

# Symmetry of Fuzzy Data

Hagit Zabrodsky\*

Shmuel Peleg\*

David Avnir<sup>†</sup>

Institute of Computer Science\* and Department of Organic Chemistry<sup>†</sup>  
The Hebrew University of Jerusalem  
91904 Jerusalem, Israel

## Abstract

*Symmetry is usually viewed as a discrete feature: an object is either symmetric or non-symmetric. Following the view that symmetry is a continuous feature, a Continuous Symmetry Measure (CSM) has been developed to evaluate symmetries of shapes and objects. In this paper we extend the symmetry measure to evaluate the imperfect symmetry of fuzzy shapes, i.e. shapes with uncertain point localization. We find the probability distribution of symmetry values for a given fuzzy shape. Additionally, for every such fuzzy shape, we find the most probable symmetric shape.*

## 1 Introduction

One of the basic features of shapes and objects is symmetry. Symmetry is considered a pre-attentive feature which enhances recognition and reconstruction of shapes and objects [2]. Symmetry is also an important parameter in physical and chemical processes and is an important criterion in medical diagnosis.

The exact mathematical definition of symmetry [4] is inadequate to describe and quantify the symmetries found in the natural world nor those found in the visual world. Furthermore, even perfectly symmetric objects lose their exact symmetry when projected onto an image plane or retina due to occlusion, self-occlusion, digitization, etc.

Previous work [5] introduced a symmetry measure to define and quantify the deviation of shapes and objects from perfect symmetry. This work was extended to deal with evaluating the deviation from perfect symmetry of incomplete data as appears in occluded shapes [6]. In most cases, however, sensing processes do not have absolute accuracy and the location of each point in a sensed pattern is given only as a probability distribution - a *fuzzy shape*. In this paper we continue to deal with evaluating symmetry of incomplete data, specifically evaluating symmetry of fuzzy or uncertain data.

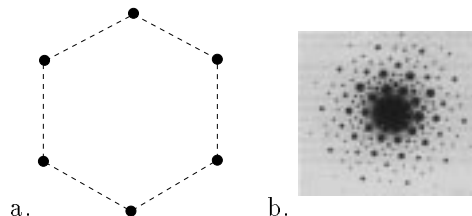


Figure 1: a) A perfectly  $D_6$ -symmetric configuration of points. b) Interference pattern of crystals.

Fig. 1a shows a perfect ( $D_6$ ) symmetric configuration of points. The location of these points (marked as dots) are given precisely. Fig. 1b shows an interference pattern created by projecting X-ray beams onto crystals. Crystal quality is measured by evaluating the symmetry of these interference patterns. These patterns represent uncertain locations (the dark blobs) of point data. Extension of the symmetry measure to quantify the symmetry content of uncertain data, can be directly applied to evaluating patterns similar to these interference patterns.

In the next section we briefly review the symmetry measure as applied to 2D shapes. In Section 3 we extend the symmetry measure to deal with uncertain or fuzzy data. In Section 4 we give mathematical derivations of the methods described in Section 3.

## 2 A Symmetry Measure

The **Symmetry Measure** as described in [5] quantifies the minimum effort necessary to turn a given shape into a symmetric shape. This effort is measured by the sum of square distances each point is moved from its location in the original shape to its location in the symmetric shape.

A shape  $P$  is represented by a sequence of  $n$  points  $\{P_i\}_{i=0}^{n-1}$ . We define a distance between every two shapes  $P$  and  $Q$ :

$$d(P, Q) = d(\{P_i\}, \{Q_i\}) = \frac{1}{n} \sum_{i=1}^n \|P_i - Q_i\|^2$$

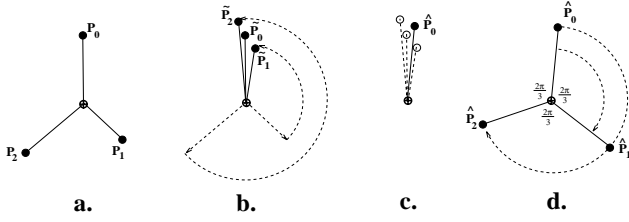


Figure 2: The  $C_3$ -symmetry Transform of 3 points: a) original points  $\{P_i\}_{i=0}^2$ . b) Fold  $\{P_i\}_{i=0}^2$  into  $\{\tilde{P}_i\}_{i=0}^2$ . c) Average  $\{\tilde{P}_i\}_{i=0}^2$  obtaining  $\hat{P}_0 = \frac{1}{3} \sum_{i=0}^2 \tilde{P}_i$ . d) Unfold  $\hat{P}_0$  obtaining  $\{\hat{P}_i\}_{i=0}^2$ .

We define the **Symmetry Transform**  $\hat{P}$  of  $P$  as the symmetric shape closest to  $P$  in terms of distance  $d$ .

The **Symmetry Measure** of  $P$  denoted  $s(P)$  is now defined as the distance to the closest symmetric shape:

$$s(P) = d(P, \hat{P})$$

The CSM of a shape  $P = \{P_i\}_{i=0}^{n-1}$  is evaluated by finding the symmetry transform  $\hat{P}$  of  $P$  and then computing:  $s(P) = \frac{1}{n} \sum_{i=0}^{n-1} \|P_i - \hat{P}_i\|^2$ . Following is a geometrical algorithm for deriving the symmetry transform of a shape  $P$  having  $n$  points with respect to rotational symmetry of order  $n$  ( $C_n$ -symmetry). Mathematical derivation and proof can be found in [7]. This method transforms  $P$  into a regular  $n$ -gon, keeping the centroid in place.

1. *Fold* the points  $\{P_i\}_{i=0}^{n-1}$  by rotating each point  $P_i$  counterclockwise about the centroid by  $2\pi i/n$  radians (Fig. 2b). The “folding” takes  $P_i$  to  $\tilde{P}_i$ , where  $\tilde{P}_0 = P_0$ .
2. Let  $\hat{P}_0$  be the average of the folded points  $\{\tilde{P}_i\}_{i=0}^{n-1}$  (Fig. 2c).
3. *Unfold* the points, obtaining the  $C_n$ -symmetric points  $\{\hat{P}_i\}_{i=0}^{n-1}$  by duplicating  $\hat{P}_0$  and rotating clockwise about the centroid by  $2\pi i/n$  radians (Fig. 2d).

A 2D shape  $P$  having  $qn$  points is represented as  $q$  sets  $\{S_r\}_{r=0}^{q-1}$  of  $n$  interlaced points  $S_r = \{P_{in+r}\}_{i=0}^{n-1}$ . The  $C_n$ -symmetry transform of  $P$  is obtained by applying the above algorithm to each set of  $n$  points separately,

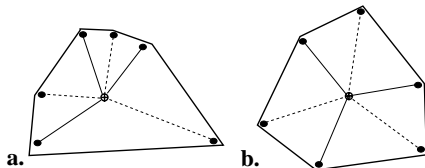


Figure 3: The  $C_3$ -symmetry transform for a 6-sided polygon. The centroid of the polygon is marked by  $\oplus$ . a) The original polygon shown as two sets of 3 points. b) The  $C_3$ -symmetric shape obtained.

where the folding is performed about the centroid of all the points (Fig. 3). The procedure for evaluating the symmetry transform for mirror-symmetry is similar (see [5]).

### 3 Symmetry of points with uncertain locations

In most cases, sensors do not have absolute accuracy and the location of each point in a sensed pattern can be given only as a probability distribution. Given sensed points with such uncertain locations, the following properties are of interest:

- The most probable symmetric configuration represented by the sensed points.
- The probability distribution of symmetry distance values for the sensed points.

#### 3.1 The most probable symmetric shape

Fig. 4a shows a configuration of points whose locations are given by a normal distribution function. The dot represents the expected location of the point and the rectangle represents the uncertainty of the location, where the width and length of the rectangle are proportional to the standard deviation. In this section we describe a method of evaluating the most probable symmetric shape under the Maximum Likelihood criterion given the sensed points. Detailed derivations and proofs are given in Section 4.1. For simplicity we describe the method with respect to rotational symmetry of order  $n$  ( $C_n$ -symmetry). The solution for mirror symmetry or any other symmetry is similar.

Given  $n$  ordered points in 2D whose locations are given as normal probability distributions with expected location  $P_i$  and covariance matrix  $\Lambda_i$ :

$Q_i \sim \mathcal{N}(P_i, \Lambda_i) \quad i=0 \dots n-1$ , we find the  $C_n$ -symmetric configuration of points at locations  $\{\hat{P}_i\}_{i=0}^{n-1}$  which is optimal under the Maximum Likelihood criterion.

Denote by  $\omega$  the unknown centroid of the most probable  $C_n$ -symmetric set of locations  $\hat{P}_i$ :

$\omega = \frac{1}{n} \sum_{i=0}^{n-1} \hat{P}_i$ . The point  $\omega$  is dependent on the location of the measurements ( $P_i$ ) and on the probability distribution associated with them ( $\Lambda_i$ ). Intuitively,  $\omega$  is positioned at that point about which the folding (described below) gives the tightest cluster of points with small uncertainty (small s.t.d.). We assume for the moment that the centroid  $\omega$  is given. A method for finding  $\omega$  is derived in Section 4.1. We use a variant of the folding method which was described in Section 2 for evaluating  $C_n$ -symmetry of a set of points:

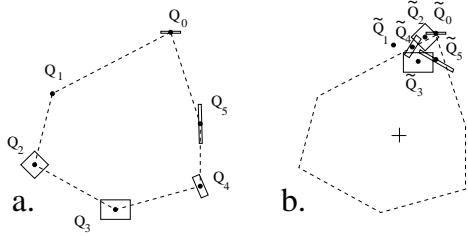


Figure 4: Folding measured points. a) A configuration of 6 measured points  $Q_0 \dots Q_5$ . The dot represents the expected location of the point and the rectangle has width and length proportional to the standard deviation. b) Each measurement  $Q_i$  was rotated by  $2\pi i/6$  radians about the centroid of the expected point locations (marked as '+') obtaining measurement  $\tilde{Q}_0 \dots \tilde{Q}_5$ .

1. The  $n$  measurements  $Q_i \sim \mathcal{N}(P_i, \Lambda_i)$  are *folded* by rotating each measurement  $Q_i$  by  $2\pi i/n$  radians about the centroid  $\omega$ . A new set of measurements  $\tilde{Q}_i \sim \mathcal{N}(\tilde{P}_i, \tilde{\Lambda}_i)$  is obtained (see Fig. 4b).
2. The folded measurements are *averaged* using a weighted average, obtaining a single point  $\hat{P}_0$ . Averaging is done by considering the  $n$  folded measurements  $\tilde{Q}_i$  as  $n$  measurements of a single point and  $\hat{P}_0$  represents the most probable location of that point under the Maximum Likelihood criterion.

$$\hat{P}_0 - \omega = \left( \sum_{j=0}^{n-1} \tilde{\Lambda}_j^{-1} \right)^{-1} \sum_{i=0}^{n-1} \tilde{\Lambda}_i^{-1} \tilde{P}_i - \omega$$

3. The “average” point  $\hat{P}_0$  is *unfolded* as described in Section 2 obtaining points  $\{\tilde{P}_i\}_{i=0}^{n-1}$  which are perfectly  $C_n$ -symmetric.

When we are given  $m = qn$  measurements, we find the most probable  $C_n$ -symmetric configuration of points, similar to the folding method of Section 2. The  $m$  measurements  $\{Q_i\}_{i=0}^{m-1}$ , are divided into  $q$  interlaced sets of  $n$  points each, and the folding method as described above is applied separately to each set of measurements. Derivations and proof of this case are also given in Section 4.1.

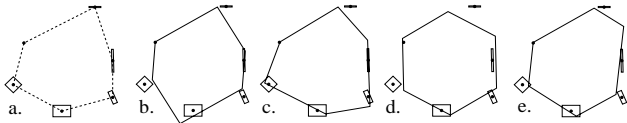


Figure 5: The most probable symmetric shapes. a) A configuration of 6 measured points and the most probable symmetric shapes with respect to b)  $C_2$ -symmetry, c)  $C_3$ -symmetry, d)  $C_6$ -symmetry, and e) mirror-symmetry.

Several examples are shown in Fig. 5, where for a given set of measurements, the most probable symmetric shapes were found. Fig. 6 shows the effects of varying the probability distribution of the measurements on the resulting symmetric shape.

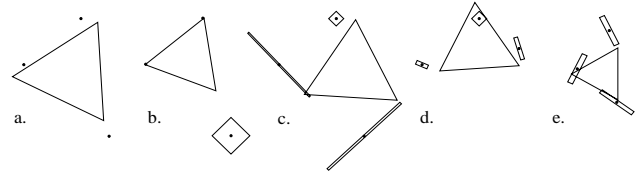


Figure 6: The most probable  $C_3$ -symmetric shape for a set of measurements after varying the a-c) the uncertainty (s.t.d.), d-e) both the uncertainty and the expected location of the measurements.

### 3.2 The probability distribution of symmetry values

Fig. 7a displays a Laue photograph ([1]) which is an interference pattern created by projecting X-ray beams onto crystals. Crystal quality is determined by evaluating the symmetry of the pattern. In this case the interesting feature is not the closest symmetric configuration, but the probability distribution of the symmetry distance values.

Consider the configuration of 2D measurements given in Fig. 4a. Each measurement  $Q_i$  is a normal probability distribution  $Q_i \sim \mathcal{N}(P_i, \Lambda_i)$ . We assume the centroid of the expectation of the measurements is at the origin. The probability distribution of the symmetry distance values of the original measurements is equivalent to the probability distribution of the location of the “average” point ( $\hat{P}_0$ ) given the folded measurements as obtained in Step 1 and Step 2 of the algorithm in Section 3.1. It is shown in Section 4.2 that this probability distribution is a  $\chi^2$  distribution of order  $n - 1$ . However, we can approximate the distribution by a gaussian distribution. Details of the derivation are given in Section 4.2.

In Fig. 7 we display distributions of the symmetry distance as obtained for the Laue photograph given in Fig. 7a. In this example we considered every dark patch as a measured point with variance proportional to the size of the patch. Thus in Fig. 7b the rectangles which are proportional in size to the corresponding dark patches of Fig. 7a, represent the standard deviation of the locations of point measurements. Note that a different analysis could be used in which the variance of the measurement location is taken as inversely proportional to the size of the dark patch.

In Fig. 8 we display distributions of the symmetry distance value for various measurements. As expected,

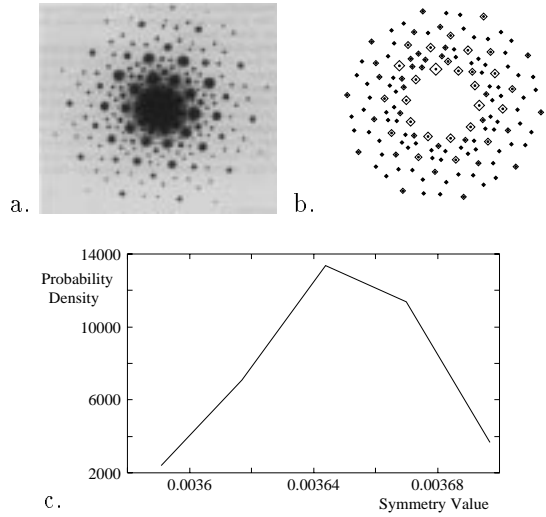


Figure 7: Probability distribution of symmetry values. a) Interference pattern of crystals. b) Probability distribution of point locations corresponding to a. c) Probability distribution of symmetry distance values with respect to  $C_{10}$ -symmetry. Expectation value = 0.003663.

the distribution of symmetry distance values becomes broader as the uncertainties (the variance of the distribution) of the measurements increase.

Medical diagnostics often use symmetry. For example cancerous tissues are quite often non symmetric and asymmetric organs may imply some abnormality or cancerous growth. Using symmetry measures these imperfect symmetries can be quantified and used to assist in medical diagnosis. A specific case is that of

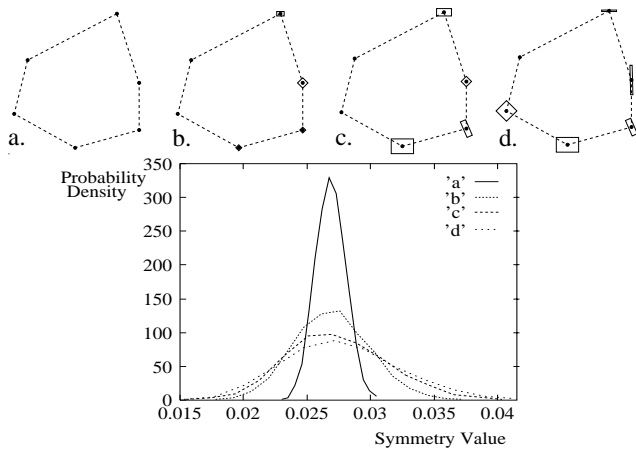


Figure 8: Probability distribution of symmetry distance values as a function of the variance of the measured points. a-d) Configurations of measured points. e) Probability distribution of symmetry distance values with respect to  $C_6$ -symmetry for the configurations a-d.

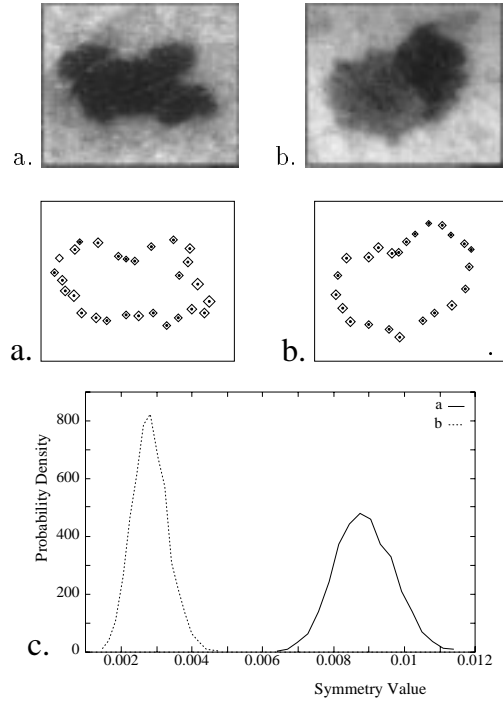


Figure 9: Probability distribution of symmetry values a-b) Two images of skin spots. c-d) Probability distribution of point locations corresponding to the skin spots of a-b respectively. e) Probability distribution of symmetry distance values with respect to mirror-symmetry for skin spots. Expectation value for skin spots a and b are 0.009013 and 0.002921 respectively.

skin cancer where the skin spot is determined to be cancerous, as a function of the “amount” of symmetry of the spot [3]. Figs 9a-b display two images of skin spots. These spots were represented by a sequence of measurements along the fuzzy contour of the spot (see Fig. 9c-d). The symmetry distribution of these sets of measurements were evaluated with respect to mirror symmetry. Notice that the skin spot of Fig. 9a has not only a higher expectation for the symmetry value but also has a broader distribution.

## 4 Mathematical derivations

### 4.1 The most probable $C_n$ -symmetric shape

In Section 3.1 we described a method of evaluating the most probable symmetric shape given a set of measurements. In this Section we derive mathematically and prove the method. For simplicity we derive the method with respect to rotational symmetry of order  $n$  ( $C_n$ -symmetry). The solution for mirror symmetry is similar.

Given  $n$  points in 2D whose positions are given as normal probability distributions:  $Q_i \sim \mathcal{N}(P_i, \Lambda_i)$ ,  $i = 0 \dots n-1$ , we find the  $C_n$ -symmetric configuration of points  $\{\hat{P}_i\}_0^{n-1}$  which is optimal under the Maximum Likelihood criterion.

Denote by  $\omega$  the center of mass of the points  $\hat{P}_i$ :  $\omega = \frac{1}{n} \sum_{i=0}^{n-1} \hat{P}_i$ . Having that  $\{\hat{P}_i\}_0^{n-1}$  are  $C_n$ -symmetric, the following is satisfied:

$$\hat{P}_i = R_i(\hat{P}_0 - \omega) + \omega \quad (1)$$

for  $i = 0 \dots n-1$  where  $R_i$  is a matrix representing a rotation of  $2\pi i/n$  radians. Given the measurements  $Q_0, \dots, Q_{n-1}$  we find the most probable  $\hat{P}_0$  and  $\omega$  by maximizing  $\text{Prob}(\{P_i\}_{i=0}^{n-1} \mid \omega, \hat{P}_0)$  under the symmetry constraints of Eq. 1.

Thus, due to the normal distribution we minimize:

$$\prod_{i=0}^{n-1} k_i \exp\left(-\frac{1}{2}(\hat{P}_i - P_i)^t \Lambda_i^{-1} (\hat{P}_i - P_i)\right)$$

where  $k_i = \frac{1}{2} \pi |\Lambda_i|^{1/2}$ . Having log being a monotonic function, we maximize:

$$\log \prod_{i=0}^{n-1} k_i \exp\left(-\frac{1}{2}(\hat{P}_i - P_i)^t \Lambda_i^{-1} (\hat{P}_i - P_i)\right)$$

Thus we find those parameters which maximize:

$$-\frac{1}{2} \sum_{i=0}^{n-1} (\hat{P}_i - P_i)^t \Lambda_i^{-1} (\hat{P}_i - P_i)$$

under the symmetry constraint of Eq. 1.

Substituting Eq. 1, taking the derivative with respect to  $\hat{P}_0$  and equating to zero we obtain:

$$\underbrace{\left(\sum_{i=0}^{n-1} R_i^t \Lambda_i^{-1} R_i\right) \hat{P}_0}_{A} + \underbrace{\sum_{i=0}^{n-1} R_i^t \Lambda_i^{-1} (I - R_i) \omega}_{B} = \underbrace{\sum_{i=0}^{n-1} R_i^t \Lambda_i^{-1} P_i}_{E} \quad (2)$$

Note that  $R_0 = I$  where  $I$  is the identity matrix.

When the derivative with respect to  $\omega$  is zero:

$$\underbrace{\left(\sum_{i=0}^{n-1} (I - R_i)^t \Lambda_i^{-1} R_i\right) \hat{P}_0}_{C} + \underbrace{\sum_{i=0}^{n-1} (I - R_i)^t \Lambda_i^{-1} (I - R_i) \omega}_{D} = \underbrace{\sum_{i=0}^{n-1} (I - R_i)^t \Lambda_i^{-1} P_i}_{F} \quad (3)$$

Notice that when all  $\Lambda_i$  are equal (i.e. all points have the same uncertainty, which is equivalent to the cases in Section 2 where point location is known with no uncertainty), Eqs. 2-3 reduce to Eqs. 3.5-3.6 in [7].

From Eq. 2 we obtain

$$\hat{P}_0 - \omega = \left(\sum_{j=0}^{n-1} R_j^t \Lambda_j^{-1} R_j\right)^{-1} \sum_{i=0}^{n-1} (R_i^t \Lambda_i^{-1} R_i) R_i^t (P_i - \omega)$$

Which gives the folding method described in Section 3.1, where  $R_i^t(P_i - \omega)$  is the location of the folded measurement (denoted  $\tilde{P}_i$  in text) and  $R_i^t \Lambda_i^{-1} R_i$  is its probability distribution (denoted  $\tilde{\Lambda}_i$  in the text). The term  $(\sum_{j=0}^{n-1} R_j^t \Lambda_j^{-1} R_j)$  is the normalization factor.

Reformulating Eqs. 2 and 3 in matrix formation we obtain:

$$\underbrace{\begin{pmatrix} A & B \\ C & D \end{pmatrix}}_U \underbrace{\begin{pmatrix} \hat{P}_0 \\ \omega \end{pmatrix}}_V = \underbrace{\begin{pmatrix} E \\ F \end{pmatrix}}_Z$$

Noting that  $U$  is symmetric we solve by inversion  $V = U^{-1}Z$  and obtain the parameters  $\omega$  and  $\hat{P}_0$ , and obtain the most probable  $C_n$ -symmetric configuration, given the measurements  $\{Q_i\}_{i=0}^{n-1}$ .

Similar to the representation in Section 2, given  $m = qn$  measurements  $\{Q_i\}_{i=0}^{m-1}$ , we consider them as  $q$  sets of  $n$  interlaced measurements:  $\{Q_{iq+j}\}_{i=0}^{n-1}$  for  $j = 0 \dots q-1$ . The derivations given above are applied to each set of  $n$  measurements separately, in order to obtain the most probable  $C_n$ -symmetric set of points  $\{\hat{P}_i\}_{i=0}^{m-1}$ . Thus the symmetry constraints that must be satisfied are:

$$\hat{P}_{iq+j} = R_i(\hat{P}_j - \omega) + \omega$$

for  $j = 0 \dots q-1$  and  $i = 0 \dots n-1$  where, again,  $R_i$  is a matrix representing a rotation of  $2\pi i/n$  radians and  $\omega$  is the centroid of all points  $\{\hat{P}_i\}_{i=0}^{m-1}$ .

As derived in Eq. 2, we obtain for  $j = 0 \dots q-1$ :

$$\underbrace{\left(\sum_{i=0}^{n-1} R_i^t \Lambda_{iq+j}^{-1} R_i\right) \hat{P}_j}_{A_j} + \underbrace{\sum_{i=0}^{n-1} R_i^t \Lambda_{iq+j}^{-1} (I - R_i) \omega}_{B_j} = \underbrace{\sum_{i=0}^{n-1} R_i^t \Lambda_{iq+j}^{-1} P_{iq+j}}_{E_j} \quad (4)$$

and equating to zero, the derivative with respect to  $\omega$ , we obtain, similar to Eq. 3:

$$\sum_{j=0}^{q-1} \underbrace{\left(\sum_{i=0}^{n-1} (I - R_i)^t \Lambda_{iq+j}^{-1} R_i\right) \hat{P}_j}_{C_j} + \underbrace{\sum_{j=0}^{q-1} \sum_{i=0}^{n-1} (I - R_i)^t \Lambda_{iq+j}^{-1} (I - R_i) \omega}_{D} = \underbrace{\sum_{j=0}^{q-1} \sum_{i=0}^{n-1} (I - R_i)^t \Lambda_{iq+j}^{-1} P_{iq+j}}_F \quad (5)$$

Rewriting Eqs. 4 and 5 in matrix formation we obtain:

$$\underbrace{\begin{pmatrix} A_0 & & & B_0 \\ & A_1 & & B_1 \\ & & \ddots & \vdots \\ & & & A_{q-1} & B_{q-1} \\ C_0 & C_1 & \cdots & C_{q-1} & D \end{pmatrix}}_U \underbrace{\begin{pmatrix} \hat{P}_0 \\ \hat{P}_1 \\ \vdots \\ \hat{P}_{q-1} \\ \omega \end{pmatrix}}_V = \underbrace{\begin{pmatrix} E_0 \\ E_1 \\ \vdots \\ E_{q-1} \\ F \end{pmatrix}}_Z$$

Noting that  $U$  is symmetric we solve by inversion  $V = U^{-1}Z$  and obtain the parameters  $\omega$  and  $\{\hat{P}_j\}_{j=0}^{q-1}$ , and obtain the most probable  $C_n$ -symmetric configuration,  $\{\hat{P}_j\}_{j=0}^{m-1}$  given the measurements  $\{Q_i\}_{i=0}^{m-1}$ .

## 4.2 Probability distribution of symmetry values

In this section we derive the probability distribution of symmetry distance values with respect to  $C_n$ -symmetry, obtained from a set of  $n$  measurements:  $Q_i \sim \mathcal{N}(P_i, \Lambda_i)$   $i = 0 \dots n-1$  (see Section 3.2).

Denote by  $X_i$  the 2-dimensional random variable having a normal distribution equal to that of  $\hat{Q}_i$  i.e.

$$\begin{aligned} E(X_i) &= R_i P_i \\ \text{Cov}(X_i) &= R_i \Lambda_i R_i^t \end{aligned}$$

where  $R_i$  denotes (as in Section 2) the rotation matrix of  $2\pi i/n$  radians.

Denote by  $Y_i$  the 2-dimensional random variable:

$$Y_i = X_i - \frac{1}{n} \sum_{j=0}^{n-1} X_j$$

in matrix notation:

$$\underbrace{\begin{pmatrix} Y_0 \\ \vdots \\ Y_{n-1} \end{pmatrix}}_Y = A \underbrace{\begin{pmatrix} X_0 \\ \vdots \\ X_{n-1} \end{pmatrix}}_X$$

or  $Y = AX$  where  $Y$  and  $X$  are of dimension  $2n$  and  $A$  is the  $2n \times 2n$  matrix:

$$A = \frac{1}{n} \begin{pmatrix} n-1 & 0 & -1 & 0 & -1 & \cdots \\ 0 & n-1 & 0 & -1 & 0 & \cdots \\ -1 & 0 & \ddots & 0 & -1 & \cdots \\ & & & \ddots & & \\ & & & & \ddots & \\ \cdots & \cdots & \cdots & \cdots & \cdots & n-1 \end{pmatrix}$$

And we have

$$E(\mathbf{X}) = \begin{pmatrix} E(X_0) \\ \vdots \\ E(X_{n-1}) \end{pmatrix} \quad \text{Cov}(\mathbf{X}) = \begin{pmatrix} \text{Cov}(X_0) & & \\ & \ddots & \\ & & \text{Cov}(X_{n-1}) \end{pmatrix}$$

$$E(\mathbf{Y}) = AE(\mathbf{X}) \quad \text{Cov}(\mathbf{Y}) = ACov(\mathbf{X})A^t$$

Given that the matrix  $ACov(\mathbf{X})A^t$ , is symmetric and positive definite, we can find a  $2n \times 2n$  matrix  $S$  diagonalizing  $\text{Cov}(\mathbf{Y})$  i.e.

$$SACov(\mathbf{X})A^tS^t = D$$

where  $D$  is a diagonal matrix (of rank  $2(n-1)$ ).

Denote by  $\mathbf{Z}$  the  $2n$ -dimensional random variable  $SAX$ .

$$\begin{aligned} E(\mathbf{Z}) &= SAE(\mathbf{X}) \\ \text{Cov}(\mathbf{Z}) &= SACov(\mathbf{X})A^tS^t = D \end{aligned}$$

Thus the random variables  $Z_i$  that compose  $\mathbf{Z}$  are independent and, being linear combinations of  $X_i$ , they are of normal distribution.

The symmetry distance, as defined in Section 2, is equivalent, in the current notations, to  $s = \mathbf{Y}^t \mathbf{Y}$ . Having  $S$  orthonormal we have

$$s = (A\mathbf{X})^t A\mathbf{X} = (SAX)^t SAX = \mathbf{Z}^t \mathbf{Z}$$

If  $\mathbf{Z}$  were a random variable of standard normal distribution, we would have  $s$  being of a  $\chi^2$  distribution of order  $2(n-1)$ . In the general case  $Z_i$  are normally distributed but not standard and  $\mathbf{Z}$  cannot be standardized globally. We approximate the distribution of  $s$  as a normal distribution with

$$\begin{aligned} E(s) &= E(\mathbf{Z})^t E(\mathbf{Z}) + \text{trace} D^t D \\ \text{Cov}(s) &= 2\text{trace}(D^t D)(D^t D) + 4E(\mathbf{Z})^t D^t D E(\mathbf{Z}) \end{aligned}$$

## 5 Conclusion

In this paper we evaluated the deviation from perfect symmetry of incomplete data. We described a method based on a continuous measure of symmetry, previously defined, for dealing with uncertain data, i.e. dealing with a configuration of measurements representing the probability distribution of point location. A direct application of this method is to quantify crystal quality by evaluating the symmetry of interference patterns obtained by projecting X-rays onto crystals. These methods can be easily extended to higher dimensions and to more complex symmetry classes.

## References

- [1] J. Auleytner. *X-Ray Methods in the Study of Defects in Single Crystals*. Pergamon Press, Warszawa, 1967.
- [2] M. Bornstein and J. Stiles-Davis. Discrimination and memory for symmetry in young children. *Developmental Psychology*, 17:82–86, 1984.
- [3] H.T. Lynch. *Skin Heredity, and Malignant Neoplasms*. Medical Examination Pub., Flushing, NY, 1972.
- [4] H. Weyl. *Symmetry*. Princeton Univ. Press, 1952.
- [5] H. Zabrodsky, S. Peleg, and D. Avnir. A measure of symmetry based on shape similarity. In *CVPR-92*, pages 703–706, Champaign, June 1992.
- [6] H. Zabrodsky, S. Peleg, and D. Avnir. Completion of occluded shapes using symmetry. In *CVPR-93*, pages 678–679, New York, June 1993.
- [7] H. Zabrodsky, S. Peleg, and D. Avnir. Symmetry as a continuous feature. *IEEE trans. PAMI*, to appear.