

# A class of spherical and elliptical distributions with Gaussian-like limit properties.

Chris Sherlock<sup>1\*</sup> and Daniel Elton<sup>1</sup>

1. *Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK.*

*\* Correspondence should be addressed to Chris Sherlock.  
(e-mail: c.sherlock@lancaster.ac.uk).*

## Summary

We present a class of spherically symmetric random variables defined by the property that as dimension increases to infinity the mass becomes concentrated in a hyperspherical shell, the width of which is negligible compared to its radius. We provide a sufficient condition for this property in terms of the functional form of the density, and then show that the property carries through to equivalent elliptically symmetric distributions, provided that the contours are not too eccentric, in a sense which we make precise. Individual components of such distributions possess a number of appealing Gaussian-like limit properties; in particular that the limiting one-dimensional marginal distribution along any component is Gaussian.

**Keywords:** *Spherically symmetric, elliptically symmetric, Gaussian limit.*

## 1 Introduction

Any spherically symmetric random variable can be represented as a mixture of spherical “shells” with distribution function proportional to  $\mathbf{1}_{\{\|\mathbf{x}\|\geq r\}}$ . We consider a class of the spherically symmetric random variables for which as dimension  $d \rightarrow \infty$  the effective range of the mixture of “shells” becomes infinitesimal relative to a typical scale from the mixture. We then generalise this class to include a subset of the corresponding elliptically symmetric random variables. This offers a relatively rich class of random variables, the components of which are shown to possess appealing Gaussian-like limit properties.

Specifically we consider sequences of spherically symmetric random variables  $\{\mathbf{X}_d\}$  which satisfy

$$\text{either } \frac{\|\mathbf{X}_d\|}{r_d} \xrightarrow{p} 1 \quad (1)$$

$$\text{or } \frac{\|\mathbf{X}_d\|}{r_d} \xrightarrow{m.s.} 1, \quad (2)$$

for some positive sequence  $\{r_d\}$ . Here and throughout this article  $\|\cdot\|$  refers to the Euclidean norm. The set of such sequences includes, for instance, the sequence of standard  $d$ -dimensional Gaussians, for which  $\|\mathbf{X}_d\|/d^{1/2} \xrightarrow{m.s.} 1$ ; indeed the Gaussian-like limit properties of the whole class arise from this fact. More generally, we provide a sufficient condition for (2) for sequences of random variables with densities of the form

$$f_d(\mathbf{x}_d) \propto \exp[-g(\|\mathbf{x}_d\|)]. \quad (3)$$

We then consider elliptically symmetric random variables, which are obtained by a sequence of (potentially) random linear transformations of spherically symmetric random variables satisfying either (1) or (2) and show that the properties (1) and (2) are unaffected by the transformation provided that the eccentricity of the elliptically symmetric random variable is not too extreme, in a sense which we make precise. Finally we show Gaussian-like limiting behaviour for individual components of a random variable from this class, both in terms of their marginal distribution, and in terms of a their maximum.

Section 2 presents the main results, which are briefly summarised and placed in context in Section 3; proofs are provided in Appendix A.

## 2 Results

Our first result provides a class of densities and associated scaling constants that satisfy (2).

**Theorem 1** *Let  $\{\mathbf{X}_d\}$  be a sequence of spherically symmetric random variables with density given by (3). Let  $g \in C^2$  satisfy*

$$\left(r \frac{d}{dr}\right)^2 g(r) \rightarrow \infty \text{ as } r \rightarrow \infty, \quad (4)$$

and let  $r_d$  be a solution to

$$r \frac{d}{dr} g(r) = d.$$

*Then there is a sequence of solutions which satisfies  $r_d \rightarrow \infty$ , where  $r_d$  is unique for sufficiently large  $d$ . Elements of this sequence and  $\mathbf{X}_d$  together satisfy (2).*

The class of interest therefore includes the exponential power family, which has densities proportional to  $\exp(-\|\mathbf{x}\|^a)$  ( $a > 0$ ), and  $r_d = (d/a)^{1/d}$ ; indeed the class includes any density with polynomial exponents. However it does not include, for example, the log-normal density proportional to  $\|\mathbf{x}\|^{-1} \exp(-0.5(\|\mathbf{x}\| - \mu)^2)$ .

Theorem 1 requires  $g \in C^2$ . Other functional forms can also lead to the desired convergence, although not necessarily with  $r_d \rightarrow \infty$ ; for example  $g(r) = \mathbf{1}_{\{r \leq 1\}}$ .

We next show that (1) and (2) continue to hold after a linear transformation is applied to each  $\mathbf{X}_d$ , providing that the resulting sequence of elliptically symmetric random variables is not too eccentric.

**Theorem 2** *Let  $\{\mathbf{X}_d\}$  be a sequence of spherically symmetric random variables and  $\{r_d\}$  a sequence of positive constants. Further let  $\{\mathbf{T}_d\}$  be a sequence of random linear maps on  $\mathbb{R}^d$  which are independent of  $\{\mathbf{X}_d\}$ . Denote the eigenvalues of  $\mathbf{T}_d^t \mathbf{T}_d$  by  $\lambda_{d,1} \geq \lambda_{d,2} \cdots \geq \lambda_{d,d} \geq 0$ , and set  $B_d := \sum_{i=1}^d \lambda_{d,i}$ . If*

$$\frac{\lambda_{d,1}}{B_d} \xrightarrow{p} 0 \tag{5}$$

then

$$\frac{\|\mathbf{X}_d\|}{r_d} \xrightarrow{p} 1 \Rightarrow \frac{d^{1/2} \|\mathbf{T}_d \mathbf{X}_d\|}{B_d^{1/2} r_d} \xrightarrow{p} 1, \tag{6}$$

$$\frac{\|\mathbf{X}_d\|}{r_d} \xrightarrow{m.s.} 1 \Rightarrow \frac{d^{1/2} \|\mathbf{T}_d \mathbf{X}_d\|}{B_d^{1/2} r_d} \xrightarrow{m.s.} 1. \tag{7}$$

The class of elliptically symmetric random variables therefore includes, for example, densities of the form  $\exp(-(\mathbf{x}^t \mathbf{\Lambda} \mathbf{x})^a)$  ( $a > 0$ ), for symmetric  $\mathbf{\Lambda}$  for which the sum of the eigenvalues is much larger than their maximum.

Our final theorem demonstrates that even if the weaker condition (1) is satisfied by a spherically symmetric sequence, then any limiting one-dimensional marginal distribution is Gaussian; it also provides a slightly weaker result for elliptically symmetric sequences as well as a limiting bound on the maximum of all of the components.

**Theorem 3** *Let the sequence of spherically symmetric random variables  $\{\mathbf{X}_d\}$  and the sequence of positive constants  $\{r_d\}$  satisfy (1), and let the sequence of  $d$ -dimensional linear maps,  $\{\mathbf{T}_d\}$ , satisfy (5).*

1. *For any sequence of unit vectors  $\{\mathbf{e}_d\}$ , which may be random, but is independent of  $\{\mathbf{X}_d\}$ ,*

$$\frac{d^{1/2}}{r_d} \mathbf{X}_d \cdot \mathbf{e}_d \xrightarrow{D} N(0, 1).$$

2. For any sequence of random unit vectors  $\{\mathbf{e}_d\}$  which is independent of  $\{\mathbf{X}_d\}$  and  $\{\mathbf{T}_d\}$ , and the  $d^{\text{th}}$  of which is distributed uniformly on the surface of a  $d$ -sphere,

$$\frac{d}{B_d^{1/2} r_d} (\mathbf{T}_d \mathbf{X}_d) \cdot \mathbf{e}_d \xrightarrow{D} N(0, 1).$$

3. Denote the  $i^{\text{th}}$  component of  $\mathbf{X}_d$  as  $X_{d,i}$ . Then

$$\frac{d^{1/2}}{(2 \log d)^{1/2} r_d} \max_{i=1 \dots d} X_{d,i} \xrightarrow{p} 1.$$

It should be noted that the first part of Theorem 3 is not simply a standard consequence of the central limit theorem. Rather it results from the fact that the standard  $d$ -dimensional Gaussian satisfies condition (1), and hence any other sequence which satisfies (1) becomes in some sense “close” to a  $d$ -dimensional Gaussian as  $d \rightarrow \infty$ ; close enough that the marginal one-dimensional distributions start to resemble each other.

The resemblance to a standard multivariate Gaussian is sufficient for a similar deterministic limit on the maximum of all of the components (Part 3); however the well known limiting Gumbel distribution for the maximum of a set of independent Gaussians (see Section A.3) is not shared by all members of this class.

### 3 Discussion

It is well known (e.g. [1]) that any given spherically (or elliptically) symmetric random variable can be represented as a mixture of Gaussians; the marginal distribution of any given component is therefore also a mixture of Gaussians. Further, [2] show that for a sequence of independent and identically distributed components, the marginal one dimensional distribution along all but a vanishingly small fraction of random unit vectors becomes closer and closer to Gaussian as dimension  $d \rightarrow \infty$ .

In a sense we have presented an intersection of these two ideas: a class of spherical and elliptical distributions for which the limiting marginal distribution is in fact Gaussian; not a mixture. Moreover the maximum component size is bounded in proportion to  $(\log d)^{1/2}$ , in a similar manner to the maximum component size of a high dimensional Gaussian. A sufficient condition for the functional form has been provided, and this is satisfied, for example, by the exponential power distribution.

The Gaussian-like limit properties are fundamental to results in [3] and [4] where it is shown that if the proposal distribution for a random walk Metropolis algorithm is chosen from this class then some aspects of the behaviour of the algorithm can become deterministic and, in particular, that the optimal acceptance rate approaches a known fixed value as  $d \rightarrow \infty$ .

# A Proofs of results

## A.1 Proof of Theorem 1

It will be helpful to define  $R_d := \|\mathbf{X}_d\|$ , and  $U_d := \log R_d$ , and to transform the problem to that of approximating a single integral:

$$\mathbb{P}(R \in (a, b)) \propto \int_a^b dr r^{d-1} \exp[-g(r)] = \int_{\log a}^{\log b} du \exp[ud - g(e^u)]. \quad (8)$$

Here and elsewhere for clarity of exposition we sometimes omit the subscript,  $d$ .

Theorem 1 is proved in three parts.

- We first show that for  $d > d_*$  (for some  $d_* > 0$ ), the density  $\propto \exp[ud - g(e^u)]$  attains a unique maximum in  $[u_*, \infty)$  for some fixed  $u_* \in \mathbb{R}$ . We will denote the value at which this maximum occurs as  $u_d$ . The required sequence of scalings will turn out to be  $r_d = \exp(u_d)$ .
- Convexity arguments are then applied to show that

$$\frac{\|\mathbf{X}_d\|}{r_d} \xrightarrow{p} 1.$$

- It is then shown that for any fixed  $k > 0$

$$\frac{1}{r_d^k} \mathbb{E} \left[ \|\mathbf{X}_d\|^k \right] \rightarrow 1.$$

Applying this with  $k = 1$  and  $k = 2$  provides the required result.

### A.1.1 Existence of an unique maximum in $[u_*, \infty)$

Define  $\eta(u) := g(e^u)$ . Clearly  $\eta : \mathbb{R} \rightarrow \mathbb{R}$  and  $\eta \in C^2$ ; also Condition (4) is equivalent to

$$\lim_{u \rightarrow \infty} \eta''(u) = \infty. \quad (9)$$

Hence we may define

$$u_* := \inf\{u : \eta''(u') > 1 \forall u' > u\}.$$

**Lemma 1** *Subject to Condition (9),  $\exists d_*$  such that for all  $d > d_*$  there is a solution  $u_d > u_*$  to the equation  $\eta'(u) = d$  which is unique in  $[u_*, \infty)$ . Moreover  $u_d \rightarrow \infty$ .*

**Proof:** For  $u > u_*$ ,  $\eta'(u) > \eta'(u_*) + u - u_*$ . Let  $d_*$  be the first positive integer greater than  $\eta'(u_*)$  then clearly there is a solution to  $\eta'(u) = d \quad \forall d \geq d_*$ .

If there are two such solutions,  $u'$  and  $u''$  with  $u' > u'' > u_*$  then we obtain a contradiction since

$$0 = \frac{\eta'(u') - \eta'(u'')}{u' - u''} = \eta''(u''') \quad \text{for some } u''' \in [u', u''].$$

Next consider successive solutions,  $u_d$  and  $u_{d+1}$  for  $d > d_*$  and again apply the intermediate value theorem:

$$\frac{1}{u_{d+1} - u_d} = \frac{\eta'(u_{d+1}) - \eta'(u_d)}{u_{d+1} - u_d} = \eta''(u') > 0,$$

for some  $u'$ , since  $u' > u_*$ . Therefore  $u_{d+1} > u_d$  and the sequence  $\{u_d : d \geq d_*\}$  is monotonic and therefore must approach a limit. Suppose that this limit is finite,  $u_d \rightarrow c$ . Then since  $\eta'$  is continuous,  $\eta'(u_d) \rightarrow \eta'(c) < \infty$ . This contradicts the fact that  $\eta'(u_d) = d$  and hence  $u_d \rightarrow \infty$ . Since  $\eta''(u) \rightarrow \infty$  as  $u \rightarrow \infty$  the proof is complete.

### A.1.2 Convergence in probability

**Lemma 2** *Let  $\{\mathbf{X}_d\}$  be a sequence of spherically symmetric random variables with density given by (3). If  $g \in C^2$  and satisfies (4) then there is a sequence  $r_d \rightarrow \infty$  such that*

$$\frac{\|\mathbf{X}_d\|}{r_d} \xrightarrow{p} 1.$$

In proving Lemma 2 we consider the log-density (up to a constant) of  $U_d$ :

$$\psi_d(u) = ud - \eta(u). \tag{10}$$

Note that Condition 9 implies that  $\psi_d''(u) \rightarrow -\infty$  as  $u \rightarrow \infty$ , and  $\psi_d''(u) < -1 \quad \forall u > u_*$ .

We now assume  $d > d_*$  and consider the integral  $\int_{-\infty}^{\infty} du \exp[\psi_d(u)]$ . This integral must be finite for all  $d$  greater than some  $d_{**}$ , since otherwise  $\{R_d\}$  cannot be an infinite sequence of random variables. For a given  $\delta \in (0, 1)$ , the area of integration is partitioned into five separate regions:

- $R_1^{(d)} := (-\infty, u_*]$ ;
- $R_2^{(d)} := (u_*, u_d + \log(1 - \delta)]$ ;
- $R_3^{(d)} := (u_d + \log(1 - \delta), u_d]$ ;

- $R_4^{(d)} := (u_d, u_d + \log(1 + \delta)]$ ;
- and  $R_5^{(d)} := (u_d + \log(1 + \delta), \infty)$ .

It will be convenient to define the respective integrals

$$I_i^{(d)} := \int_{R_i^{(d)}} du \exp[\psi_d(u)], \quad (i = 1, \dots, 5). \quad (11)$$

Note that

$$I_3^{(d)} + I_4^{(d)} \propto \int_{(1-\delta)\exp(u_d)}^{(1+\delta)\exp(u_d)} dr f_r(r)$$

where  $f_r(r)$  is the density of  $R$ . The required convergence in probability will therefore be proven if we can show that by taking  $d$  large enough, each of  $I_1^{(d)}$ ,  $I_2^{(d)}$  and  $I_5^{(d)}$  can be made arbitrarily small compared with either  $I_3^{(d)}$  or  $I_4^{(d)}$ .

The next three propositions arise from convexity arguments and will be applied repeatedly to bound certain ratios of integrals.

**Proposition 1** *Let  $\psi : [u_*, \infty) \rightarrow \mathbb{R}$  have  $\psi''(u) < 0$ . For any  $u_0, u_1 \in [u_*, \infty)$ ,*

$$\int_{u_0}^{u_1} du e^{\psi(u)} \geq e^{\psi(u_1)} \frac{u_1 - u_0}{\psi(u_0) - \psi(u_1)} (e^{\psi(u_0) - \psi(u_1)} - 1).$$

**Proof:** Define the interval  $K := [u_0, u_1]$  if  $u_1 > u_0$ , and  $[u_1, u_0]$  otherwise. By the concavity of  $\psi$ ,

$$\psi(u) \geq \psi(u_1) + \frac{\psi(u_1) - \psi(u_0)}{u_1 - u_0} (u - u_1) \quad \forall u \in K.$$

Hence

$$\int_{u_0}^{u_1} du e^{\psi(u)} \geq e^{\psi(u_1)} \int_{u_0}^{u_1} du \exp \left[ \frac{\psi(u_1) - \psi(u_0)}{u_1 - u_0} (u - u_1) \right].$$

The result follows on evaluating the right hand integral.

**Proposition 2** *Let  $\psi : [u_*, \infty) \rightarrow \mathbb{R}$  have  $\psi''(u) \leq 0$ . For any  $u_0, u_1 \in [u_*, \infty)$  with  $u_1 > u_0$  and  $\psi(u_0) > \psi(u_1)$ ,*

$$\int_{u_1}^{\infty} du e^{\psi(u)} \leq e^{\psi(u_1)} \frac{u_1 - u_0}{\psi(u_0) - \psi(u_1)}.$$

**Proof:** By the concavity of  $\psi$ ,

$$\psi(u) \leq \psi(u_1) + \psi'(u_1)(u - u_1) \leq \psi(u_1) + \frac{\psi(u_1) - \psi(u_0)}{u_1 - u_0} (u - u_1) \quad \forall u \in [u_1, \infty).$$

Hence

$$\int_{u_1}^{\infty} du e^{\psi(u)} \leq e^{\psi(u_1)} \int_{u_1}^{\infty} du \exp \left[ \frac{\psi(u_1) - \psi(u_0)}{u_1 - u_0} (u - u_1) \right]$$

Since  $(\psi(u_1) - \psi(u_0))/(u_1 - u_0)$  is negative, the result follows on evaluating the right hand integral.

The proof for the following is almost identical to that of Proposition 2 and is therefore omitted.

**Proposition 3** *Let  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  have  $\psi''(u) \leq 0$ . For any  $u_0, u_1 \in \mathbb{R}$  with  $u_1 < u_0$  and  $\psi(u_0) > \psi(u_1)$ ,*

$$\int_{-\infty}^{u_1} du e^{\psi(u)} \leq e^{\psi(u_1)} \frac{u_0 - u_1}{\psi(u_0) - \psi(u_1)}.$$

**Corollary 1**

$$\frac{I_5}{I_4 + I_5} \leq \exp [\psi(u_d + \log(1 + \delta)) - \psi(u_d)].$$

**Proof:** Set  $u_0 = u_d$  and  $u_1 = u_d + \log(1 + \delta)$  in Propositions 1 and 2 to obtain

$$\frac{I_4}{I_5} \geq \exp [\psi(u_d) - \psi(u_d + \log(1 + \delta))] - 1.$$

But

$$\frac{I_5}{I_4 + I_5} = \frac{1}{1 + I_4/I_5}$$

and so the result follows.

**Corollary 2**

$$\frac{I_2}{I_2 + I_3} \leq \exp [\psi(u_d + \log(1 - \delta)) - \psi(u_d)].$$

**Proof:** Define

$$\psi_c(u) := \begin{cases} \psi(u) & u > u_* \\ \psi(u_*) - (u_* - u)\psi'(u_*) & u < u_*. \end{cases}$$

By definition,  $\psi_c(u) : \mathbb{R} \rightarrow \mathbb{R}$ , and  $\psi''(u) \leq 0 \forall u \in \mathbb{R}$ . Let

$$I_{1c} := \int_{R_1} du \exp [\psi_c(u)],$$

and note that  $\psi'(u_*) > 0$  since  $\psi''(u) \leq 0 \forall u \geq u_*$ , and  $\psi'(u_d) = 0$  with  $u_d > u_*$ . Hence  $\int_{R_1} du \exp[\psi_c(u)]$  exists.

Set  $u_0 = u_d$  and  $u_1 = u_d + \log(1 - \delta)$  in Propositions 1 and 3 to obtain

$$\frac{I_3}{I_{1c} + I_2} \geq \exp[\psi(u_d) - \psi(u_d + \log(1 - \delta))] - 1.$$

But

$$\frac{I_2}{I_2 + I_3} \leq \frac{I_{1c} + I_2}{I_{1c} + I_2 + I_3} = \frac{1}{1 + I_3/(I_{1c} + I_2)}$$

and so the result follows.

We now consider  $I_1^{(d)}$  and use the fact that  $\int_{R_1} du \exp[\psi_d(u)]$  must exist for all  $d > d_{**}$  (for some  $d_{**} > 0$ ) for  $\{R_d\}$  to be an infinite sequence of random variables. Also note that  $\psi_d(u) - \psi_k(u) = (d - k)u$ , which is an increasing function for  $d > k$ .

**Corollary 3** *If  $I_1^{(k)} < \infty$  for some  $k > 0$  and if for all  $d > k$ ,  $\psi_d(u) - \psi_k(u)$  is an increasing function of  $u$  then*

$$\frac{I_1^{(d)}}{I_2^{(d)} + I_3^{(d)}} \leq \frac{e^{-\psi_k(u_*)} I_1^{(k)}}{u_d - u_*}.$$

**Proof:** By the monotonicity of  $\psi_d - \psi_k$ ,

$$I_1^{(d)} = \int_{-\infty}^{u_*} du e^{\psi_d(u) - \psi_k(u)} e^{\psi_k(u)} \leq e^{\psi_d(u_*) - \psi_k(u_*)} I_1^{(k)}.$$

By Proposition 1 with  $u_0 = u_d$  and  $u_1 = u_*$

$$I_2^{(d)} + I_3^{(d)} \geq e^{\psi_d(u_*)} \frac{u_d - u_*}{\psi_d(u_d) - \psi_d(u_*)} (e^{\psi_d(u_d) - \psi_d(u_*)} - 1) \geq e^{\psi_d(u_*)} (u_d - u_*),$$

where the last statement follows since for  $x > 0$ ,  $e^x > 1 + x$ .

The result follows from combining the two inequalities.

We next combine Corollaries 1, 2, and 3 to prove the sufficient condition for the required convergence in probability. We show that if Condition 9 is satisfied, then

$$\frac{I_1^{(d)} + I_2^{(d)} + I_5^{(d)}}{I_1^{(d)} + I_2^{(d)} + I_3^{(d)} + I_4^{(d)} + I_5^{(d)}} \rightarrow 0 \text{ as } d \rightarrow \infty. \quad (12)$$

By Lemma 1  $u_d \rightarrow \infty$  as  $d \rightarrow \infty$ , and so from Corollary 3

$$\frac{I_1^{(d)}}{I_2^{(d)} + I_3^{(d)}} \rightarrow 0.$$

Since  $u_d \rightarrow \infty$ , given some  $\delta \in (0, \delta_0)$  and any  $M > 0$ , we may choose a  $d_0$  such that for all  $d > d_0$  and all  $\delta_* \in (0, \delta)$ ,

$$(\log(1 + \delta))^2 \eta''(u_d + \log(1 + \delta_*)) \geq M.$$

Taylor expand  $\psi_d$  about  $u_d$ , recalling that  $\psi'_d(u_d) = 0$  and  $\psi''_d(u) = -\eta''(u)$ :

$$\psi_d(u_d) - \psi_d(u_d + \log(1 + \delta)) = \frac{1}{2} (\log(1 + \delta))^2 \eta''(u_d + \log(1 + \delta_*)) \geq \frac{1}{2} M,$$

for some  $\delta^* \in (0, \delta)$ . From Corollary 1 we therefore see that

$$\frac{I_5^{(d)}}{I_4^{(d)} + I_5^{(d)}} \leq e^{-\frac{1}{2}M}.$$

Similarly, from Corollary 2

$$\frac{I_2^{(d)}}{I_2^{(d)} + I_3^{(d)}} \leq e^{-\frac{1}{2}M}.$$

But

$$\frac{I_1^{(d)} + I_2^{(d)} + I_5^{(d)}}{I_1^{(d)} + I_2^{(d)} + I_3^{(d)} + I_4^{(d)} + I_5^{(d)}} \leq \frac{I_1^{(d)}}{I_2^{(d)} + I_3^{(d)}} + \frac{I_2^{(d)}}{I_2^{(d)} + I_3^{(d)}} + \frac{I_5^{(d)}}{I_4^{(d)} + I_5^{(d)}}$$

and each of the terms on the right hand side can be made as small as desired by taking  $d$  large enough.

### A.1.3 Convergence of $k$ -th moment

**Proposition 4** *Let  $r_d$  be the (eventually) unique solution to the equation*

$$r g'(r) = d - 1,$$

*If  $g(r)$  satisfies (4) then for any fixed  $k > 0$*

$$\lim_{d \rightarrow \infty} \frac{r_d}{r_{d+k}} = 1.$$

**Proof:** Without loss of generality assume that  $r_{d+k} > r_d$ . Hence by the Intermediate Value Theorem, there exists a value  $r_* \in [r_d, r_{d+k}]$  such that

$$\frac{k r_{d+k}}{r_{d+k} - r_d} > \frac{r_* k}{r_{d+k} - r_d} = r_* \frac{r_{d+k} g'(r_{d+k}) - r_d g'(r_d)}{r_{d+k} - r_d} = r_* \frac{d}{dr} (r g'(r)) \Big|_{r_*} \rightarrow \infty.$$

Thus

$$\frac{r_{d+k} - r_d}{r_{d+k}} \rightarrow 0,$$

and the result follows.

**Lemma 3** For fixed  $k > 0$ ,

$$\frac{1}{r_d^k} \mathbb{E}_d [R^k] \rightarrow 1.$$

**Proof:** Set

$$I_1 := \int_0^\infty dr r^{d-1} \exp[-g(r)] \quad \text{and} \quad I_2 := \int_0^\infty dr r^{d-1+k} \exp[-g(r)].$$

If  $g(r)$  satisfies (4) then so does  $g(r) - k \log(r)$ . Therefore, from Lemma 2, given  $\epsilon > 0$  and  $\delta > 0$  there is a  $d_1$  such that for all  $d > d_1$ ,

$$(1 - \epsilon)I_2 < \int_{r_{d+k}(1-2\delta)}^{r_{d+k}(1+2\delta)} dr r^{d-1+k} \exp[-g(r)] < I_2.$$

Furthermore, by Proposition 4, there is a  $d_2$  such that for all  $d > d_2$ ,

$$r_d(1 - \delta) < r_{d+k} < (1 + \delta)r_d.$$

Therefore, since the integrand is positive, for all  $d > \max(d_1, d_2)$ ,

$$\begin{aligned} (1 - \epsilon)I_2 &< \int_{r_d(1-2\delta)(1-\delta)}^{r_d(1+2\delta)(1+\delta)} dr r^{d-1+k} \exp[-g(r)] \\ &\leq r_d^k (1 + 2\delta)^k (1 + \delta)^k \int_{r_d(1-2\delta)(1-\delta)}^{r_d(1+2\delta)(1+\delta)} dr r^{d-1} \exp[-g(r)]. \end{aligned}$$

Similarly

$$\begin{aligned} I_2 &> \int_{r_d(1-2\delta)(1+\delta)}^{r_d(1+2\delta)(1-\delta)} dr r^{d-1+k} \exp[-g(r)] \\ &\geq r_d^k (1 - 2\delta)^k (1 + \delta)^k \int_{r_d(1-2\delta)(1+\delta)}^{r_d(1+2\delta)(1-\delta)} dr r^{d-1} \exp[-g(r)]. \end{aligned}$$

Applying Lemma 2 again, there is a  $d_3$  such that for all  $d > d_3$ ,

$$(1 - \epsilon)I_1 < \int_{r_d(1-2\delta)(1+\delta)}^{r_d(1+2\delta)(1-\delta)} dr r^{d-1} \exp[-g(r)] < I_1.$$

Therefore, for all  $d > \max(d_1, d_2, d_3)$ ,

$$\begin{aligned} (1 - \epsilon)I_2 &< r_d^k (1 + 2\delta)^k (1 + \delta)^k I_1 \\ I_2 &> r_d^k (1 - 2\delta)^k (1 + \delta)^k (1 - \epsilon)I_1. \end{aligned}$$

Hence

$$(1 - 2\delta)^k (1 + \delta)^k (1 - \epsilon) < \frac{1}{r_d^k} \frac{I_2}{I_1} < (1 + 2\delta)^k (1 + \delta)^k (1 - \epsilon)^{-1}.$$

The result follows since  $\delta$  and  $\epsilon$  can be made arbitrarily small.

## A.2 Proof of Theorem 2

Any spherically symmetric random variable can be decomposed into a uniform angular component and a radial distribution. We may therefore create an invertible map from any  $d$  dimensional spherically symmetric random variable  $\mathbf{V}$  with a continuous radial distribution function to a standard  $d$  dimensional Gaussian,  $\mathbf{Z}$ . We will apply the following map: set

$$\|\mathbf{Z}\| = F_{\|\mathbf{Z}\|}^{-1} (F_{\|\mathbf{V}\|} (\|\mathbf{V}\|)), \quad (13)$$

where where  $F_{\|\mathbf{V}\|}(\cdot)$  and  $F_{\|\mathbf{Z}\|}(\cdot)$  are the distribution functions of  $\|\mathbf{V}\|$  and  $\|\mathbf{Z}\|$  respectively, and then fix

$$\mathbf{Z} = \frac{\|\mathbf{Z}\|}{\|\mathbf{V}\|} \mathbf{V}. \quad (14)$$

This mapping is key both to the proofs of both Theorem 2 and of Theorem 3. To simplify the exposition in both this section and Section A.3 we define

$$\mathbf{V}_d := \frac{1}{r_d} \mathbf{X}_d. \quad (15)$$

The following is therefore equivalent to (6).

**Lemma 4** Define  $\{\mathbf{V}_d\}$ ,  $\{\mathbf{T}_d\}$ ,  $\{\lambda_{d,i}\}$  and  $\{B_d\}$  as in (15) and the statement of Theorem 2. If (5) holds and  $\|\mathbf{V}_d\| \xrightarrow{p} 1$  then

$$\frac{d}{B_d} \mathbf{V}_d^t \mathbf{T}_d^t \mathbf{T}_d \mathbf{V}_d \xrightarrow{p} 1.$$

**Proof:** For some  $\delta > 0$ , let

$$A_d := \left\{ \mathbf{T}_d : \frac{\lambda_{d,1}}{\sum_{i=1}^d \lambda_{d,i}} < \delta^3 \right\}.$$

For now fix  $d$  and  $\mathbf{T}_d \in A_d$ , and suppress the subscript  $d$ . Denote the spectral decomposition of  $\mathbf{T}^t \mathbf{T}$  as  $\mathbf{L}^t \mathbf{\Lambda} \mathbf{L}$ , where  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_d)$ . We will initially consider the Gaussian  $\mathbf{Z}$  and define  $\mathbf{Z}^* = \mathbf{L} \mathbf{Z}$ ; since  $\mathbf{L}$  is orthonormal it follows that  $\mathbf{Z}^* \sim N(\mathbf{0}, \mathbf{I}_d)$ .

Define

$$W = \frac{\mathbf{Z}^t \mathbf{T}^t \mathbf{T} \mathbf{Z}}{B} = \frac{\mathbf{Z}^{*t} \mathbf{\Lambda} \mathbf{Z}^*}{B}.$$

Then for fixed  $d$ ,

$$\begin{aligned}\mathbb{E}_{\mathbf{Z}} [W] &= \frac{1}{B} \mathbb{E}_{\mathbf{Z}} \left[ \sum_{i=1}^d \lambda_i Z_i^{*2} \right] = 1 \\ \text{Var}_{\mathbf{Z}} [W] &= \frac{1}{B^2} \text{Var}_{\mathbf{Z}} \left[ \sum_{i=1}^d \lambda_i Z_i^{*2} \right] \\ &= 2 \frac{\sum_{i=1}^d \lambda_i^2}{\left( \sum_{i=1}^d \lambda_i \right)^2} \leq 2 \frac{\lambda_1}{\sum_{i=1}^d \lambda_i} < 2\delta^3.\end{aligned}$$

Chebyshev's inequality gives

$$\begin{aligned}\mathbb{P}_{\mathbf{Z}} (|W - \mathbb{E}_{\mathbf{Z}} [W]| > \delta) &< \frac{\text{Var}_{\mathbf{Z}} [W]}{\delta^2} \\ \text{i.e. } \mathbb{P}_{\mathbf{Z}} (|W - 1| > \delta) &< 2\delta.\end{aligned}$$

By (5) there is a  $d_0$  such that for all  $d > d_0$ ,  $\mathbb{P}(\mathbf{T}_d \notin A_d) < \delta$ . Thus for all  $d > d_0$

$$\mathbb{P}_{\mathbf{T}, \mathbf{Z}} (|W - 1| > \delta) \leq \mathbb{P} (|W - 1| > \delta | \mathbf{T} \in A) + \mathbb{P} (\mathbf{T} \notin A) < 3\delta.$$

Hence  $W_d \xrightarrow{p} 1$ . Now

$$\frac{d}{B_d} \mathbf{V}_d^t \mathbf{T}_d^t \mathbf{T}_d \mathbf{V}_d = \frac{d}{\|\mathbf{Z}_d\|^2} \|\mathbf{V}_d\|^2 W_d,$$

and since each of the three terms converge in probability to 1, so does the product.

We now turn to the proof of convergence in mean square, and first show an equivalence of the expected second moments of the norms.

**Proposition 5** *For  $\{\mathbf{V}_d\}$ ,  $\{\mathbf{T}_d\}$ ,  $\{\lambda_{d,i}\}$  and  $\{B_d\}$  be as defined in (15) and the statement of Theorem 2,*

$$\mathbb{E} [\|\mathbf{V}_d\|^2] = \frac{d}{B_d} \mathbb{E} [\|\mathbf{T}_d \mathbf{V}_d\|^2].$$

**Proof:** For clarity of exposition we suppress the subscript  $d$ . Since  $\mathbf{V}$  is spherically symmetric we may without loss of generality consider it with axes along the principle components of  $\mathbf{T}$ . Then

$$\mathbb{E} [\|\mathbf{T}\mathbf{V}\|^2] = \mathbb{E} \left[ \sum_{i=1}^d \lambda_i^2 V_i^2 \right] = \sum_{i=1}^d \lambda_i^2 \mathbb{E} [V_i^2].$$

But, again,  $\mathbf{V}$  is spherically symmetric so this is

$$\sum_{i=1}^d \lambda_i^2 \mathbb{E} [V_1^2] = \frac{1}{d} \sum_{i=1}^d \lambda_i^2 \sum_{j=1}^d \mathbb{E} [V_j^2] = \frac{B_d}{d} \mathbb{E} [\|\mathbf{V}\|^2].$$

Turning now to convergence in mean square itself, note that by Proposition 5,

$$\mathbb{E} \left[ \left( \left\| \frac{d^{1/2}}{B_d} \mathbf{T}_d \mathbf{V}_d \right\| - 1 \right)^2 \right] - \mathbb{E} [(\|\mathbf{V}_d\| - 1)^2] = -2 \left( \mathbb{E} \left[ \left\| \frac{d^{1/2}}{B_d} \mathbf{T}_d \mathbf{V}_d \right\| \right] - \mathbb{E} [\|\mathbf{V}_d\|] \right).$$

But (2) implies that  $\|\mathbf{V}_d\| \xrightarrow{m.s.} 1$ , and hence it is sufficient to show that

$$\mathbb{E} \left[ \left\| \frac{d^{1/2}}{B_d} \mathbf{T}_d \mathbf{V}_d \right\| \right] - \mathbb{E} [\|\mathbf{V}_d\|] \rightarrow 0. \quad (16)$$

Now by Lemma 4 and Proposition 5,

$$\left\| \frac{d^{1/2}}{B_d} \mathbf{T}_d \mathbf{V}_d \right\| \xrightarrow{p} 1 \quad \text{and} \quad \mathbb{E} \left[ \left\| \frac{d^{1/2}}{B_d} \mathbf{T}_d \mathbf{V}_d \right\|^2 \right] = \mathbb{E} [\|\mathbf{V}_d\|^2] \rightarrow 1.$$

We now require Scheffe's Lemma, which states that for any sequence of random variables  $\{Y_d\}$ , if  $\mathbb{E}[Y_d^2] \rightarrow 1$  and  $Y_d \xrightarrow{p} 1$  then  $\mathbb{E}[Y_d] \rightarrow 1$ . Hence  $\mathbb{E} \left[ \left\| \frac{d^{1/2}}{B_d} \mathbf{T}_d \mathbf{V}_d \right\| \right] \rightarrow 1$ . Now (2) also implies that  $\mathbb{E} [\|\mathbf{V}_d\|] \rightarrow 1$ , and hence (16) is satisfied.

### A.3 Proof of Theorem 3

Throughout this section we define  $\mathbf{Z}_d$  and  $\mathbf{V}_d$  as in Section A.2. We first prove Part 1.

Given  $\delta > 0$ , it will be convenient to define the following event:

$$A_d := \left| \frac{\|\mathbf{Z}_d\|}{d^{1/2}} \frac{1}{\|\mathbf{V}_d\|} - 1 \right| < \delta. \quad (17)$$

Now for  $\mathbf{e}_d$  independent of  $\mathbf{V}_d$  (and  $\mathbf{Z}_d$ ),

$$\mathbb{P}(d^{1/2} \mathbf{V}_d \cdot \mathbf{e}_d \leq a) = \mathbb{P} \left( \mathbf{Z}_d \cdot \mathbf{e}_d \leq a \frac{\|\mathbf{Z}_d\|}{d^{1/2}} \frac{1}{\|\mathbf{V}_d\|} \right) = \Phi \left( a \frac{\|\mathbf{Z}_d\|}{d^{1/2}} \frac{1}{\|\mathbf{V}_d\|} \right),$$

and so

$$\Phi(a(1 - \delta)) < \mathbb{P}(d^{1/2} \mathbf{V}_d \cdot \mathbf{e}_d \leq a \mid A^c) < \Phi(a(1 + \delta)).$$

For any event  $E$ ,

$$|\mathbb{P}(E) - \mathbb{P}(E|A)| = \mathbb{P}(A^c) |\mathbb{P}(E|A^c) - \mathbb{P}(E|A)| \leq \mathbb{P}(A^c)$$

and in particular,

$$|\mathbb{P}(d^{1/2} \mathbf{V}_d \cdot \mathbf{e}_d \leq a) - \mathbb{P}(d^{1/2} \mathbf{V}_d \cdot \mathbf{e}_d \leq a \mid A)| \leq \mathbb{P}(A^c).$$

Given  $\epsilon > 0$ , by (1) we may define  $d_0$  such that for all  $d > d_0$ ,  $\mathbb{P}(A_d^c) < \epsilon$ . Thus for all  $d > d_0$

$$\Phi(a(1 - \delta)) - \epsilon < \mathbb{P}_{\mathbf{V}_d}(d^{1/2} \mathbf{V}_d \cdot \mathbf{e}_d \leq a) < \Phi(a(1 + \delta)) + \epsilon.$$

By taking  $d$  large enough we can make  $\delta$  and  $\epsilon$  as small as desired. Moreover, since  $\Phi(\cdot)$  is bounded and monotonic  $\exists \delta^* > 0$  such that  $|\Phi(a(1 + \delta)) - \Phi(a)| < \epsilon \forall \delta$  with  $|\delta| < \delta^*$ , and hence

$$\lim_{d \rightarrow \infty} \mathbb{P} \left( d^{1/2} \mathbf{V}_d \cdot \mathbf{e}_d \leq a \right) = \Phi(a). \quad (18)$$

To prove Part 2, first note that whereas  $\mathbf{Z}_d \cdot \mathbf{e}_d \sim N(0, 1)$ ,  $(\mathbf{T}_d \mathbf{Z}_d) \cdot \mathbf{e}_d \sim N(0, \|\mathbf{T}_d \mathbf{e}_d\|^2)$ , and so

$$\begin{aligned} \mathbb{P}_{\mathbf{V}_d} \left( \frac{d}{B_d^{1/2}} (\mathbf{T}_d \mathbf{V}_d) \cdot \mathbf{e}_d \leq a \right) &= \mathbb{P}_{\mathbf{Z}_d} \left( \frac{1}{\|\mathbf{T}_d \mathbf{e}_d\|} (\mathbf{T}_d \mathbf{Z}_d) \cdot \mathbf{e}_d \leq a \frac{B_d^{1/2}}{d^{1/2} \|\mathbf{T}_d \mathbf{e}_d\|} \frac{\|\mathbf{Z}_d\|}{d^{1/2}} \frac{1}{\|\mathbf{V}_d\|} \right) \\ &= \Phi \left( a \frac{B_d^{1/2}}{d^{1/2} \|\mathbf{T}_d \mathbf{e}_d\|} \frac{\|\mathbf{Z}_d\|}{d^{1/2}} \frac{1}{\|\mathbf{V}_d\|} \right). \end{aligned}$$

But a unit vector  $\mathbf{e}_d$  chosen uniformly at random can be written as  $\mathbf{Z}_d^* / \|\mathbf{Z}_d^*\|$  for some standard  $d$ -dimensional Gaussian  $\mathbf{Z}_d^*$ . Hence by Theorem 2

$$\frac{d^{1/2}}{B_d^{1/2}} \|\mathbf{T}_d \mathbf{e}_d\| = \frac{d^{1/2}}{\|\mathbf{Z}_d^*\|} \frac{\|\mathbf{T}_d \mathbf{Z}_d^*\|}{B_d^{1/2}} \xrightarrow{p} 1.$$

We now define the event

$$A_d^* := \left| \frac{B_d^{1/2}}{d^{1/2} \|\mathbf{T}_d \mathbf{e}_d\|} \frac{\|\mathbf{Z}_d\|}{d^{1/2}} \frac{1}{\|\mathbf{V}_d\|} - 1 \right| < \delta,$$

and the proof follows as for Part 1.

In proving Part 3 we require the following standard result (e.g. [5], Theorem 1.5.3). Set

$$a_d := (2 \log d)^{-\frac{1}{2}} \quad \text{and} \quad b_d := (2 \log d)^{\frac{1}{2}} - \frac{1}{2} a_d [\log \log d + \log(4\pi)].$$

Also let  $G(\cdot)$  be the distribution function of a Gumbel random variable, and let  $Z_1, \dots, Z_d$  be independent and identically distributed  $N(0, 1)$  random variables. Then

$$G_d(c) := \mathbb{P} \left( \frac{1}{a_d} \left( \max_{i=1 \dots d} Z_i - b_d \right) \leq c \right) \rightarrow G(c).$$

Replacing  $c$  with  $c_d := [\log \log d + \log(4\pi)] / 2 = ((2 \log d)^{1/2} - b_d) / a_d$ , or with  $c_d^* := -\alpha(2 \log d)^{1/2} + [\log \log d + \log(4\pi)] / 2 = ((1 - \alpha)(2 \log d)^{1/2} - b_d) / a_d$  ( $\alpha > 0$ ) gives

$$\begin{aligned} \mathbb{P} \left( \max_{i=1 \dots d} Z_i \leq (2 \log d)^{1/2} \right) &\sim G(c_d) \rightarrow 1 \\ \mathbb{P} \left( \max_{i=1 \dots d} Z_i \leq (1 - \alpha)(2 \log d)^{1/2} \right) &\sim G(c_d^*) \rightarrow 0. \end{aligned}$$

Choose  $\delta$  in (17) small enough that  $(1 - \delta)(2 + \epsilon)^{1/2} > 2^{1/2}$ . Then

$$\begin{aligned}
& \mathbb{P} \left( d^{1/2} \max_{i=1\dots d} V_{d,i} \leq ((2 + \epsilon) \log d)^{1/2} | A_d \right) \\
&= \mathbb{P} \left( \max_{i=1\dots d} Z_{d,i} \leq \frac{\|\mathbf{Z}_d\|}{d^{1/2}} \frac{1}{\|\mathbf{V}_d\|} ((2 + \epsilon) \log d)^{1/2} \mid A_d \right) \\
&> \mathbb{P} \left( \max_{i=1\dots d} Z_{d,i} \leq (1 - \delta)((2 + \epsilon) \log d)^{1/2} \mid A_d \right) \\
&> \mathbb{P} \left( \max_{i=1\dots d} Z_{d,i} \leq (2 \log d)^{1/2} \mid A_d \right) \\
&> \mathbb{P} \left( \max_{i=1\dots d} Z_{d,i} \leq (2 \log d)^{1/2} \right) - \mathbb{P}(A_d^c).
\end{aligned}$$

Similarly by choosing  $\delta$  in (17) small enough that  $(1 + \delta)(2 - \epsilon)^{1/2} > 2^{1/2}(1 - \alpha)$  for some small  $\alpha > 0$ ,

$$\mathbb{P} \left( d^{1/2} \max_{i=1\dots d} V_{d,i} \geq ((2 - \epsilon) \log d)^{1/2} | A_d \right) > \mathbb{P} \left( \max_{i=1\dots d} Z_{d,i} \geq (1 - \alpha)(2 \log d)^{1/2} \right) - \mathbb{P}(A_d^c).$$

In each case the first term tends to 1 and  $\mathbb{P}(A_d^c) \rightarrow 0$ , proving the desired result.

## References

- [1] Kai Tai Fang, Samuel Kotz, and Kai Wang Ng. *Symmetric multivariate and related distributions*, volume 36 of *Monographs on Statistics and Applied Probability*. Chapman and Hall Ltd., London, 1990.
- [2] Persi Diaconis and David Freedman. Asymptotics of graphical projection pursuit. *Ann. Statist.*, 12(3):793–815, 1984.
- [3] Chris Sherlock and Gareth Roberts. Optimal scaling of the random walk Metropolis on elliptically symmetric unimodal targets. *Bernoulli*, 15(3):774–798, 2009.
- [4] Chris Sherlock. Optimal scaling of the random walk Metropolis: general criteria for the 0.234 acceptance rule. 2010. Submitted.
- [5] M. R. Leadbetter, Georg Lindgren, and Holger Rootzén. *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York, 1983.