105}$$

The two sets of correlations, (78) and (105), are compared in Fig. 2. The general correlation function $\mathcal{E}(\mathbf{a},\,\mathbf{b})$ derived in equation (75) can thus serve to distinguish the geodesic distances $\mathscr{D}(\mathbf{a},\,\mathbf{b})$ on the groups SU(2) and SO(3) [6].

2. Why did we lose the strong correlation for SO(3)?

It is crucial to appreciate that even when we choose SO(3) as the codomain of the function $\mathscr{A}(\mathbf{r},\,\lambda)$ the correct statistical procedure that must be followed is the one described in subsections II C and II D above. This is because, as Weatherall himself notes, the Lie algebras of SU(2) and SO(3) are isomorphic to each other. In other words, the local algebraic or tangent space structures on SU(2) and SO(3) are identical, but not their metrical structures in the sense of geodesic distances. Thus the above statistical procedure, tailored to the graded basis (102), leading up to the general expression (76) for correlations and beyond, is equally inevitable in the case of SO(3). Comparing the two sets of correlations resulting from this procedure—one for the prescription (3) and other for the prescription (103)—it is then easy to see why we have lost the strong correlation in the second case. We started out with S^3 as a codomain of $\mathscr{A}(\mathbf{r},\,\lambda)$ and then, for the case of SO(3), we identified each point of S^3 with its antipodal point. But in doing so we lost the following spinorial rotation symmetry satisfied by $\mathbf{q}(\psi,\,\mathbf{r})$, as described in equations (80) and (81) above:

$$\mathbf{q}(\psi+2\kappa\pi,\,\mathbf{r})\;=\;-\,\mathbf{q}(\psi,\,\mathbf{r})\;\;\text{for}\;\;\kappa=1,3,5,7,\dots\tag{106}$$

In other words, by identifying the antipodal points of S^3 we lost the sensitivity to spinorial sign changes. As a result, $\mathbf{q}(\psi,\,\mathbf{r})$ now represents the state of a rotating body that returns to itself after any and all multiples of 2π rotation:

$$\mathbf{q}(\psi+2\kappa\pi,\,\mathbf{r})\;=\;+\,\mathbf{q}(\psi,\,\mathbf{r})\;\;\text{for any}\;\;\kappa=0,1,2,3,\dots\tag{107}$$

This is the real reason why we lost the strong correlation for the SO(3) case. The reason Weatherall has argued for is an artifact of his bad choice of measurement functions. It stems from a failure to appreciate the unified nature of the graded basis (102) and the associated fact that the scalars $\{-1,\,+1\}$ and the bivectors $\mathbf{L}(\mathbf{r},\,\lambda)$ occur as image points within the same codomain S^3 in my model. Thus the loss of correlation has nothing to do with the fact that ultimately the measurement functions must map to the image subset $\{-1,\,+1\}$. They *manifestly do* in my model [cf. Eqs. (16), (17), and (21)]. The raw scalar numbers $\mathscr{A} = \pm 1$ and $\mathscr{B} = \pm 1$ mapped to the image subset $\{-1, +1\}$—according to my model—are indeed the numbers used by Alice and Bob for calculating the correlation in the usual manner. And when, at the end of their experiment, they evaluate the statistical quantity $\mathcal{S}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}')$ involving these numbers as

$$\mathcal{S}(\mathbf{a}, \mathbf{a}', \mathbf{b}, \mathbf{b}') = \lim_{n \gg 1} \left[\frac{1}{n} \sum_{k=1}^{n} \left\{ \mathscr{A}_{\mathbf{a}}(\lambda^k)\, \mathscr{B}_{\mathbf{b}}(\lambda^k) + \mathscr{A}_{\mathbf{a}}(\lambda^k)\, \mathscr{B}_{\mathbf{b}'}(\lambda^k) + \mathscr{A}_{\mathbf{a}'}(\lambda^k)\, \mathscr{B}_{\mathbf{b}}(\lambda^k) - \mathscr{A}_{\mathbf{a}'}(\lambda^k)\, \mathscr{B}_{\mathbf{b}'}(\lambda^k) \right\} \right], \tag{108}$$

they will inevitably find that it exceeds the bound of 2 and extends to the bound of $2\sqrt{2}$. This conclusion may seem odd from the perspective of Bell-type reasoning, but the evidence presented for it in subsection II F is incontrovertible.

IV. CONCLUDING REMARKS

In addition to the main issue discussed above, it is instructive to reflect on the broader reasons why Weatherall's model fails. We can in fact identify at least six erroneous steps which engender the failure of his model from the start:

1. Choice of incomplete codomain of the measurement functions $\mathscr{A}(\mathbf{r}, \lambda)$ (although with correct image points ± 1).

2. Neglect of putting scalars and bivectors on equal footing within a single, comprehensive, real number system.

3. Failure to implement a spinor representation of SU(2) by recognizing the significance of spinorial sign changes.

4. Lack of appreciation of the role played by the parallelizing torsion within S^3 for the existence and strength of strong correlations (or not recognizing the discipline of parallelization as the true source of strong correlations).

5. Failure to appreciate how errors propagate within S^3, when taken as a codomain of the measurement functions.

6. Neglect of the correct statistical procedure in the derivations of both the correlation and the upper bound $2\sqrt{2}$.

Although interconnected, any one of these reasons is sufficient for the failure of Weatherall's model. Recognizing this, I must conclude that, contrary to first impressions, Weatherall's thinly veiled criticism of my work is entirely vacuous.

Acknowledgments

I am grateful to Lucien Hardy for several months of correspondence which led to improvements in section II D. I am also grateful to Martin Castell for his hospitality in the Materials Department of the University of Oxford. This work was funded by a grant from the Foundational Questions Institute (FQXi) Fund, a donor advised fund of the Silicon Valley Community Foundation on the basis of proposal FQXi-MGA-1215 to the Foundational Questions Institute.

References

[1] J. O. Weatherall, *The Scope and Generality of Bell's Theorem*, arXiv:1212.4854
[2] J. Christian, *Disproof of Bell's Theorem*, arXiv:1103.1879. See also http://libertesphilosophica.info/blog/.
[3] J. Christian, *Disproof of Bell's Theorem: Illuminating the Illusion of Entanglement* (BrownWalker Press, Florida, 2012).
[4] J. Christian, *Refutation of Some Arguments Against my Disproof of Bell's Theorem*, arXiv:1110.5876
[5] J. Christian, *Refutation of Richard Gill's Argument Against my Disproof of Bell's Theorem*, arXiv:1203.2529
[6] J. Christian, *Macroscopic Observability of Spinorial Sign Changes under 2pi Rotations*, arXiv:1211.0784
[7] J. Christian, *What Really Sets the Upper Bound on Quantum Correlations?*, arXiv:1101.1958
[8] J. Christian, *On the Origins of Quantum Correlations*, arXiv:1201.0775. See also http://libertesphilosophica.info/blog/.
[9] J. S. Bell, Physics **1**, 195 (1964); see also Rev. Mod. Phys. **38**, 447 (1966), and Dialectica **39**, 86 (1985).
[10] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[11] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[12] C. Doran and A. Lasenby, *Geometric Algebra for Physicists* (Cambridge University Press, Cambridge, 2003).
[13] D. Hestenes, Am. J. Phys. **71**, 104 (2003).
[14] J. W. Milnor, *Topology from the Differentiable Viewpoint* (Princeton University Press, Princeton, NJ, 1965).
[15] J. L. Rodgers and W. A. Nicewander, The American Statistician **42**, 59 (1988).
[16] T. Frankel, *The Geometry of Physics: An Introduction* (Cambridge University Press, 1997), p 501.
[17] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980); L. J. Landau, Phys. Lett. A **120**, 54 (1987).
[18] A. Aspect *et al.*, Phys. Rev. Lett. **49**, 91 (1982); G. Weihs *et al.*, Phys. Rev. Lett. **81**, 5039 (1998).