

## INFERENCE IN STATISTICAL SHAPE THEORY: ELLIPTICAL CONFIGURATION DENSITIES

FRANCISCO J. CARO-LOPERA

*Department of Basic Mathematics, Centro de Investigación en Matemáticas  
Callejón de Jalisco s/n, Mineral de Valenciana, 36240 Guanajuato, Guanajuato, México*

*Email: fjcaro@cimat.mx*

JOSÉ A. DÍAZ-GARCÍA

*Universidad Autónoma Agraria Antonio Narro, Department of Statistics and Computation  
25315 Buenavista, Saltillo, Coahuila, México*

*Email: jadiaz@uaaan.mx*

GRACIELA GONZÁLEZ-FARÍAS

*Department of Statistics, Centro de Investigación en Matemáticas  
Callejón de Jalisco s/n, Mineral de Valenciana, 36240 Guanajuato, Guanajuato, México*

*Email: farias@cimat.mx*

### SUMMARY

The inference procedure for any elliptical configuration density is set out in this work in terms of published efficient algorithms involving infinite confluent hypergeometric type series of zonal polynomials. The polynomial configuration density study is proposed and then applied in a subfamily of the Kotz configuration densities, including the normal distribution; the inference procedure is then based on polynomial densities, which can be computed easily. Finally, the polynomial distributions are applied to the type of experiment readily available in other publications on shape literature.

*Keywords and phrases:* Inference in statistical shape theory, elliptical configuration densities, zonal polynomials, Kotz configuration density.

*AMS Classification:* 62H10, 62E15, 15A09, 15A52.

## 1 Introduction

Statistical shape theory based on the Euclidean transformation has been studied extensively, see Dryden and Mardia (1998) and the references therein. Assuming a isotropic matrix variate Gaussian distribution, Goodall and Mardia (1993) (corrected by Díaz-García *et al.* (2003) and revised again by Caro-Lopera *et al.* (2009)) proposed a shape density based on affine transformations, which they termed configuration density. This constitute a open distributional problem for elliptical generalisations and inference based on exact distributions.

Recently, Caro-Lopera *et al.* (2009), derived the noncentral configuration density under an elliptical model and, by using partition theory, a number of explicit configuration densities were obtained; i.e. configuration densities associated with the matrix variate symmetric Kotz type distributions (including the matrix variate normal distribution), the matrix variate Pearson type VII distributions (including the matrix variate  $t$  and Cauchy distributions), the matrix variate symmetric Bessel distribution (including the matrix variate Laplace distribution) and the matrix variate symmetric logistic distribution. The configuration density of any elliptical model was set in terms of zonal polynomials, which can now be efficiently following Koev and Edelman (2006), and in consequence, the inference problem can be studied and solved with the exact densities instead of the usual constraints and asymptotic distributions, and approximations of the statistical shape theory now function (see Goodall and Mardia (1993), Dryden and Mardia (1998) and the references therein). The general procedure becomes very clear and the underlying problem, that of the programming problem, is simply a question of time.

Thus, two perspectives can be explored; first, the inference based on exact distributions and second, their applications in shape theory.

The general procedure for performing inference of any elliptical model is proposed and set out in such a manner that the published efficient numerical algorithms for to compute a confluent hypergeometric function of a matrix argument in terms of zonal polynomials, can be used; this is outlined in section 2.

Moreover, a further simplification of the closed computational problem is also proposed, namely the study of polynomial configuration densities (Section 3); a subfamily of these densities is derived and as a simple example of their use, exact inferences for testing configuration location differences in certain applied problems are provided in Section 4. These applications are in the fields of biology (mouse vertebrae, gorilla skulls, girl and boy craniofacial studies), medicine (brain MR scans of schizophrenic patients) and image analysis (postcode recognition).

## 2 Inference for Elliptical Configuration Models

First we recall the basic definitions of elliptical distributions and configurations (see Gupta and Varga (1993) and Goodall and Mardia (1993), respectively).

We say that  $\mathbf{X} : N \times K$  has a matrix variate elliptically contoured distribution if its density with respect to the Lebesgue measure is given by:

$$f_{\mathbf{X}}(\mathbf{X}) = \frac{1}{|\boldsymbol{\Sigma}|^{K/2} |\boldsymbol{\Theta}|^{N/2}} h \left\{ \text{tr} [(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \boldsymbol{\Theta}^{-1}] \right\},$$

where  $\boldsymbol{\mu} : N \times K$ ,  $\boldsymbol{\Sigma} : N \times N$ ,  $\boldsymbol{\Theta} : K \times K$ ,  $\boldsymbol{\Sigma}$  positive definite ( $\boldsymbol{\Sigma} > \mathbf{0}$ ),  $\boldsymbol{\Theta} > \mathbf{0}$ . The function  $h : \Re \rightarrow [0, \infty)$  is termed the generator function, and is such that  $\int_0^\infty u^{NK-1} h(u^2) du < \infty$ . Such a distribution is denoted by  $\mathbf{X} \sim \mathcal{E}_{N \times K}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Theta}, h)$ .

**Definition 2.1.** Two figures  $\mathbf{X} : N \times K$  and  $\mathbf{X}_1 : N \times K$  have the *same configuration*, or *affine shape*, if  $\mathbf{X}_1 = \mathbf{X}\mathbf{E} + \mathbf{1}_N \mathbf{e}'$ , for some translation  $\mathbf{e} : K \times 1$  and a nonsingular  $\mathbf{E} : K \times K$ .

The configuration coordinates are constructed in two steps, summarised in the expression

$$\mathbf{L}\mathbf{X} = \mathbf{Y} = \mathbf{U}\mathbf{E}. \quad (2.1)$$

The matrix  $\mathbf{U} : N - 1 \times K$  contains configuration coordinates of  $\mathbf{X}$ . Let  $\mathbf{Y}_1 : K \times K$  be nonsingular and  $\mathbf{Y}_2 : q = N - K - 1 \geq 1 \times K$ , such that  $\mathbf{Y} = (\mathbf{Y}_1' | \mathbf{Y}_2)'$ . Define also  $\mathbf{U} = (\mathbf{I} | \mathbf{V})'$ , then  $\mathbf{V} = \mathbf{Y}_2 \mathbf{Y}_1^{-1}$  and  $\mathbf{E} = \mathbf{Y}_1$  where  $\mathbf{L}$  is an  $N - 1 \times N$  Helmert sub-matrix.

Now, from Caro-Lopera (2008) and Caro-Lopera *et al.* (2009) the configuration density under a non-isotropic elliptical distribution is given by

**Theorem 1.** If  $\mathbf{Y} \sim \mathcal{E}_{N-1 \times K}(\boldsymbol{\mu}_{N-1 \times K}, \boldsymbol{\Sigma}_{N-1 \times N-1} \otimes \mathbf{I}_K, h)$ , for  $\boldsymbol{\Sigma} > \mathbf{0}$ ,  $\boldsymbol{\mu} \neq \mathbf{0}_{N-1 \times K}$ , then the configuration density is given by

$$\begin{aligned} & \frac{\pi^{K^2/2} \Gamma_K \left( \frac{N-1}{2} \right)}{|\boldsymbol{\Sigma}|^{K/2} |\mathbf{U}' \boldsymbol{\Sigma}^{-1} \mathbf{U}|^{N-1} \Gamma_K \left( \frac{K}{2} \right)} \sum_{t=0}^{\infty} \frac{1}{t! \Gamma \left( \frac{K(N-1)}{2} + t \right)} \sum_{r=0}^{\infty} \frac{1}{r!} [\text{tr} (\boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})]^r \\ & \times \sum_{\tau} \frac{\left( \frac{N-1}{2} \right)_{\tau}}{\left( \frac{K}{2} \right)_{\tau}} C_{\tau} (\mathbf{U}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \mathbf{U} (\mathbf{U}' \boldsymbol{\Sigma}^{-1} \mathbf{U})^{-1}) S, \end{aligned} \quad (2.2)$$

where

$$S = \int_0^{\infty} h^{(2t+r)}(y) y^{\frac{K(N-1)}{2} + t - 1} dy < \infty.$$

Our proposal is to use the elliptically contoured distribution to model population configurations (2.2) for certain particular cases. For this purpose, we consider a random sample of  $n$  independent and identically distributed observations  $\mathbf{U}_1, \dots, \mathbf{U}_n$  obtained from

$$\mathbf{Y}_i \sim E_{N-1 \times K}(\boldsymbol{\mu}_{N-1 \times K}, \sigma^2 \mathbf{I}_{N-1} \otimes \mathbf{I}_K, h), \quad i = 1, \dots, n,$$

by means of (2.1).

Now we define the configuration population parameters. Let  $CD(\mathbf{U}; \boldsymbol{\mu}, \sigma^2)$  be the exact configuration density, where  $\boldsymbol{\mu}$  is the location parameter matrix of the configuration population (for the sake of simplicity, the configuration location) and  $\sigma^2$  is the population scale parameter. Both the  $\boldsymbol{\mu}$  and the  $\sigma^2$  parameters are to be estimated. More

exactly, let  $\boldsymbol{\mu} \neq \mathbf{0}_{N-1 \times K}$  be the parameter matrix of the elliptical density  $\mathbf{Y}$  considered in theorem 1; if we write it as  $\boldsymbol{\mu} = (\boldsymbol{\mu}'_1 \mid \boldsymbol{\mu}'_2)'$ , where  $\boldsymbol{\mu}_1 : K \times K$  (nonsingular) and  $\boldsymbol{\mu}_2 : q = N - K - 1 \geq 1 \times K$ , then, according to (2.1), we can define the configuration location parameter matrix  $\mathbf{U} : N - 1 \times K$  as follows:  $\mathbf{U} = (\mathbf{I}_K \mid \boldsymbol{\nu})'$  where  $\boldsymbol{\nu} = \boldsymbol{\mu}_2 \boldsymbol{\mu}_1^{-1}$ ; and  $\boldsymbol{\nu} : q = N - K - 1 \geq 1 \times K$  contains  $q \times K$  configuration location parameters to estimate. Then, taking into account this remark and using the same notation as in Dryden and Mardia (1998, pp. 144-145), we have:

$$\log L(\mathbf{U}_1, \dots, \mathbf{U}_n; \boldsymbol{\nu}, \sigma^2) = \sum_{i=1}^n \log CD(\mathbf{U}_i; \boldsymbol{\nu}, \sigma^2).$$

Finally, the maximum likelihood estimators for the location and scale parameters are

$$(\widetilde{\boldsymbol{\nu}}, \widetilde{\sigma}^2) = \arg \sup_{\boldsymbol{\nu}, \sigma^2} \log L(\mathbf{U}_1, \dots, \mathbf{U}_n; \boldsymbol{\nu}, \sigma^2). \quad (2.3)$$

For the numerical optimisation, we can use a number of routines, which, clearly, are based on the initial point for estimation. In our case, consider the Helmertised landmark data  $\mathbf{Y}_i \sim E_{N-1 \times K}(\boldsymbol{\mu}_{N-1 \times K}, \sigma^2 \mathbf{I}_{N-1} \otimes \mathbf{I}_K, h)$   $i = 1, \dots, n$  (see (2.1)) and let  $\widetilde{\boldsymbol{\mu}} = (\widetilde{\boldsymbol{\mu}}'_1 \mid \widetilde{\boldsymbol{\mu}}'_2)'$  and  $\widetilde{\sigma}^2$  be the maximum likelihood estimators of the location parameter matrix  $\boldsymbol{\mu}_{N-1 \times K}$  and the scale parameter  $\sigma^2$  of the elliptical distribution under consideration. Accordingly, given that

$$\mathbf{U}'_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \mathbf{U}_i (\mathbf{U}'_i \boldsymbol{\Sigma}^{-1} \mathbf{U}_i)^{-1} = \mathbf{Y}'_i \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \mathbf{Y}_i (\mathbf{Y}'_i \boldsymbol{\Sigma}^{-1} \mathbf{Y}_i)^{-1},$$

then an initial point can be  $\mathbf{x}_0 = (\text{vec}'(\boldsymbol{\nu}_0'), \sigma_0^2)$ , where  $\boldsymbol{\nu}_0 = \widetilde{\boldsymbol{\mu}}_2 \widetilde{\boldsymbol{\mu}}_1^{-1}$  and  $\sigma_0^2 = \widetilde{\sigma}^2$ .

The exact inference procedure is outlined in the following sections.

## 2.1 Available Distributions: Families of Isotropic Elliptical Configuration Densities

As a first step, let us consider a list of configuration densities, as derived in full in Caro-Lopera *et al.* (2009). The classical elliptical configuration densities included are the Kotz, Pearson type VII, Bessel and Logistic types. Note, however, that any elliptical function  $h(\cdot)$  which satisfies Theorem 1 would be appropriate.

Most of the applications in statistical theory of shape are based on the isotropic model (see Dryden and Mardia (1998)), and thus in the case of the noncentral elliptical configuration density, by taking  $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}_{N-1}$  in the general configuration density, we obtain a list of suitable distributions for inference; when these are expanded in terms of zonal polynomials, they can be computed efficiently following Koev and Edelman (2006). Note that a more enriched structure can be considered, see for example  $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_{N-1}^2)$  (which assume a different scale parameter in each landmark component), and similar diagonal structures.

We shall not write out such densities here, but limit ourselves to drawing an inference with a special Kotz subfamily; however, it should be noted that the four step inference procedure can be studied with the densities provided in Caro-Lopera *et al.* (2009).

The particular Kotz density which shall be studied in the applications is as follows:

*Corollary 2.1.* If  $\mathbf{Y} \sim E_{N-1 \times K}(\boldsymbol{\mu}_{N-1 \times K}, \sigma^2 \mathbf{I}_{N-1} \otimes \mathbf{I}_K, h)$  and  $T = 1$ , then the Kotz type isotropic noncentral configuration density simplifies to

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right) \text{etr}\left(\frac{R}{\sigma^2} \boldsymbol{\mu}' \mathbf{U} (\mathbf{U}' \mathbf{U})^{-1} \mathbf{U}' \boldsymbol{\mu} - \frac{R}{\sigma^2} \boldsymbol{\mu}' \boldsymbol{\mu}\right)}{\pi^{Kq/2} |\mathbf{I}_K + \mathbf{V}' \mathbf{V}|^{\frac{N-1}{2}} \Gamma_K\left(\frac{K}{2}\right)} \times {}_1F_1\left(-\frac{q}{2}; \frac{K}{2}; -\frac{R}{\sigma^2} \boldsymbol{\mu}' \mathbf{U} (\mathbf{U}' \mathbf{U})^{-1} \mathbf{U}' \boldsymbol{\mu}\right), \quad (2.4)$$

and where  $R = \frac{1}{2}$ , we obtain the normal configuration density.

## 2.2 Choosing the Elliptical Configuration Density

It is here that we see the main advantage of working with elliptical models, i. e. the possibility of choosing a distribution for the landmark data, recalling that the main assumptions for inference in the present study are supported by independent and identically elliptically contoured distributed observations

$$\mathbf{Y}_i \sim E_{N-1 \times K}(\boldsymbol{\mu}_{N-1 \times K}, \sigma^2 \mathbf{I}_{N-1} \otimes \mathbf{I}_K, h), \quad i = 1, \dots, n.$$

According to our assumptions, we can consider Schwarz (1978) as an appropriate technique for choosing the elliptical model. Explicitly, the procedure is as follows: consider  $k$  elliptical models, then perform the maximisation of the likelihood function separately for each model  $j = 1, \dots, k$ , obtaining say,  $M_j(\mathbf{Y}_1, \dots, \mathbf{Y}_n)$ ; then Schwarz's criterion for a large sample is given by

$$\text{Choose the model for which } \log M_j(\mathbf{Y}_1, \dots, \mathbf{Y}_n) - \frac{1}{2} k_j \log n \text{ is largest,}$$

where  $k_j$  is the dimension (number of parameters) of the model  $j$ .

*Remark 1.* The preceding result can be implemented in choosing a shape model, i.e. given an independent and identically distributed random sample of landmark data and a list of shape distributions: pre-shape, size and shape, shape, reflection shape, reflection size and shape, cone, disk (all of these being supported by Euclidean transformations) and configuration (supported by affine transformations), we can select the best shape-transformation model. However, given that the shape distributions based on Euclidean transformations are infinite series of zonal polynomials and that they cannot be computed exactly (see Dryden and Mardia (1998) and Koev and Edelman (2006)), then in order to compare these models with the polynomial configuration densities, first we need to solve the computation problems of the Euclidean models. These comparisons shall be considered in a later study.

### 2.3 Configuration Location

Once the elliptical model has been selected, the estimators of location and scale parameters of configuration can be found by means of (2.3). The crucial point here is the computation of the configuration density; if the selected model is the Gaussian one, then the Matlab algorithms for confluent hypergeometric functions of a matrix argument given by Koev and Edelman (2006) give the solution very efficiently, and in fact this solves the inference problem described by Goodall and Mardia (1993). We emphasize that the cited computation of the  ${}_1F_1(a; c; \mathbf{X})$  series is restricted to its truncation and that is an open question that is addressed in the last section of Koev and Edelman (2006). Nevertheless, the fast algorithms make it possible to constitute numerical experiments until a given precision is reached, and thus the optimisation problem continues to be in terms of truncation and set precision. However, this obviously occurs it is an intrinsic problem of any numerical optimisation problem.

On the other hand, if the selected model is not Gaussian, it might be considered that the problem remains an open one; fortunately, the configuration densities can be computed efficiently by using the method described in Koev and Edelman (2006).

First let us denote the elliptical configuration density of theorem 1 by

$$A {}_1P_1(f(t) : a; c; \mathbf{X}), \quad (2.5)$$

where

$$A = \frac{\pi^{K^2/2} \Gamma_K\left(\frac{N-1}{2}\right)}{|\boldsymbol{\Sigma}|^{\frac{K}{2}} |\mathbf{U}'\boldsymbol{\Sigma}^{-1}\mathbf{U}|^{\frac{N-1}{2}} \Gamma_K\left(\frac{K}{2}\right)}, \quad {}_1P_1(f(t) : a; c; \mathbf{X}) = \sum_{t=0}^{\infty} \frac{f(t)}{t!} \sum_{\tau} \frac{(a)_{\tau}}{(c)_{\tau}} C_{\tau}(\mathbf{X}),$$

$$f(t) = \sum_{r=0}^{\infty} \frac{[\text{tr}(\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})]^r}{r! \Gamma\left(\frac{K(N-1)}{2} + t\right)} \int_0^{\infty} h^{(2t+r)}(y) y^{\frac{K(N-1)}{2} + t - 1} dy,$$

$$\mathbf{X} = \mathbf{U}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\mathbf{U}(\mathbf{U}'\boldsymbol{\Sigma}^{-1}\mathbf{U})^{-1}, \quad a = \frac{N-1}{2}, \quad c = \frac{K}{2}.$$

Unfortunately, the configuration density  $A {}_1P_1(f(t) : a; c; \mathbf{X})$  is an infinite series, given that  $a$  and  $c$  are positive (recall that  $N$  is the number of landmarks,  $K$  is the dimension and  $N - K - 1 \geq 1$ ). Hence a truncation is needed if we want to use it directly by the computation of zonal polynomials.

Now, expression (2.5) belongs to a general class of series, studied by Koev and Edelman (2006, eq. (1.6), p. 3), which can be evaluated at the same computational cost as the efficiently computed hypergeometric function of a matrix argument.

In principle, thus, the configuration densities can be evaluated efficiently using the fast algorithms proposed by Koev and Edelman (2006) and the corresponding inference problem can be solved numerically. At this stage, using for example the compatible Matlab routine *fminsearch* (unconstrained nonlinear optimisation) with the modified Matlab files of Koev and Edelman (2006), we have the estimators for the configuration location and the scale parameter of the “best” elliptical model chosen using Schwarz’s criterion. We arrive then, at the final step.

## 2.4 Hypothesis Testing

Finally, given that the likelihood can be evaluated and optimised, a type of likelihood ratio test can be performed to examine questions such as a particular configuration for a population, or the for differences in configuration between two populations, or the one-dimensional uniform shear of two populations.

In statistical shape analysis, large sample standard likelihood ratio tests are those most frequently used, see for example Dryden and Mardia (1998), by mean of Wilk's theorem. Explicitly, they are used to test whether  $H_0 : \mathbf{U} \in \Omega_0$  versus  $H_a : \mathbf{U} \in \Omega_1$ , where  $\Omega_0 \subset \Omega_1 \subseteq \mathbb{R}^{Kq}$ , with  $\dim(\Omega_0) = p < Kq$  and  $\dim(\Omega_a) = r \leq Kq$ . Thus, the  $-2$  log-likelihood ratio is given by

$$-2 \log \Lambda = 2 \sup_{H_a} \log L(\mathbf{U}, \sigma^2) - 2 \sup_{H_0} \log L(\mathbf{U}, \sigma^2),$$

and by Wilk's theorem for large samples, the distribution of the null hypothesis  $H_0$  (see Dryden and Mardia (1998)) follows

$$-2 \log \Lambda \approx \chi_{r-p}^2.$$

In a similar way we can test differences in configurations between two populations; assuming that the latter hypothesis is rejected, then an interesting test can be performed on the one-dimensional uniform shear of two populations, which determines the amount of deformation axis by axis. Note that classical statistical shape analysis (pre-shape, size and shape, shape, reflection shape, reflection size and shape, cone, disk), which is based on Euclidean transformations assumes that any shape is uniformly deformed in any dimension, which certainly is very idealistic, but the configuration density accepts different degrees of uniform shearing among the axes.

Explicitly, if we wish to test uniform shear in the  $i$  coordinate of two populations, then the testing procedure lies in  $H_0 : \mu_1 \mathbf{B} = \mu_2 \mathbf{B}$  versus  $H_a : \mu_1 \mathbf{B} \neq \mu_2 \mathbf{B}$ , where  $\mathbf{B} = (0, \dots, i, \dots, 0)'$  and the configuration density  $\mathbf{U}$  becomes  $\mathbf{UB}$ . Note that the new configuration density is simpler, since it is just a vector density and is thus easier to compute.

Thus, the whole inference procedure of the above four steps can be carried out for particular landmark data (for example see Dryden and Mardia (1998) and Bookstein (1991)), and thus far we can consider the inference problem numerically solvable.

## 3 Further Simplifications: Polynomial Configuration Densities

Indeed, the whole elliptical configuration problem may be overcome, and there are interesting simplifications which open up promising lines for future work. In this section, we make an introductory exploration of the problem, before concluding the work with some applications.

Although the zonal polynomials are computed very fast, the problem is now one of the convergence and truncation of the above series in order to perform the numerical optimisation. In this respect, Koev and Edelman (2006) state:

“Several problems remain open, among them automatic detection of convergence . . . and it is unclear how to tell when convergence sets in. Another open problem is to determine the best way to truncate the series.”

Thus the numerical difficulties implicit in the truncation of any configuration density series of type (2.5) motivate two areas of investigation: on the one hand, the continuing the numerical approach begun by (Koev and Edelman (2006)) with the confluent hypergeometric functions and extending it to the case of Kotz, Pearson type VII, Bessel or Logistic-type configuration series, for example; or the other, that of proposing a theoretical approach to solve the problem analytically.

Let us now set out the second question, while leaving its implications for future study.

First represent the configuration density as in (2.5).

The above series can be polynomials if we use the following basic principle.

**Lemma 3.1.** *Let  $N - K - 1 \geq 1$  as usual, and consider the infinite configuration density*

$$CD_1 = A {}_1P_1 \left( f(t) : \frac{N-1}{2}; \frac{K}{2}; \mathbf{X} \right).$$

*If the dimension  $K$  is even (odd) and the number of landmarks  $N$  is odd (even), respectively, then the equivalent configuration density*

$$CD_2 = A {}_1P_1 \left( g(t) : - \left( \frac{N-1}{2} - \frac{K}{2} \right); \frac{K}{2}; h(\mathbf{X}) \right),$$

*is a polynomial of degree  $K \left( \frac{N-1}{2} - \frac{K}{2} \right)$  in the latent roots of the matrix  $\mathbf{X}$  (otherwise the series is infinite), for suitable  $f(t)$ ,  $g(t)$  and  $h(\mathbf{X})$ .*

*Proof.* Recall that  $\tau = (t_1, \dots, t_K)$ ,  $t_1 \geq t_2 \geq \dots \geq t_K \geq 0$ , is a partition of  $t$  and

$$(\alpha)_\tau = \prod_{i=1}^K \left( \alpha - \frac{1}{2} (i-1) \right)_{t_i},$$

where

$$(\alpha)_t = \alpha(\alpha+1) \cdots (\alpha+t-1), \quad (\alpha)_0 = 1.$$

Now, if  $K$  is even (odd) and  $N$  is odd (even) then  $-\left(\frac{N-1}{2} - \frac{K}{2}\right) = -\frac{q}{2}$  is a negative integer and clearly  $(-\frac{q}{2})_\tau = 0$  for every  $t \geq \frac{Kq}{2} + 1$ , then  $CD_2$  is a polynomial of degree  $\frac{Kq}{2}$  in the latent roots of  $\mathbf{X}$ .  $\square$

Thus, the truncation problem being addressed, that of an infinite configuration density, can be solved by finding an equivalent polynomial configuration density according to the preceding lemma and by selecting an appropriate number of landmarks in the figure.

Given an elliptical configuration density  $CD_1$  indexed by function  $f(t)$ ,  $a = \frac{N-1}{2} > 0, c = \frac{K}{2} > 0$ , the crucial point consist in finding an integral representation valid for  $c-a = -\frac{q}{2} < 0$  leading to an equivalent elliptical configuration density  $CD_2$  indexed by a function  $g(t)$ .



Then the series  $CD_2$  becomes a polynomial when  $K$  is even (odd) and  $N$  is odd (even), respectively.

We have previously seen this type of relation, when  $f(t)$  is a constant, i.e. in corollary 2.1; in this case lemma 3.1 is reduced to the Kummer relations and the corresponding configuration densities (which include Gaussian density) are polynomials when we select an odd (even) number of landmarks  $N$  according to an even (odd) dimension of  $K$ , respectively. The implications of the polynomial distributions for applications avoid the above-mentioned open problem of truncation, as described in Koev and Edelman (2006).

The above discussion is important for a generalisation of Kummer type relations; for example, equalities for non constant  $f(t)$ , i.e. expressions of  $g(t)$  and  $h(\mathbf{X})$  for non R-normal models (2.4). Some advances in this direction are available from the authors, for example, concerning the generalised Kummer relations for a Kotz type distribution ( $T$  positive integer), while a Pearson type VII model one based on a Beta type integral representation has ratified that  ${}_1P_1(f(t) : a; c; \mathbf{X}) = {}_1P_1(g(t) : c - a; c; -\mathbf{X})$ , for the corresponding  $f, g$ , but in the case of  $c - a > 0$ . The next step is to prove the relations for  $c - a < 0$ , by a Laplace representation type, and then lemma 3.1 can be applied to Kotz type and Pearson type VII configuration densities and the respective series become polynomials.

Meanwhile, fortunately, we can perform inference with the polynomials of Corollary 2.1, especially with the Gaussian case  $R = \frac{1}{2}$ .

*Corollary 3.1.* If  $\mathbf{Y} \sim \mathcal{N}_{N-1 \times K}(\boldsymbol{\mu}_{N-1 \times K}, \sigma^2 \mathbf{I}_{N-1} \otimes \mathbf{I}_K)$ ,  $K$  is even (odd) and  $N$  is odd (even), respectively, then the polynomial isotropic noncentral normal configuration density is given by

$$\frac{\Gamma_K\left(\frac{N-1}{2}\right)}{\pi^{Kq/2} |\mathbf{I}_K + \mathbf{V}'\mathbf{V}|^{\frac{N-1}{2}} \Gamma_K\left(\frac{K}{2}\right)} \text{etr} \left( \frac{1}{2\sigma^2} \boldsymbol{\mu}'\mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\boldsymbol{\mu} - \frac{1}{2\sigma^2} \boldsymbol{\mu}'\boldsymbol{\mu} \right) \\ \times {}_1F_1 \left( -\frac{q}{2}; \frac{K}{2}; -\frac{1}{2\sigma^2} \boldsymbol{\mu}'\mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\boldsymbol{\mu} \right),$$

and it is a polynomial of degree  $K\left(\frac{N-1}{2} - \frac{K}{2}\right)$  in the latent roots of

$$\frac{1}{2\sigma^2} \boldsymbol{\mu}'\mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\boldsymbol{\mu}.$$

## 4 Applications

In this section we consider a planar classical application in statistical shape analysis. The following situations are sufficiently studied by shape based on Euclidean transformations and asymptotic formulae. Here, exact inference is utilised, in the sense that we shall use the exact densities and compute the likelihood exactly by using zonal polynomial theory.

We shall test configuration differences under the exact gaussian configuration density, with applications include biology (mouse vertebrae, gorilla skulls, girl and boy craniofacial studies), medicine (brain MR scans of schizophrenic patients) and image analysis (postcode recognition).

First we start with the two dimensional case. Corollary 3.1 then becomes:

*Corollary 4.1.* If  $\mathbf{Y} \sim \mathcal{N}_{N-1 \times 2}(\boldsymbol{\mu}_{N-1 \times 2}, \sigma^2 \mathbf{I}_{N-1} \otimes \mathbf{I}_2)$ , and  $N$  is odd, then the polynomial two dimensional isotropic noncentral normal configuration density is given by

$$\frac{\Gamma_2\left(\frac{N-1}{2}\right)}{\pi^{N-3} |\mathbf{I}_2 + \mathbf{V}'\mathbf{V}|^{\frac{N-1}{2}} \Gamma_K(1)} \text{etr} \left( \frac{1}{2\sigma^2} \boldsymbol{\mu}'\mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\boldsymbol{\mu} - \frac{1}{2\sigma^2} \boldsymbol{\mu}'\boldsymbol{\mu} \right) \\ \times {}_1F_1 \left( -\frac{N-3}{2}; 1; -\frac{1}{2\sigma^2} \boldsymbol{\mu}'\mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\boldsymbol{\mu} \right),$$

and it is a polynomial of degree  $N - 3$  in the two latent roots of

$$\frac{1}{2\sigma^2} \boldsymbol{\mu}'\mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\boldsymbol{\mu}.$$

Then, we apply the confluent hypergeometric given in the appendix to the class of problems and consider it an area for future study with other elliptical models and situations.

#### 4.1 Biology: Mouse Vertebrae

This problem has been studied in depth by Dryden and Mardia (1998). The data come from an investigation into the effects of selection for body weight on the shape of mouse vertebrae, with the experiments considering the second thoracic vertebra T2 of 30 control (C), 23 large (L) and 23 small (S) bones. The control group contains unselected mice, the large group contains mice selected at each generation according to large body weight and the small group was selected for small body weight. In order to apply the polynomial densities we do not consider the third landmark of the total of 6 proposed (see Dryden and Mardia (1998, p. 10), and the data given in p. 313-316).

Inference is based on (A.1), a confluent hypergeometric polynomial of degree two; then after a very simple computation we have the following configuration locations of the three groups.

Table 1: Mouse vertebra

Group	$\widetilde{\mathcal{V}}_{11}$	$\widetilde{\mathcal{V}}_{12}$	$\widetilde{\mathcal{V}}_{21}$	$\widetilde{\mathcal{V}}_{22}$	$\widetilde{\sigma}^2$
Control	-0.10789	0.15594	0.0005299	-0.97056	0.00165
Large	-0.084983	0.12436	0.00049203	-1.0787	0.0021303
Small	-0.092342	0.21291	0.00052577	-1.018	0.0019863

Then the likelihood ratios (based on  $-2 \log \Lambda \approx \chi_4^2$ ) for the paired tests  $H_0 : \boldsymbol{\mathcal{U}}_1 = \boldsymbol{\mathcal{U}}_2$  vs  $H_a : \boldsymbol{\mathcal{U}}_1 \neq \boldsymbol{\mathcal{U}}_2$ , give the p-values:  $2.8E - 9$  for C-L;  $177.769E - 7$  for L-S and  $3E - 10$  for C-S. Thus, we can say that there is strong evidence for different configuration changes, the most important of which are between small and large, as expected.

## 4.2 Biology: Gorilla Skulls

In this application, Dryden and Mardia (1998) investigate the cranial differences between 29 male and 30 female apes by studying 8 anatomical landmarks. In order to obtain a polynomial configuration density, we removed landmark  $o$  (see Dryden and Mardia (1998, p. 11), and the data in p. 317-318) and the corresponding confluent hypergeometric is a polynomial of fourth degree, see (A.2).

The estimators of the configuration location and scale parameters are given below.

Table 2: Gorilla skulls

Group	$\widetilde{\mathcal{V}}_{11}$	$\widetilde{\mathcal{V}}_{12}$	$\widetilde{\mathcal{V}}_{21}$	$\widetilde{\mathcal{V}}_{22}$
Female	-0.28033	0.31315	-0.42269	-0.59672
Male	-0.33313	0.42484	-0.43594	-0.5734

...	$\widetilde{\mathcal{V}}_{31}$	$\widetilde{\mathcal{V}}_{32}$	$\widetilde{\mathcal{V}}_{41}$	$\widetilde{\mathcal{V}}_{42}$	$\widetilde{\sigma}^2$
...	0.27398	-1.4695	0.7363	-1.2665	0.0042665
...	0.30563	-1.306	0.73169	-1.0594	0.010057

Thus the likelihood ratio (based on  $-2 \log \Lambda \approx \chi_8^2$ ) for  $H_0 : \mathbf{u}_1 = \mathbf{u}_2$  vs  $H_a : \mathbf{u}_1 \neq \mathbf{u}_2$  of the configuration location cranial difference between the sexes of the apes gives a p-value of  $12.74E - 13$ . This clearly ratifies the conclusion of strong evidence for differences between the female and male configuration locations.

## 4.3 Biology: The university School Study Subsample

In this experiment, Bookstein (1991) studies sex shape differences between 8 craniofacial landmarks for 36 normal Ann Arbor boys and 26 girls aged about 8 years. In order to obtain a polynomial configuration density, we discard the Sella landmark (see Bookstein (1991, pp. 401-405)), and then the hypergeometric functions is a polynomial of fourth degree, see (A.2).

Thus the likelihood ratio (based on  $-2 \log \Lambda \approx \chi_8^2$ ) for  $H_0 : \mathbf{u}_1 = \mathbf{u}_2$  vs  $H_a : \mathbf{u}_1 \neq \mathbf{u}_2$  of the configuration location craniofacial difference between the boys and girls gives a p-value of 0.7053. The difference between these two configuration locations is insignificant. A similar overall conclusion was drawn by Bookstein (1991), although a more detailed study of the landmark subset is required for possible differences to be detected, as Bookstein (1991) ratifies in a different shape context.

## 4.4 Medicine: Brain MR Scans of Schizophrenic Patients

Let us return to the applications described in Dryden and Mardia (1998), in their study of 13 landmarks on a near midsagittal two dimensional slice from magnetic resonance (MR)

Table 3: The university school study subsample

Group	$\widetilde{\mathcal{V}}_{11}$	$\widetilde{\mathcal{V}}_{12}$	$\widetilde{\mathcal{V}}_{21}$	$\widetilde{\mathcal{V}}_{22}$
Male	-1.2425	2.1948	0.46435	-1.3752
Female	-1.2483	2.2331	0.43685	-1.3845

...	$\widetilde{\mathcal{V}}_{31}$	$\widetilde{\mathcal{V}}_{32}$	$\widetilde{\mathcal{V}}_{41}$	$\widetilde{\mathcal{V}}_{42}$	$\widetilde{\sigma}^2$
...	-0.91487	0.66127	0.15775	-0.069042	0.0032908
...	-0.92903	0.70439	0.1616	-0.077236	0.0059142

brain scans of 14 schizophrenic patients and 14 normal patients. Given that the number of two dimensional landmarks is odd, we preserve them and thus obtain a 10 degree confluent hypergeometric polynomial, which is straightforwardly computed, see (A.5).

Table 4: Brain MR scans of schizophrenic patients

Group	$\widetilde{\mathcal{V}}_{11}$	$\widetilde{\mathcal{V}}_{12}$	$\widetilde{\mathcal{V}}_{21}$	$\widetilde{\mathcal{V}}_{22}$	$\widetilde{\mathcal{V}}_{31}$	$\widetilde{\mathcal{V}}_{32}$
Normal	-0.64099	2.6942	-1.2744	-2.8323	-0.42155	-1.003
Squizo.	-0.68623	2.393	-1.145	-2.8484	-0.37349	-1.0744

$\widetilde{\mathcal{V}}_{41}$	$\widetilde{\mathcal{V}}_{42}$	$\widetilde{\mathcal{V}}_{51}$	$\widetilde{\mathcal{V}}_{52}$	$\widetilde{\mathcal{V}}_{61}$	$\widetilde{\mathcal{V}}_{62}$
-0.31011	-2.3094	-0.30236	-3.5261	0.36	-0.90135
-0.23173	-2.1929	-0.20173	-3.3226	0.38123	-0.84316

$\widetilde{\mathcal{V}}_{71}$	$\widetilde{\mathcal{V}}_{72}$	$\widetilde{\mathcal{V}}_{81}$	$\widetilde{\mathcal{V}}_{82}$	$\widetilde{\mathcal{V}}_{91}$	$\widetilde{\mathcal{V}}_{92}$
0.1597	-2.2205	0.8518	-0.7578	1.8686	0.86501
0.20429	-2.109	0.84683	-0.56588	1.7948	0.88466

$\widetilde{\mathcal{V}}_{10,1}$	$\widetilde{\mathcal{V}}_{10,2}$	$\widetilde{\sigma}^2$
-0.14205	0.20718	0.010843
-0.079005	0.1378	0.054064

Dryden and Mardia (1998) given a warning about the small sample size of this experiment; obviously it would be possible to account for the opposite result on the basis  $-2 \log \Lambda \approx \chi_{20}^2$ , which means a p-value of 0.9174. Dryden and Mardia (1998) concluded that there were mean shape differences, but the configuration difference is definitely in-

significant. The most important fact here is the geometrical meaning of the data, because it certainly differs from preceding values, which have an explicit geometrical explanation.

#### 4.5 Image Analysis: Postcode Recognition

This, again, is a 13 landmark problem, which supposes a 10 degree confluent hypergeometric polynomial. In this case, Dryden and Mardia (1998) studied 30 random samples of the handwritten digit 3 for the purposes of postcode recognition. The study data are available in Dryden and Mardia (1998, pp. 318-320).

The Table 5 shows, the configuration location and scale parameter estimates, together with the configuration coordinates of a template number 3 digit, with two equal sized arcs, and 13 landmarks (two coincident) lying on two regular octagons (see Dryden and Mardia (1998, p. 153)).

Table 5: Postcode recognition

Group	$\widetilde{\mathcal{V}}_{11}$	$\widetilde{\mathcal{V}}_{12}$	$\widetilde{\mathcal{V}}_{21}$	$\widetilde{\mathcal{V}}_{22}$	$\widetilde{\mathcal{V}}_{31}$	$\widetilde{\mathcal{V}}_{32}$
Digit 3	-0.79087	1.9432	-2.1073	1.5875	-2.713	0.81862
Template	-2.0908	2.2071	-4.0409	2.8051	-4.5904	2.2904

$\widetilde{\mathcal{V}}_{41}$	$\widetilde{\mathcal{V}}_{42}$	$\widetilde{\mathcal{V}}_{51}$	$\widetilde{\mathcal{V}}_{52}$	$\widetilde{\mathcal{V}}_{61}$	$\widetilde{\mathcal{V}}_{62}$
-2.8084	-0.066901	-2.5712	0.71315	-2.6934	1.2955
-4.2069	1.3688	-3.3126	1.7582	-3.5881	2.7053

$\widetilde{\mathcal{V}}_{71}$	$\widetilde{\mathcal{V}}_{72}$	$\widetilde{\mathcal{V}}_{81}$	$\widetilde{\mathcal{V}}_{82}$	$\widetilde{\mathcal{V}}_{91}$	$\widetilde{\mathcal{V}}_{92}$
-3.1548	1.6802	-3.8004	1.34	-4.0517	0.33141
-5.4996	4.0629	-7.5557	4.8428	-8.2514	4.4208

$\widetilde{\mathcal{V}}_{10,1}$	$\widetilde{\mathcal{V}}_{10,2}$	$\widetilde{\sigma}^2$
-3.7659	-0.6583	0.22904
-6.9108	2.8899	

Thus the likelihood ratio based on  $-2 \log \Lambda \approx \chi_{20}^2$ , gives an approximately zero p-value. This result was corroborated with a probability of  $\approx 0.0002$  by Dryden and Mardia (1998, p. 153), under a shape model. In any case, there is strong evidence that the configuration location does not have the configuration of the ideal template for the digit 3.

Finally, let us note that the remaining planar applications in Dryden and Mardia (1998) and Bookstein (1991), etc. can be studied with the polynomial configuration densities and

exact formulae for zonal polynomials; in fact, the three dimensional applications available in the literature (see Goodall and Mardia (1993)) and others in genetics for 3D DNA part, etc., can be studied in an exact form with the help of corollary 2.1 via lemma 3.1 and exact formulae for zonal polynomials of third degree in James (1964), avoiding the open truncation problems implicit in Koev and Edelman (2006). Moreover, some comparisons among the shape models can be performed via Remark 1 and other tests for uniform deformations (see Subsection 2.4) can be performed. This shall be considered in a subsequent study. Of course, the study of polynomial configuration densities associated to Pearson, Bessel, logistic and the general Kotz distribution shall facilitate exact inference and avoid the truncation problem addressed here, but this shall depend on developments in integration and series representation. In fact, potentially, the distribution of the likelihood ratio could be studied by using low-degree polynomial configuration densities. These questions are currently being investigated.

## Acknowledgements

The authors wish to express their gratitude to the editor and the referee for their helpful comments and suggestions. This research work was partially supported by CONACYT-México, research grant no. 45974-F, 81512 and IDI-Spain, grant MTM2005-09209. This paper was written during J. A. Díaz-García's stay as a visiting professor at the Department of Statistics and O. R. of the University of Granada, Spain.

## References

- Bookstein, F. L. (1991). *Morphometric tools for landmark data*. Cambridge, England: Cambridge University Press.
- Caro-Lopera, F. J. (2008) Noncentral Elliptical Configuration Density. Ph.D. Thesis, CIMAT, A.C. México, 2008.
- Caro-Lopera, F. J., Díaz-García, J. A., and González-Farías, G. (2009). Elliptical configuration distribution. *Journal of Multivariate Analysis*. doi:10.1016/j.jmva.2009.03.004. In press.
- Caro-Lopera, F. J., Díaz-García, J. A., and González-Farías, G. (2007). A formula for Jack polynomials of the second order. *Applicationes Mathematicae*, (Warsaw) **34**, 113–119.
- Díaz-García, J. A., Gutiérrez, J. R., and Ramos, R. (2003). Size-and-Shape Cone, Shape Disk and Configuration Densities for the Elliptical Models. *Brazilian Journal of Probability and Statistics*, **17**, 135–146.
- Dryden, I. L., and Mardia, K. V. (1998). *Statistical shape analysis*. John Wiley and Sons, Chichester.

- Goodall, C. R., and Mardia, K. V. (1993). Multivariate Aspects of Shape Theory. *The Annals of Statistics*, **21**, 848–866.
- Gupta, A. K., and Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Kluwer Academic Publishers, Dordrecht.
- James, A. T. (1968). Calculation of zonal polynomial coefficients by use of the Laplace-Beltrami operator, *The Annals of Mathematical Statistics*, **39**, 1711–1718.
- James, A. T. (1964). Distributions of matrix variate and latent roots derived from normal samples, *The Annals of Mathematical Statistics*, **35**, 475–501.
- Koev, P., and Edelman, A. (2006). The efficient evaluation of the hypergeometric function of a matrix argument, *Mathematics of Computation*, **75**, 833–846.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**, 461–464.

## A A Finite Series for Planar Applications of Maximum of 21 Landmarks

Given that most applications in shape theory come from two-dimensional images (see Dryden and Mardia (1998)), it is important to give explicit expressions for the polynomials  ${}_1F_1(- (N-3)/2; 1; -\boldsymbol{\mu}'\mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\boldsymbol{\mu}/2\sigma^2)$  involved in corollary 4.1 when  $N = 5, 7, 9, \dots$  is small. Let  $x, y$  be the eigenvalues of  $\boldsymbol{\Omega} = \boldsymbol{\mu}'\mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}'\boldsymbol{\mu}/2\sigma^2$ , then we have for  $N = 5, 7, \dots, 21$  the following list of polynomials of degree  $Kq/2 = N - 3$ ; these expressions are useful for exact inference of the corresponding configuration densities. In this case, we use the exact formulae for zonal polynomials given by James (1968) -see also Caro-Lopera *et al.* (2007). In fact, all the applications studied in Dryden and Mardia (1998) have maximum of 21 landmarks (which implies a polynomial of 18 degrees in the two eigenvalues of the corresponding matrix), and therefore the following confluent hypergeometric expressions are sufficient for their corresponding configuration analysis. Note that the cited applications demand formulae for second-order zonal polynomials up to maximum of 18 degree; these expressions have been available since the 1960s, and so the numerical algorithms of Koev and Edelman (2006), very useful for infinite series, but with the addressed problem of truncations, are not needed here. It has been possible to study the exact inference on configuration densities since the configurations were first proposed by Goodall and Mardia (1993).

Observe that the selection of an odd number of landmarks for planar applications suggests that one of the available remarks usually studied for approximation methods could be deleted. Clearly, it is also possible to reduce by one any group of preset even landmarks; however we shall leave the decision to an expert. In view of the number of odd landmarks,

we suggest examination of some problems studied by Dryden and Mardia (1998) but in the context of polynomial Gaussian configuration densities (in parentheses the original number of landmarks studied by Dryden and Mardia (1998)).

Series with up to 15 landmarks are easily computed (the figure in parentheses indicates the number of landmarks in the original source of Dryden and Mardia (1998)(DM) and Bookstein (1991)(B), respectively):

- $N = 5$ : Mouse vertebra (6)(DM),

$$1 + y + x + 2yx \quad (\text{A.1})$$

- $N = 7$ : Gorilla skulls (8)(DM), the university school study subsample (8)(B),

$$1 + 2y + 2x + \frac{1}{2}y^2 + 7yx + \frac{1}{2}x^2 + 2y^2x + 2yx^2 + \frac{2}{3}y^2x^2 \quad (\text{A.2})$$

- $N = 9$ :

$$\begin{aligned} & 1 + 3y + 3x + \frac{3}{2}y^2 + 15yx + \frac{3}{2}x^2 + \frac{1}{6}y^3 + \frac{17}{2}y^2x \\ & + \frac{17}{2}yx^2 + \frac{1}{6}x^3 + y^3x + \frac{16}{3}y^2x^2 + yx^3 + \frac{2}{3}y^3x^2 \\ & + \frac{2}{3}y^2x^3 + \frac{4}{45}y^3x^3 \end{aligned} \quad (\text{A.3})$$

- $N = 11$ : Sooty mangabeys (12)(DM).

$$\begin{aligned} & 1 + 4x + 22y^2x + 4y + \frac{81}{4}y^2x^2 + 5y^2x^3 + 5y^3x^2 + 22yx^2 \\ & + \frac{31}{6}y^3x + \frac{31}{6}yx^3 + 26yx + \frac{1}{24}x^4 + \frac{1}{3}yx^4 + \frac{1}{3}y^4x \\ & + \frac{2}{315}y^4x^4 + \frac{4}{45}y^4x^3 + \frac{4}{45}y^3x^4 + \frac{1}{3}y^4x^2 + \frac{58}{45}y^3x^3 \\ & + \frac{1}{3}y^2x^4 + 3y^2 + 3x^2 + \frac{2}{3}y^3 + \frac{2}{3}x^3 + \frac{1}{24}y^4 \end{aligned} \quad (\text{A.4})$$

- $N = 13$ : Brain MR scans of schizophrenic patients (13)(DM), postcode recognition (13)(DM)

$$\begin{aligned} & 1 + 5x + 45y^2x + 5y + \frac{655}{12}y^2x^2 + \frac{241}{12}y^2x^3 + \frac{241}{12}y^3x^2 + 45yx^2 \\ & + \frac{95}{6}y^3x + \frac{95}{6}yx^3 + 40yx + \frac{5}{24}x^4 + \frac{49}{24}yx^4 + \frac{49}{24}y^4x + \frac{1}{12}yx^5 \\ & + \frac{4}{14175}y^5x^5 + \frac{2}{315}y^5x^4 + \frac{2}{315}y^4x^5 + \frac{2}{45}y^5x^3 + \frac{46}{315}y^4x^4 \\ & + \frac{2}{45}y^3x^5 + \frac{1}{9}y^5x^2 + \frac{47}{45}y^4x^3 + \frac{47}{45}y^3x^4 + \frac{1}{9}y^2x^5 \\ & + \frac{1}{12}y^5x + \frac{8}{3}y^4x^2 + \frac{689}{90}y^3x^3 + \frac{8}{3}y^2x^4 + 5y^2 \\ & + 5x^2 + \frac{5}{3}y^3 + \frac{5}{3}x^3 + \frac{5}{24}y^4 + \frac{1}{120}y^5 + \frac{1}{120}x^5 \end{aligned} \quad (\text{A.5})$$



- $N = 15$ :

$$\begin{aligned}
& 1 + 6x + 80y^2x + 6y + \frac{1445}{12}y^2x^2 + \frac{353}{6}y^2x^3 + \frac{353}{6}y^3x^2 + 80yx^2 \\
& + \frac{75}{2}y^3x + \frac{75}{2}yx^3 + 57yx + \frac{5}{8}x^4 + \frac{29}{4}yx^4 + \frac{29}{4}y^4x + \frac{71}{120}yx^5 \\
& + \frac{134}{14175}y^5x^5 + \frac{34}{315}y^5x^4 + \frac{34}{315}y^4x^5 + \frac{1}{36}y^6x^2 + \frac{23}{45}y^5x^3 \\
& + \frac{263}{210}y^4x^4 + \frac{23}{45}y^3x^5 + \frac{1}{36}y^2x^6 + \frac{1}{60}y^6x + \frac{35}{36}y^5x^2 + \frac{181}{30}y^4x^3 \\
& + \frac{181}{30}y^3x^4 + \frac{35}{36}y^2x^5 + \frac{1}{60}yx^6 + \frac{2}{135}y^6x^3 + \frac{2}{135}y^3x^6 + \frac{1}{315}y^6x^4 \\
& + \frac{1}{315}y^4x^6 + \frac{4}{14175}y^6x^5 + \frac{4}{467775}y^6x^6 + \frac{4}{14175}y^5x^6 + \frac{1}{720}x^6 \\
& + \frac{1}{720}y^6 + \frac{71}{120}y^5x + \frac{187}{16}y^4x^2 + \frac{5339}{180}y^3x^3 + \frac{187}{16}y^2x^4 + \frac{15}{2}y^2 \\
& + \frac{15}{2}x^2 + \frac{10}{3}y^3 + \frac{10}{3}x^3 + \frac{5}{8}y^4 + \frac{1}{20}y^5 + \frac{1}{20}x^5
\end{aligned} \tag{A.6}$$

- $N = 17$ :

$$\begin{aligned}
& 1 + 7x + 7y + \frac{1}{945}y^4x^7 + 77yx + \frac{21}{2}y^2 + \frac{21}{2}x^2 + \frac{259}{2}y^2x \\
& + \frac{259}{2}yx^2 + \frac{931}{4}y^2x^2 + \frac{35}{6}y^3 + \frac{35}{6}x^3 + \frac{455}{6}y^3x + \frac{455}{6}yx^3 \\
& + \frac{567}{4}y^3x^2 + \frac{567}{4}y^2x^3 + \frac{3199}{36}y^3x^3 + \frac{8}{42567525}y^7x^7 + \frac{35}{24}y^4 \\
& + \frac{35}{24}x^4 + \frac{469}{24}y^4x + \frac{469}{24}yx^4 + \frac{1799}{48}y^4x^2 + \frac{1799}{48}y^2x^4 \\
& + \frac{1}{945}y^7x^4 + \frac{3457}{144}y^4x^3 + \frac{3457}{144}y^3x^4 + \frac{416}{63}y^4x^4 + \frac{2}{14175}y^5x^7 \\
& + \frac{7}{40}y^5 + \frac{7}{40}x^5 + \frac{287}{120}y^5x + \frac{287}{120}yx^5 + \frac{1121}{240}y^5x^2 + \frac{1121}{240}y^2x^5 \\
& + \frac{73}{24}y^5x^3 + \frac{73}{24}y^3x^5 + \frac{2}{14175}y^7x^5 + \frac{107}{126}y^5x^4 + \frac{107}{126}y^4x^5 \\
& + \frac{523}{4725}y^5x^5 + \frac{4}{467775}y^7x^6 + \frac{4}{467775}y^6x^7 + \frac{7}{720}y^6 + \frac{7}{720}x^6 \\
& + \frac{97}{720}y^6x + \frac{97}{720}yx^6 + \frac{4}{15}y^6x^2 + \frac{4}{15}y^2x^6 + \frac{19}{108}y^6x^3 + \frac{19}{108}y^3x^6 \\
& + \frac{47}{945}y^6x^4 + \frac{47}{945}y^4x^6 + \frac{31}{4725}y^6x^5 + \frac{31}{4725}y^5x^6 + \frac{184}{467775}y^6x^6 \\
& + \frac{1}{5040}y^7 + \frac{1}{5040}x^7 + \frac{1}{360}y^7x + \frac{1}{360}yx^7 + \frac{1}{180}y^7x^2 \\
& + \frac{1}{180}y^2x^7 + \frac{1}{270}y^7x^3 + \frac{1}{270}y^3x^7
\end{aligned} \tag{A.7}$$

- $N = 19$ :

$$\begin{aligned}
& 1 + 8x + 8y + \frac{31}{1890}y^4x^7 + 100yx + 14y^2 + 14x^2 + 196y^2x \\
& + 196yx^2 + \frac{819}{2}y^2x^2 + \frac{28}{3}y^3 + \frac{28}{3}x^3 + \frac{413}{3}y^3x + \frac{413}{3}yx^3 \\
& + \frac{896}{3}y^3x^2 + \frac{896}{3}y^2x^3 + \frac{10073}{45}y^3x^3 + \frac{44}{3869775}y^7x^7 + \frac{35}{12}y^4 \\
& + \frac{35}{12}x^4 + \frac{133}{3}y^4x + \frac{133}{3}yx^4 + \frac{1183}{12}y^4x^2 + \frac{1183}{12}y^2x^4 + \frac{31}{1890}y^7x^4 \\
& + \frac{1357}{18}y^4x^3 + \frac{1357}{18}y^3x^4 + \frac{104071}{4032}y^4x^4 + \frac{41}{14175}y^5x^7 + \frac{7}{15}y^5 \\
& + \frac{7}{15}x^5 + \frac{217}{30}y^5x + \frac{217}{30}yx^5 + \frac{1}{40320}x^8 + \frac{491}{30}y^5x^2 + \frac{491}{30}y^2x^5 \\
& + \frac{1}{2520}y^8x + \frac{9151}{720}y^5x^3 + \frac{9151}{720}y^3x^5 + \frac{41}{14175}y^7x^5 + \frac{247}{56}y^5x^4 \\
& + \frac{247}{56}y^4x^5 + \frac{3089}{4050}y^5x^5 + \frac{122}{467775}y^7x^6 + \frac{122}{467775}y^6x^7 + \frac{1}{1080}y^2x^8 \\
& + \frac{7}{180}y^6 + \frac{7}{180}x^6 + \frac{11}{18}y^6x + \frac{11}{18}yx^6 + \frac{2017}{1440}y^6x^2 + \frac{2017}{1440}y^2x^6 \\
& + \frac{1189}{1080}y^6x^3 + \frac{1189}{1080}y^3x^6 + \frac{1}{2520}yx^8 + \frac{2921}{7560}y^6x^4 + \frac{2921}{7560}y^4x^6 \\
& + \frac{319}{4725}y^6x^5 + \frac{319}{4725}y^5x^6 + \frac{1129}{187110}y^6x^6 + \frac{1}{40320}y^8 + \frac{1}{630}y^7 \\
& + \frac{1}{630}x^7 + \frac{127}{5040}y^7x + \frac{1}{1080}y^8x^2 + \frac{127}{5040}yx^7 + \frac{7}{120}y^7x^2 + \frac{7}{120}y^2x^7 \\
& + \frac{5}{108}y^7x^3 + \frac{5}{108}y^3x^7 + \frac{1}{1350}y^8x^3 + \frac{1}{1350}y^3x^8 + \frac{1}{3780}y^8x^4 \\
& + \frac{1}{3780}y^4x^8 + \frac{2}{42525}y^8x^5 + \frac{2}{42525}y^5x^8 + \frac{2}{467775}y^8x^6 + \frac{2}{467775}y^6x^8 \\
& + \frac{8}{42567525}y^8x^7 + \frac{8}{42567525}y^7x^8 + \frac{2}{638512875}y^8x^8
\end{aligned} \tag{A.8}$$

- $N = 21$ : Microfossils (21)(DM).

$$\begin{aligned}
& 1 + 9x + 9y + \frac{109}{840}y^4x^7 + \frac{1}{8100}y^3x^9 + 126yx + 18y^2 + 18x^2 \\
& + 282y^2x + 282yx^2 + \frac{1343}{2}y^2x^2 + 14y^3 + 14x^3 + \frac{1}{7560}y^9x^2 + 231y^3x \\
& + 231yx^3 + \frac{2}{1403325}y^6x^9 + \frac{1141}{2}y^3x^2 + \frac{1141}{2}y^2x^3 + \frac{1}{85050}y^9x^5 \\
& + \frac{7462}{15}y^3x^3 + \frac{349}{1576575}y^7x^7 + \frac{1}{18900}y^9x^4 + \frac{21}{4}y^4 + \frac{21}{4}x^4 + \frac{357}{4}y^4x \\
& + \frac{357}{4}yx^4 + \frac{1}{18900}y^4x^9 + \frac{903}{4}y^4x^2 + \frac{903}{4}y^2x^4 + \frac{109}{840}y^7x^4 + \frac{4013}{20}y^4x^3 \\
& + \frac{4013}{20}y^3x^4 + \frac{184099}{2240}y^4x^4 + \frac{2}{638512875}y^8x^9 + \frac{1613}{56700}y^5x^7 + \frac{21}{20}y^5 \\
& + \frac{21}{20}x^5 + \frac{91}{5}y^5x + \frac{91}{5}yx^5 + \frac{1}{4480}x^8 + \frac{2809}{60}y^5x^2 + \frac{2809}{60}y^2x^5
\end{aligned}$$

$$\begin{aligned}
& + \frac{23}{5760} y^8 x + \frac{10133}{240} y^5 x^3 + \frac{10133}{240} y^3 x^5 + \frac{1613}{56700} y^7 x^5 + \frac{353047}{20160} y^5 x^4 \\
& + \frac{353047}{20160} y^4 x^5 + \frac{1711063}{453600} y^5 x^5 + \frac{1061}{311850} y^7 x^6 + \frac{1061}{311850} y^6 x^7 + \frac{1}{8100} y^9 x^3 \\
& + \frac{2}{189} y^2 x^8 + \frac{7}{60} y^6 + \frac{7}{60} x^6 + \frac{41}{20} y^6 x + \frac{41}{20} y x^6 + \frac{2563}{480} y^6 x^2 + \frac{2563}{480} y^2 x^6 \\
& + \frac{21041}{4320} y^6 x^3 + \frac{21041}{4320} y^3 x^6 + \frac{23}{5760} y x^8 + \frac{2}{1403325} y^9 x^6 + \frac{643}{315} y^6 x^4 \\
& + \frac{643}{315} y^4 x^6 + \frac{50333}{113400} y^6 x^5 + \frac{50333}{113400} y^5 x^6 + \frac{2}{638512875} y^9 x^8 \\
& + \frac{49279}{935550} y^6 x^6 + \frac{1}{7560} y^2 x^9 + \frac{1}{4480} y^8 + \frac{1}{140} y^7 + \frac{1}{140} x^7 + \frac{71}{560} y^7 x \\
& + \frac{2}{189} y^8 x^2 + \frac{71}{560} y x^7 + \frac{4}{42567525} y^7 x^9 + \frac{3361}{10080} y^7 x^2 + \frac{3361}{10080} y^2 x^7 \\
& + \frac{221}{720} y^7 x^3 + \frac{221}{720} y^3 x^7 + \frac{1}{20160} y x^9 + \frac{1}{20160} y^9 x + \frac{53}{5400} y^8 x^3 \\
& + \frac{53}{5400} y^3 x^8 + \frac{79}{18900} y^8 x^4 + \frac{79}{18900} y^4 x^8 + \frac{4}{97692469875} y^9 x^9 \\
& + \frac{157}{170100} y^8 x^5 + \frac{157}{170100} y^5 x^8 + \frac{1}{85050} y^5 x^9 + \frac{52}{467775} y^8 x^6 \\
& + \frac{52}{467775} y^6 x^8 + \frac{1}{362880} x^9 + \frac{62}{8513505} y^8 x^7 + \frac{62}{8513505} y^7 x^8 \\
& + \frac{2}{8292375} y^8 x^8 + \frac{1}{362880} y^9 + \frac{4}{42567525} y^9 x^7
\end{aligned} \tag{A.9}$$