Fake uniformity in a shape inversion formula

Christian Rau
Shantou University

Abstract. We revisit a shape inversion formula derived by Panaretos in the context of a particle density estimation problem with unknown rotation of the particle. A distribution is presented which imitates, or "fakes", the uniformity or Haar distribution that is part of that formula.

1 Introduction

Stochastic geometry makes extensive use of uniform, or Haar distributed, rotations; for example, in constructing random geometric objects (hyperplanes, polytopes, or other objects) that have a uniform orientation. The uniformity often makes it possible to reconstruct a three-dimensional object from two-dimensional projections or sections; such tasks belong to the realm of stereology, an area which is connected to both stochastic geometry and spatial statistics. This paper is a cautionary note, illustrating that the said uniformity may, in the case of a natural geometric functional, be imitated, or "faked", by a non-uniform distribution.

The natural functional that we consider here is a central ingredient in a *shape inversion* formula from Panaretos (2009), which he used to tackle a particle reconstruction problem arising from electron microscopy data. This formula, which is given in Proposition 1 below, allows to recover salient features, which may be referred to as "landmarks", of an object in $\mathbb{R}^d$ ($d \geq 2$) from its projections on an arbitrary fixed $(d-1)$-dimensional subspace. An important complicating feature of the problem is that the object is subject to a prior random and unknown uniform rotation before it is imaged. For practically relevant imaging tasks, where the object may be a protein fragment, we have $d = 3$, and the subspace is the imaging plane of the microscope. In this introduction, we limit ourselves to those notions that are necessary to formulate the shape inversion formula and our theorem, and refer to Section 4 for a precise statement of the statistical model.

The precise definition of uniform or Haar (probability) measure on the rotation group is as follows. Recall that $SO(d)$, the group of $d \times d$ rotation matrices, consists of those real matrices A satisfying $A^\mathsf{T} A = I$, the $d \times d$ identity matrix, and $\det A = 1$. A random rotation A has the Haar distribution, written as $A \sim \mu$, if the distribution of QA is the same as that of A, for any nonrandom $Q \in SO(d)$.

Here we examine how—and indeed, if—the formula given in Proposition 1 changes if one replaces μ by a less symmetric rotation. A motivation for this question comes from the fact that the group $SO(d)$ *acts* on the d-dimensional unit sphere $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : \|x\| = 1\}$ via matrix-vector multiplication $x \mapsto Ax$. In measure theory, this connection between $SO(d)$ and $\mathbb{S}^{d-1}$ is frequently exploited, and the present paper may be seen as another modest instance. The following fact, which is relevant to this paper, illustrates this connection: if $A \sim \mu$ and v is any nonrandom point on $\mathbb{S}^{d-1}$, then Av has the uniform distribution σ (normalized Lebesgue surface area measure) on $\mathbb{S}^{d-1}$. We note in passing that one may conversely construct μ from σ [Schneider and Weil (2008), pages 584–585].

In Section 2, after some preliminaries, we state the shape inversion formula of Panaretos (2009), and formulate the fake uniformity problem. In Section 3, the Cayley distribution with $\kappa = 1$ is found to be a distribution which fakes uniformity for the Gram matrix functional. Finally, Section 4 gives the statistical model in detail, and discusses the impact of fake uniformity.

2 (Closely) faking the value of a functional

We begin by describing a cousin of the problem studied in the present note. This cousin problem is mathematically more sophisticated but, on the other hand, does not make any reference to rotations. For a compact convex set $K \subset \mathbb{R}^d$ and $u \in \mathbb{S}^{d-1}$, denote by K_u the orthogonal projection of K onto the hyperplane $\{x \in \mathbb{R}^d : \langle x, u \rangle = 0\}$, where $\langle \cdot, \cdot \rangle$ is the usual inner product. Let $v_{d-1}(\cdot)$ denote $(d-1)$-dimensional volume. Consider the functional

$$s(\Phi, K) = \int_{\mathbb{S}^{d-1}} v_{d-1}(K_u)\Phi(u)\,d\sigma(u), \qquad (2.1)$$

where Φ is an integrable function on $\mathbb{S}^{d-1}$. Since we consider projections, we may assume Φ to satisfy $\Phi(u) = \Phi(-u)$ for all u. For the constant function $\Phi \equiv 1$, Cauchy's surface area formula [Groemer (1996), pages 45–46], yields that $s(\Phi, K)$ is the Lebesgue surface area of K. (Note that our σ, unlike in Groemer (1996), is already normalised.) In Groemer (1996), pages 297ff, the following *inverse problem* was considered: If $s(\Phi, K)$ is close to $s(1, K)$ for all K with surface area bounded by some constant, can it be inferred that Φ is close to 1? The answer is negative, unless smoothness assumptions on Φ are imposed; Groemer (1996) proved this with tools from harmonic analysis on $\mathbb{S}^{d-1}$.

The "fake uniformity" from the title of the present note is a negative answer to the cousin problem on $SO(d)$ which we shall formulate at the end of this section, and a manifestation of a comment on ill-definedness in Panaretos (2009), page 3303. Our result is not of a limiting nature, and will not require harmonic analysis for its proof.

The shape inversion formula of Panaretos (2009) is as follows. Let V be any real $d \times \ell$ matrix ($d \geq 2$, $\ell \geq 1$), and $H = \mathrm{diag}(1, \ldots, 1, 0)$ be the projection matrix

(with respect to the standard basis of $\mathbb{R}^d$) which zeroes the last component of a vector in $\mathbb{R}^d$. We interpret the columns of V as the location vectors of the so-called *landmarks* associated with the unknown particle to be reconstructed.

Proposition 1 (Panaretos (2009), Theorem 4.1).

$$\int_{SO(d)} \text{Gram}(HAV)\mu(\text{d}A) = \frac{d-1}{d}\text{Gram}(V),$$

where $\text{Gram}(W) = W^{\mathsf{T}}W$ *for any matrix* W, *so that the entry with index* (i,j) *in* $\text{Gram}(W)$ *is the inner product of the* ith *and* jth *columns of* W.

Note that $\text{Gram}(AW) = \text{Gram}(W)$ for any $A \in O(d)$, the group of $d \times d$ orthogonal matrices. Hence, $\text{Gram}(V)$ "nearly" encodes *shape* if the latter were understood to be the information which remains if "we are not interested in location, orientation or scale of the resulting configuration" [Kendall (1977), page 428]—however, information on reflections is lost, and this will be seen to be the reason for fake uniformity.

Proposition 1 says that the original shape can be reconstructed from the projected shape. As noted in Panaretos (2009), page 3286, this feature is shared with (2.1). We shall replace μ in Proposition 1 by a distribution which has the weaker symmetry property of *conjugation-invariance*.

Definition 1. A random rotation A is conjugation-invariant if $Q^{\mathsf{T}}AQ$ has the same distribution as A, for any nonrandom $Q \in SO(d)$.

If we assume that A has a density f with respect to μ, then conjugation-invariance may be expressed by the requirement that $f(Q^{\mathsf{T}}PQ) = f(P)$ for all Q as in Definition 1. We note that conjugation-invariant functions are also called *central* [Faraut (2008), page 132].

In the case $d = 3$, conjugation-invariance has the following geometric meaning. Recall that Euler's theorem says that every rotation in $\mathbb{R}^3$ has an axis and angle; if the rotation is assumed to be counter-clockwise and in the interval $(0, \pi)$, then the orientation of the axis is given by a well-defined vector in $\mathbb{S}^2$. It can be shown that the conjugation-invariant rotations in $\mathbb{R}^3$ which have a density with respect to Haar measure μ are precisely those for which (i) the oriented rotation axis U is uniformly distributed on $\mathbb{S}^2$, and (ii) the rotation angle Θ has a density, and Θ and U are independent [Schindler (1997), Theorem 2.2, page 109].

Let us call the Gram matrix $\text{Gram}(W)$ a "functional" of a given matrix W, even though this is an abuse of terminology, since functionals are usually scalar-valued. (This reservation is not serious, see Remark 2 below.) For the Gram functional, can the role of μ in Proposition 1 be faked by another distribution? We shall affirm this by stating an offending alternative random rotation in the next section.

3 Cayley distribution and main theorem

The Cayley distribution was introduced in Schaeben (1997) under the name of de la Vallée Poussin distribution, and independently in León, Massé and Rivest (2006); we adopt the name from the latter reference, as it has also been used in the R package documentation of Stanfill, Hofmann and Genschel (2014). We state the Cayley density for the case $d = 3$. This is the case where we can give an explanation of the result in Theorem 1 through geometry (see the discussion around (3.6)); however, the proof itself, which is based on integration on $\mathbb{S}^{d-1}$, carries over to the general case $d \geq 3$. The Cayley distribution is given by the density ($\Gamma(\cdot)$ is the Gamma function and $\mathrm{tr}(\cdot)$ is the trace)

$$f_\kappa^{\mathrm{CAY}}(R) = \frac{\sqrt{\pi}\,\Gamma(\kappa+2)}{2^{2\kappa}\Gamma(\kappa+1/2)}\left(1+\mathrm{tr}(R)\right)^\kappa$$

$$= \frac{\sqrt{\pi}\,\Gamma(\kappa+2)}{2^\kappa\Gamma(\kappa+1/2)}(1+\cos\theta)^\kappa, \qquad 0 \leq \theta \leq \pi,$$

which only depends on the rotation angle θ, and thus is conjugation-invariant; the parameter $\kappa \geq 0$ measures spread around the median I, with the case $\kappa = 0$ corresponding to μ. The density of Θ is [León, Massé and Rivest (2006), page 424],

$$f_\Theta^{\mathrm{CAY}}(\theta) = \frac{\Gamma(\kappa+2)}{\sqrt{\pi}\,2^\kappa\Gamma(\kappa+1/2)}(1+\cos\theta)^\kappa(1-\cos\theta), \qquad 0 \leq \theta \leq \pi, \quad (3.1)$$

where the factor $(1 - \cos\theta)$ shows the preference of μ for large rotations, see also Schindler (1997), Remark 2.4, pages 109–110. Write e_k for the kth column of I ($k = 1, \ldots, d$), where the dimension d will always be clear from the context, and write f_{v_R} for the density of $v_R = Re_d$ with respect to σ, where R is any conjugation-invariant distribution with a density with respect to μ. For the Cayley distribution, the density f_{v_R} is known for any d, see León, Massé and Rivest (2006), Prop. 3.3, page 419; in particular

$$f_{v_R}^{\mathrm{CAY}}(w) = 2^{-\kappa}(\kappa+1)(1+w_3)^\kappa, \qquad w = (w_2, w_2, w_3) \in \mathbb{S}^2. \quad (3.2)$$

A density f on $\mathbb{S}^{d-1}$ that depends only on w_d is called zonal (with respect to e_d). Then $f(w) = \overline{f}(t)$ for a function $\overline{f}(t)$ which is defined on the interval $[-1, 1]$, and integrates to a constant (depending on d) which can be evaluated with formula (3.5) below.

We now extend Proposition 1 to conjugation-invariant rotations.

Theorem 1. *Let $P \in SO(d)$ be a rotation with Haar density f_P, and such that there exists a nonrandom $M \in SO(d)$ such that $PM^\mathsf{T} = R$ is conjugation-invariant. Then, with V as in Proposition 1,*

$$\int_{SO(d)} \mathrm{Gram}(HAV) f_P(A)\mu(\mathrm{d}A) = V^\mathsf{T} M^\mathsf{T}(I - D^2)MV,$$

where $D^2 = \text{diag}((1 - \tau_2)/(d - 1), \ldots, (1 - \tau_2)/(d - 1), \tau_2)$, and

$$\tau_2 = \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\Gamma(\frac{d-1}{2})} \int_{-1}^{1} t^2 (1 - t^2)^{(d-3)/2} \overline{f}_{v_{R^\mathsf{T}}}(t) \, dt. \tag{3.3}$$

Proof. Similar to Panaretos (2009), pages 3285–3286, one obtains

$$\int_{SO(d)} \text{Gram}(HAV) f_P(A) \mu(dA) = V^\mathsf{T} M^\mathsf{T} \cdot \{I - \mathrm{E}(v_{R^\mathsf{T}} v_{R^\mathsf{T}}^\mathsf{T})\} \cdot MV,$$

$$\mathrm{E}(v_{R^\mathsf{T}} v_{R^\mathsf{T}}^\mathsf{T}) = \left(\int_{\mathbb{S}^{d-1}} w_i w_j f_{v_{R^\mathsf{T}}}(w) \sigma(dw) \right)_{i,j}. \tag{3.4}$$

That (3.4) vanishes for $i \neq j$ follows from symmetry considerations applied in conjunction with the following standard integration formula [Faraut (2008), Proposition 9.1.2, page 189], where σ_0 is the uniform distribution on the "equator" $\mathbb{S}_0^{d-2} = \mathbb{S}^{d-1} \cap \{x : x_d = 0\}$ and g an integrable function:

$$\int_{\mathbb{S}^{d-1}} g(x) \sigma(dx)$$
$$= \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\Gamma(\frac{d-1}{2})} \int_0^\pi \left(\int_{\mathbb{S}_0^{d-2}} g((\sin\theta)u + (\cos\theta)e_d) \sigma_0(du) \right) \sin^{d-2}\theta \, d\theta. \tag{3.5}$$

That same formula applied to the case $i = j = d$, for which g is zonal, yields (3.3). Symmetry considerations also imply that all entries with $i = j < d$ coincide. Finally, $\mathrm{E}\{\text{tr}(D^2)\} = \mathrm{E}(v_{R^\mathsf{T}} v_{R^\mathsf{T}}^\mathsf{T}) = 1$, hence $\mathrm{E}(v_{R^\mathsf{T}} v_{R^\mathsf{T}}^\mathsf{T}) = D^2$, with D^2 as in the theorem. $\square$

Remark 1. In the case $d = 3$, the distributions of R and $R^\mathsf{T} = R^{-1}$ coincide, as follows readily from the axis-angle representation of $SO(3)$ in Section 2, together with the observation that a rotation by the amount $\theta \in \mathbb{R}$ around the oriented axis u is the same as a rotation by the amount $-\theta$ around $-u$. However, the distributions of R and R^T do not coincide in general for $d > 3$; see the characterisations in Said et al. (2010), Prop. 2, page 2768.

For the Cayley distribution with $d = 3$, Theorem 1, Remark 1 and (3.2) give via integration by parts

$$\int_{SO(3)} \text{Gram}\{HAV\} f_\kappa^{\text{CAY}}(A) \mu(dA)$$
$$= \text{Gram}(V) - \text{Gram}(D_\kappa^{\text{CAY}} MV),$$
$$D_\kappa^{\text{CAY}} = \text{diag}\left(\sqrt{\frac{2(\kappa + 1)}{6 + 5\kappa + \kappa^2}}, \sqrt{\frac{2(\kappa + 1)}{6 + 5\kappa + \kappa^2}}, \sqrt{\frac{2 + \kappa + \kappa^2}{6 + 5\kappa + \kappa^2}} \right).$$

We see that D_κ^{CAY} is a scalar matrix (i.e., a scalar multiple of I) not only for the case $\kappa = 0$, but also for $\kappa = 1$. This latter case is what we call *fake uniformity* of the Gram matrix functional. We may understand its genesis geometrically as follows: the nonconstant factor of (3.1) for $\kappa = 1$ is

$$(1 + \cos\theta)(1 - \cos\theta) = 1 - \cos^2\theta = \frac{1}{2}(1 - \cos 2\theta). \qquad (3.6)$$

From the formula for the transformation of the density of a real random variable X under scaling $X \mapsto cX$ with $c \in \mathbb{R}$, we conclude from (3.6) that the density of Θ with respect to Lebesgue measure on $[0, \pi]$ in the fake case ($\kappa = 1$) is obtained by halving Θ in the Haar case ($\kappa = 0$). Associate each point $u = (u_1, u_2, u_3)$ on the hemisphere $\{u \in \mathbb{S}^2 : u_3 \geq 0\}$ with its reflection $T(u) = (u_1, u_2, -u_3)$. Identifying T with its matrix with respect to the canonical basis $\{e_1, e_2, e_3\}$, we observe that $T \in O(3) \setminus SO(3)$, and that $HTV = HV$ for any landmark matrix V. Ignoring the set of points on the equator $u_3 = 0$, which has measure zero and is negligible, we may now, by suitable choice of either the upper or lower hemisphere, produce in the fake case $\kappa = 1$ the same projected configuration as for the Haar case $\kappa = 0$. Similar to Remark 1, the counter-clockwise angle α changes through reflection to $\pi - \alpha$. Also, note that the foregoing description effectively defines a coupling between the cases $\kappa = 0$ and $\kappa = 1$. Alternatively, one may base the argument on the distribution of v_R from (3.2), rather than the distribution of Θ; this reasoning is, however, not as appealing geometrically.

Remark 2. An examination of the proof of the theorem reveals that it suffices to consider the case $\ell = 1$: a single landmark vector is enough. This surprising fact, which seems to be at odds with the notion of "shape" as a configuration of several points, is true because conjugation-invariance is quite a strong symmetry property.

Remark 3. While we suspect that there are yet earlier references, Proposition 1 is a straightforward consequence of Grinberg and Rubin (2004), Lemma 2.5, page 796, in conjunction with Remark 2. The fake case, however, does not seem to have such a near-precedent.

4 Reconstruction from orthogonal views and fake uniformity

In this section, we give some insights into the consequences of fake uniformity with regard to the tomographic reconstruction problem introduced in Panaretos (2009) mentioned in Section 1. First, we state the random tomography model that he introduced, and recall some issues already known from the case of Haar distributed rotations. We again limit ourselves to the case $d = 3$.

   In the random tomography model, the unknown particle is construed as a three-dimensional compactly supported probability density $\rho(x) = \rho(x_1, x_2, x_3)$ on $\mathbb{R}^3$.

An observed image of ρ is a (discretized to a regular grid in practice) projection of ρ at a random angle, that is, it is given by the compactly supported random field

$$\check{\rho}(x_1, x_2) = \int_{-\infty}^{+\infty} \rho(U^{-1}x)\,dx_3, \qquad (4.1)$$

where $U \sim \mu$ is called the orientation of ρ; below we write $U\rho(x) = \rho(U^{-1}x)$. The stochastic Radon transform of length $N \geq 1$ is a sample $\{\check{\rho}_1, \ldots, \check{\rho}_N\}$ of N independent and identically distributed (i.i.d.) copies of $\check{\rho}$, generated by using a sample $\{U_1, \ldots, U_N\}$ of N i.i.d. copies of U.

As shown in Panaretos (2009), the level of ill-posedness inherent in the problem does not allow the recovery of ρ itself. However, recovery of $[\rho] = \{A\rho : A \in O(3)\}$, the equivalence class of ρ which identifies ρ with any rotated or reflected version $A\rho$, *is* possible.

In the context of statistical estimation, a low-dimensional parametrization of the particle ρ and an arbitrary projection $\check{\rho}$ is essential. To this end, the roughly spherical "blobs" that are evident in typical images of protein fragments led Panaretos (2009) to approximate ρ by a finite Gaussian mixture with fixed isotropic covariance matrices. Each component mean is termed a landmark; by themselves, the landmarks give a rough but useful approximation of ρ or $\check{\rho}$. The landmarks are encoded in the columns of the matrix V in Proposition 1 and Theorem 1. The reason why one focuses on $[\rho]$ rather than ρ in the reconstruction is also the reason why one aims to reconstruct $\mathrm{Gram}(V)$ rather than V. Proposition 1 and Theorem 1 provide the connection between the Gram matrices for the original three-dimensional and the projected two-dimensional landmarks.

In the context of devising statistical procedures from Proposition 1 in conjunction with (suitable versions of) the law of large numbers and central limit theorem, two issues stand out, as stated in Panaretos and Konis (2011), Sections 5.1–5.2, pages 2586–2589. We give them in reverse order.

The second issue was already noted in Panaretos (2009), page 3286, and is discussed further in Panaretos and Konis (2011): for averaging Gram matrices, one needs to associate individual landmarks across different images, even though individual labels are *a priori* not encoded in the averaged Gram matrix. The randomness in the model is not related to this point (except through issues that may arise from the possibly different number of images at different "views", where "view" is defined in the next paragraph), and hence neither is fake uniformity.

The first issue from Panaretos and Konis (2011) is whether one may use fewer projections in the law-of-large-numbers approximation of the integral on the left-hand side of the equation in Proposition 1. We repeatedly make use, as done in Panaretos and Konis (2011), of the notion of "view" of the particle. This should be carefully distinguished from the notion of orientation defined around (4.1). The notion of "view" adopted in Panaretos and Konis (2011), as well as here, is relevant only to two-dimensional images. By definition, the viewing *direction* is given by the vector $w_n \in \mathbb{S}^2$, which is unique up to sign, in the representation

$HU_n = I - w_n w_n^\mathsf{T}$, where $H(x_1, x_2, x_3) = (x_1, x_2, 0)$. (Thus we regard viewing directions as axial.) The "raw view" of the particle ρ (or its landmarks) is the image of ρ (or its landmarks) under the map HU_n. We may choose to omit the vanishing third (x_3) coordinate, leaving us with what is called a "profile" in Panaretos (2009) and Panaretos and Konis (2011). The *view* of the projected particle (or its landmarks) is obtained when the "raw view" (in $\mathbb{R}^3$) is identified with (i) the image of any rotation that leaves the projection plane of H invariant; (ii) the image obtained from reversal of the viewing direction along the x_3 axis. Thus any two images of the particle ρ comprise identical views if ρ is subjected to the action $U\rho$ defined immediately after (4.1) but now with $U \in O(3)$, rather than $SO(3)$; and views are equivalence classes with factor group $O(2)$, rather than $SO(2)$, because of view-reversal. See Panaretos and Konis (2011), Figure 10, page 2593, for an example.

As shown in Panaretos and Konis (2011), Lemma 5.1, page 2588, three orthogonal views suffice to reconstruct the Gram matrix. While, as noted in Panaretos and Konis (2011), pages 2588–2589, it is impossible to ensure that the views selected from the available imagery are indeed orthogonal, the procedure they developed fared well enough in their practical example.

In the fake case, while orientation is no longer Haar distributed, the viewing directions generated by the rotation R are still uniform. Hence, fake uniformity cannot be identified within the tomographic model—unless the modal rotation M from Theorem 1 is different enough from I to disturb a set of three approximately orthogonal views from the Haar case. Hence, fake uniformity does indeed constitute a problem from a modelling viewpoint.

Acknowledgments

References

Faraut, J. (2008). *Analysis on Lie Groups: An Introduction*. Cambridge: Cambridge University Press. MR2426516

Grinberg, E. and Rubin, B. (2004). Radon inversion on Grassmannians via Gårding–Gindikin fractional integrals. *Annals of Mathematics (2)* **159**, 783–817. MR2081440

Groemer, H. (1996). *Geometric Applications of Fourier Series and Spherical Harmonics*. Cambridge: Cambridge University Press. MR1412143

Kendall, D. (1977). The diffusion of shape. *Advances in Applied Probability* **9**, 428–430.

León, C., Massé, J.-C. and Rivest, L.-P. (2006). A statistical model for random rotations. *Journal of Multivariate Analysis* **97**, 412–430. MR2234030

Panaretos, V. (2009). On random tomography with unobservable projection angles. *The Annals of Statistics* **37**, 3272–3306. MR2549560

Panaretos, V. and Konis, K. (2011). Sparse approximations of protein structure from noisy random projections. *The Annals of Applied Statistics* **5**, 2572–2602. MR2907127

Said, S., Lageman, C., Le Bihan, N. and Manton, J. (2010). Decompounding on compact Lie groups. *IEEE Transactions on Information Theory* **56**, 2766–2777. MR2683433

Schaeben, H. (1997). A simple standard orientation density function: The hyperspherical de la Vallée Poussin kernel. *Physica Status Solidi B, Basic Research* **200**, 367–376.

Schindler, W. (1997). On the efficient simulation of a particular class of random rotations relevant to computer graphics. *Journal of Computational and Applied Mathematics* **81**, 107–114. MR1459360

Schneider, R. and Weil, W. (2008). *Stochastic and Integral Geometry*. Berlin: Springer. MR2455326

Stanfill, B., Hofmann, H. and Genschel, U. (2014). Rotations: An R package for $SO(3)$ data. *The R Journal* **6**, 68–78.

Department of Mathematics
Shantou University
Shantou Guangdong 515063
P.R. China
E-mail: christianrau080@gmail.com