

# Simple comparison of atomic population and shape atomic populations distributions between two molecular structures with a coherent number of atoms

Ramon Carbó-Dorca

Received: 7 June 2012 / Accepted: 15 November 2012 / Published online: 28 November 2012  
© Springer Science+Business Media New York 2012

**Abstract** For any molecular structure and under arbitrarily varied theoretical computational levels, the concept of atomic population, initially proposed by Mulliken and now present with many variants in quantum chemical studies, can be used to obtain a set of polarized atomic charges, as it is well known. However, it has not been studied yet how atomic populations can be also the basis for constructing discrete probability distributions in the form of shape atomic population distributions. This kind of discrete molecular probability distributions can be easily used to perform numerical comparisons between various theoretical levels employed (varying computational methods and basis sets), among diverse molecular isomers or amid molecular states. Distance dissimilarity and cosine similarity indices or Shannon entropy can be employed to compare the pairs of atomic populations and at the same time to obtain new molecular descriptor parameters.

**Keywords** Atomic populations · Mulliken atomic populations · Global atomic populations · Shape populations · Comparison of atomic populations · Euclidian and Minkowski distances · Cosine similarity index · Shannon entropy · Rao distribution comparison

## 1 Introduction

In some recent studies [1–3] performed at our laboratory, about molecular density functions and electrostatic potentials under the atomic shell approximation (ASA) [4–14], it has been proven that good quality maps of these molecular quantum functions can be obtained without effort for a large variety of structures. These examples

---

Ramon Carbó-Dorca (✉)  
Institut de Química Computacional, Universitat de Girona, Girona 17071, Catalonia, Spain  
e-mail: quantumqsar@hotmail.com

permit to consider ASA functions reliable enough for quantum similarity calculations, confirming the way as they have been employed for this purpose since several years from now, see for example reference [15] for a modern application.

The present contribution is aimed to provide simple tests for studying the characteristics of atomic populations, as such molecular parameter set, initiated by Mulliken [16] long time ago, corresponds to one of the basic information sets needed to construct reliable ASA density functions.

Because of all these considerations, the development of tools to assess population differences appears of some importance. The present work tries to set up the basic concepts which can be used in order to compare molecular atomic populations.

Using this line of thought in what follows there will be defined several new molecular parametric concepts related to atomic populations. Such new theoretical ideas, like global atomic populations and shape atomic populations, have been not yet described, as far the author knows.

### 1.1 Basic definitions

Given any molecule  $M$ , the possible attached atomic populations of any conceivable origin will be denoted hereafter by the column vector:

$$|Q^M\rangle = \{Q_I^M | I = 1, \nu_M\},$$

where  $\nu_M$  is the number of atoms forming the molecule.

Independently of the computational origin of the vector elements collecting the atomic populations and previously to any other consideration, it must be taken into account a property concerning the atomic population vector, which will fulfill without exception the equation:

$$\langle\langle Q^M \rangle\rangle = \sum_{I \in M} Q_I^M = N_M, \quad (1)$$

where  $N_M$  is the number of electrons of the molecule  $M$ .

Moreover, another capital property must hold universally in this field: the positive definiteness, which has to be associated at the same time to all the elements of an arbitrarily computed atomic population vector, attached in turn to any molecule  $M$ , that is:

$$\forall M \wedge \forall I \in M : Q_I^M \in \mathbf{R}^+.$$

### 1.2 The atomic numbers vector

The two properties earlier mentioned will automatically be fulfilled, whenever one constructs another vector of the same dimension as the previous one, but now containing the set of the positive definite atomic numbers, attached like the atomic populations

to the atoms of the same molecule  $M$ :  $|Z^M\rangle = \{Z_I^M \mid I = 1, N_M\}$ . Performing a complete sum over the atomic numbers vector  $|Z^M\rangle$  it is obtained a related characteristic, like the one described in Eq. (1):

$$\langle |Z^M\rangle \rangle = \sum_{I \in M} Z_I^M = N_M. \quad (2)$$

The difference between Eqs. (1) and (2) though, consists in that one can look at the last sum, as a composite scalar product, namely:

$$N_M = \sum_{A \in M} n_A^M Z_A, \quad (3)$$

where now the sum in Eq. (3) runs over the different kinds of atoms of the molecule  $M$ . Then, necessarily the set  $\{n_A^M\}$  corresponds to the number of atoms of kind  $A$  present into molecule  $M$ . Finally, now the set  $\{Z_A\}$  contains the atomic numbers of every atomic kind represented in  $M$ .

### 1.3 Global atomic populations

In a similar trivial way, the atomic populations can be summed up choosing the same atom kind, say  $K$ , within the complete population sum (1), yielding:

$$\sum_{K \in M} \left( \sum_{I \in K} Q_{K;I}^M \right) = N_M, \quad (4)$$

where now the molecule  $M$  population set, which can be ordered as a matrix:  $\{Q_{K;I}^M \mid K = 1, n_A^M; I = 1, n_K^M\} \wedge \sum_{K \in M} n_K^M = v_M$  corresponds to the same values as defined before, but reordered employing an extra index, which indicates the atomic kind attached to a given population. The parameters  $\{n_K^M\}$  as in Eq. (3), collect the number of atoms for each different atomic kind present into the molecule  $M$ .

In this manner it is easy to see that one can define a *global* atomic population for every atomic kind:  $\{G_K^M\}$ , which can be defined as:

$$\forall K \in M: G_K^M = \sum_{I \in K} Q_{K;I}^M \rightarrow \sum_{K \in M} G_K^M = N_M. \quad (5)$$

For example, in Benzene one has only two atomic kinds namely:  $\{C,H\}$ , also:  $n_C^M = n_H^M = 6 \wedge v_M = 12 \wedge N_M = 42$ , and using Spartan program [17] under HF and a 3-21G basis set, with computed Mulliken atomic populations up to two significant figures, only two different atomic populations due to the high molecular symmetry are present:

$$Q_C = 6.24 \wedge Q_H = 0.76,$$

then the global atomic populations are readily computed, being their sum the number of electrons; the benzene situation can be resumed using the simple scheme:

$$G_C = 6 \cdot 6.24 = 37.44 \wedge G_H = 6 \cdot 0.76 = 4.56 \rightarrow G_C + G_H = 42.$$

## 2 Shape atomic populations

When observing Eq. (1) (insisting on the fact that all the properties discussed here will hold, whatever the origin of the atomic populations used), one cannot avoid the temptation to define an obvious probability-like distribution, just dividing every atomic population value by the number of electrons in the molecule:

$$\sum_{I \in M} \left( N_M^{-1} \right) Q_I^M = \sum_{I \in M} \pi_I^M = 1. \quad (6)$$

The set  $|\pi^M\rangle = \{\pi_I^M | I = 1, v_M\}$  which can be associated to *shape* atomic populations as defined and used in Eq. (6):  $\forall I = 1, v_M : \pi_I^M = \left( N_M^{-1} \right) Q_I^M$ , behave as a discrete probability distribution, with an associated cardinality equal to the number of atoms  $v_M$  present in the molecule  $M$ . Over the shape atomic populations it can be also used the equivalent reordering employed in Eq. (4), collecting atoms of the same kind and summing up reaching for a global quantity. For instance, it can be written:

$$\forall K = 1, n_A^M \in M : \pi_K^M = \sum_{I \in K} \pi_{KI}^M.$$

Using the data of the previous benzene HF 3-21G calculation and the attached global atom populations as computed beforehand, one can easily obtain in this case the global shape atomic populations as:

$$\pi_C = 0.89 \wedge \pi_H = 0.11.$$

Equation (2) can be transformed using the same ideas into such a probability distribution too:

$$\sum_{I \in M} \left( N_M^{-1} \right) Z_I^M = \sum_{I \in M} \theta_I^M = 1,$$

With such a definition, the set of bulk shape populations, which can be easily written as:  $|\theta^M\rangle = \{\theta_I^M = \left( N_M^{-1} \right) Z_I^M | I = 1, v_M\}$ , also becomes a probability distribution. Of course, this distribution can be rearranged by classes, composed in turn with as many equal elements, as the number of atomic kinds present into the molecule.

In the benzene case there will be 6 bulk shape atomic population numbers with the same value:  $\theta_C = 1/7$ , corresponding to the carbon atoms and six more:  $\theta_H = 1/42$ , which can be associated to the hydrogen atoms.

### 3 Comparison of two shape populations

Once realized the easy transformation of atomic populations into shape atomic population distributions, it can be envisaged the possibility to compare two different molecular calculations possessing a coherent number of atoms. In a first instance, they can be compared with the usual mathematical tools furnished by simple geometrical considerations.

There are already well-known dissimilarity and similarity indices which can be employed for this task, and have been used from the early times of quantum similarity [18], and which continue to be studied in modern times, see for example references [19–22]. In fact, the main computational possibilities can be primarily organized into distances and cosines, as it will be explained below. Other more sophisticated statistical comparison procedures will be briefly discussed at the end.

#### a. Euclidian distance-like measures

In any case, one can compare a distribution issued from an atomic population, obtained according to some criterion, with the bulk shape atomic populations; that is, a squared Euclidian distance can be set up:

$$\Delta^{(2)} = \sum_{I \in M} (\theta_I^M - \pi_I^M)^2. \quad (7)$$

The positive definite scalar  $\Delta^{(2)}$  can be used as a measure of how far can be considered shape population distributions from the non-polarized molecular frame.

#### a.1 Naïve examples

- 1) When trying to obtain the difference between the previous benzene HF 3-21G calculation and the bulk molecule distribution, which can be described by means of:  $\theta_C = \frac{6}{7} = 0.86 \wedge \theta_H = \frac{1}{7} = 0.14$ . As one can easily obtain in this case the global atomic shape populations as shown before, then the squared distance corresponds to:  $\Delta^{(2)} = 1.8 \times 10^{-3}$ .
- 2) Taking the benzene molecule as an example again, within a HF calculation but now under an extended 6311+G\*\* basis set, one obtains Mulliken populations rounded up to two decimals, with values:  $Q_C = 6.13 \wedge Q_H = 0.87$ , which imply global populations in benzene like:  $G_C = 36.8 \wedge G_H = 4.98$ . These values yield a shape atomic population probability distribution, which can be written as:  $\pi_C = 0.88 \wedge \pi_H = 0.12$ . This distribution can be compared with the bulk distribution, as described before in the first example. The squared distance (7), in this large basis set benzene computation can be easily obtained as a quite small quantity  $\Delta^{(2)} = 8 \times 10^{-4}$ , a value coherent with the approach to the null polarization of benzene one can apparently expect as the basis set grows.
- 3) The difference obtained between both basis sets and the bulk molecular atomic numbers is small, but larger in the smaller basis set than in the larger one.
- 4) The previously obtained figures can be enhanced taking the square root of expression (7) and transforming them into Euclidian distances. For the 3-21G case it is obtained:  $\Delta = 0.042$  and for the large basis set one finds:  $\Delta = 0.028$ .

The simple mathematical structure described so far for comparison, can undoubtedly generate an order on the quite large different kinds of atomic populations sets, which can be computed within the same molecule.

#### a.2 Using distances between two different shape populations

For example: populations defined according to two criteria over the same molecule can be compared, when obtained with the same basis set, or alternatively, values of the same kind of shape populations but obtained in diverse basis sets can be also compared.

Even the effect of population distortions provoked by bond rotations, stretching and bending, or electronic excitations, even ionization (because shape populations sum the unit), diverse isomers (for instance, in absence of chiral fields two enantiomers have the same shape populations and the measure  $\Delta^{(2)}$  will appear the same when exchanging R and S isomers distributions) can be assessed and compared with some basic or reference shape population in this distance-like way.

In any case Eq. (7), can be also generalized and written like:

$$\Delta_{ab}^{(2)} = \sum_{I \in M} \left( a \pi_I^M - b \pi_I^M \right)^2 \quad (8)$$

where the supraindices written on the left side of the shape populations, denote different shape populations for the same molecule, which can be associated to different theoretical levels or to diverse conformers or isomers, as suggested above.

#### a.3 Trivial examples

For example, the previously computed benzene 3-21G Mulliken shape population can be compared with the large basis set one. The squared distance between the two HF basis set calculation levels corresponds to an irrelevant quadratic difference:  $\Delta_{ab}^{(2)} = 2 \times 10^{-4}$ , associated to an Euclidian distance of:  $\Delta_{ab} = 0.014$ .

As another possible example, it will be instructive to compare the large basis set HF calculation with the same basis set within B3LYP DFT calculation, which yields Mulliken populations like:  $Q_C = 6.13 \wedge Q_H = 0.87$ . In this case the distance (8) will be easily calculated when compared with the HF calculation of the same accuracy yielding:  $\Delta_{ab}^{(2)} = \Delta_{ab} = 0$ . Thus, apparently, from this point of view HF and DFT at such computation level provide the same population information.

#### a.4 Comparisons between borazine and benzene

In order to provide another simple example, employing the previously described techniques, one can try to evaluate the numerical differences between two isoelectronic molecules possessing coherent number of atoms, like benzene and borazine, for instance. The benzene molecule has been already described from the point of view of shape atomic populations; the borazine, within the same computational characteristics (HF under 3-21G basis set) provide four atomic kinds with different Mulliken populations:

$$Q_B = 4.25 \wedge Q_N = 8.01 \wedge Q_{H(N)} = 0.66 \wedge Q_{H(B)} = 1.07.$$

The distance-like measure between borazine and benzene can be constructed as:  $\Delta_{ab}^{(2)} = 0.037$ .

Comparison of borazine with bulk borazine shape distribution will show a smaller distance, comparable to the ones obtained in the benzene case:  $\Delta_{0a}^{(2)} = 9 \times 10^{-3} \rightarrow \Delta_{0a} = 0,095$ .

In these figures one can grasp the whole degree of polarization and the difference between both isoelectronic molecules, which is larger than the distances of both structures to their bulk shape populations.

#### a.5 Borazine in a 6311+G\*\* basis set

Borazine under a HF calculation with the same large basis already set used in benzene too, provides the following Mulliken atomic populations:

$$Q_B = 4.59 \wedge Q_N = 7.66 \wedge Q_{H(N)} = 0.70 \wedge Q_{H(B)} = 1.05$$

which deliver global atomic populations like:

$$G_B = 13.77 \wedge G_N = 22.98 \wedge G_{H(N)} = 2.1 \wedge G_{H(B)} = 3.15$$

and those permit to compute the shape atomic population distribution as:

$$\pi_B = 0.33 \wedge \pi_N = 0.55 \wedge \pi_{H(N)} = 0.05 \wedge \pi_{H(B)} = 0.075.$$

It can be instructive to compare this distribution with the bulk one, which yields:  $\Delta_{0a}^{(2)} = 3.5 \times 10^{-3} \rightarrow \Delta = 0.059$

This seems to indicate that, at the same time as has been obtained with benzene, atomic populations in larger basis sets tend to be nearest to the bulk non polarized distribution than the smaller basis set results.

Comparison with the HF 3-21G basis set borazine calculation provides:

$$\Delta_{ab}^{(2)} = 1.23 \times 10^{-3} \rightarrow \Delta_{ab} = 0,035$$

Indicating not only the obvious result that borazine is more polarized than benzene, but also that the difference between both small and large basis sets is also larger than in benzene.

#### b. Minkowski-like distances

Instead of using Euclidian distances one can also employ Minkowski absolute value sums. Then, Eqs. (7) and (8) will have their Minkowski counterparts redefined as:

$$M = \sum_{I \in M} \left| \theta_I^M - \pi_I^M \right|$$

Which for the benzene molecule yields for the 3-21G calculation:  $M = 0.069$  and for the large basis set:  $M = 0.037$ .

In order to compare two different calculations one can define the Minkowski counterpart of Eq. (8) as:

$$M_{ab} = \sum_{I \in M} \left| {}^a \pi_I^M - {}^b \pi_I^M \right|.$$

For example, comparing both benzene HF calculations at the two basis set levels one obtains:  $M_{ab} = 0.031$ . Also the borazine pair of calculations under the same framework, compared as in benzene, provides a greater value:  $M_{ab} = 0.054$ .

It can be seen how the ordering encountered in the previous Euclidian example is preserved in the Minkowski framework. Of course, the comparison of the HF and DFT calculation in the same basis set 6311+G\*\* in benzene will yield a null Minkowski distance too. In borazine DFT Mulliken atomic populations yield:

$$Q_B = 4.71 \wedge Q_N = 7.52 \wedge Q_{H(N)} = 0.723 \wedge Q_{H(B)} = 1.050.$$

Comparison of HF and DFT B3LYB under a large basis set by means of Minkowski distance yields:  $M_{ab} = 0.02$ , a relevant difference.

### c. Cosine-like similarity measures

The comparison possibilities associated to distance-like measures as studied above are appealing, but there is not only such a possible comparison type. Other kinds of measures can be designed, among others, a cosine-like similarity measure, which can be obviously related to the Carbó index of QSM [18]. It is easy to write the following shape atomic population vector definitions:

$$\left| {}^a \pi^M \right\rangle = \left\{ {}^a \pi_I^M \right\} \wedge \left| {}^b \pi^M \right\rangle = \left\{ {}^b \pi_I^M \right\}.$$

In order to construct this difference measure in an elegant way, the needed norms and scalar product of these shape atomic population vectors can be written in the following form:

$$\left\langle {}^a \pi^M \left| {}^a \pi^M \right\rangle = \chi_{aa} \wedge \left\langle {}^b \pi^M \left| {}^b \pi^M \right\rangle = \chi_{bb} \wedge \left\langle {}^a \pi^M \left| {}^b \pi^M \right\rangle = \chi_{ab} = \chi_{ba} = \left\langle {}^b \pi^M \left| {}^a \pi^M \right\rangle.$$

Therefore it is straightforward to define the cosine-like measure:

$$r_{ab} = \frac{\chi_{ab}}{(\chi_{aa} \chi_{bb})^{\frac{1}{2}}} \in [0, 1],$$

or its trivial squared value. For example, comparison of the two benzene HF calculations under the different basis sets yields:

$$\chi_{aa} = 0.80 \wedge \chi_{bb} = 0.79 \wedge \chi_{ab} = 0.78 \rightarrow r_{ab}^2 = 0.96$$

Consequently, in this manner one can see that it exists at least two complementary ways to assess the differences between shape populations, defined upon a given molecular structure and the possible isomers, or between molecules of the same number of atoms.



### c.1 A benzene and borazine example

In this example it has been employed the cosine-like similarity measure invariant property upon homothetic scaling of all the involved vectors, see for example reference [22] for more details. Due to this property it is irrelevant in the cosine case to use directly atomic populations or alternatively shape atomic populations.

For instance, the cosine-like measure between both benzene–borazine molecules can be easily obtained as:

$$\langle a|a \rangle = 237.10 \wedge \langle b|b \rangle = 251.41 \wedge \langle a|b \rangle = 233.45 \rightarrow r_{ab} = 0.956.$$

Both molecules can be also compared in turn with the bulk shape populations providing, for benzene a cosine-like index:

$$\langle 0a|0a \rangle = 222 \wedge \langle 0a|a \rangle = 229.2 \rightarrow r_{0aa} = 0.999;$$

while, borazine when compared with the bulk distribution it is obtained:

$$\langle 0b|0b \rangle = 228 \wedge \langle 0b|b \rangle = 237.15 \rightarrow r_{0bb} = 0.991.$$

These simple examples can be sufficient to grasp the power of such cosine-like comparison procedures. One can conclude that benzene is almost equal to its bulk shape distribution, while borazine is slightly different, but both benzene–borazine shape distributions are quite different between them.

## 4 Shannon information content and other distance possibilities

This kind of proposed study will be not complete, whenever skipping even a few words about the possibility to use some other techniques for comparison purposes, and index generation by means of atomic populations.

It can be also obtained some numerical insight, for instance, computing the information content of a shape atomic population probability distribution. Indeed, any vector attached to a molecular distribution of this kind, like:  ${}^a\pi^M = \{{}^a\pi_I^M\}$ , can be handled into the well-known Shannon entropy formula:

$$S = - \sum_{I \in M} {}^a\pi_I^M \log_2 ({}^a\pi_I^M).$$

For the two benzene HF computational levels studied so far, one can easily write, using natural logarithms, thus giving in nets the entropy:

- 1) 3-21G:  $S = 0.35$
- 2) 6311+G\*\*: $S = 0.37$
- 3) Bulk:  $S = 0.41$

While for borazine one can write:

- 1) 3-21G:  $S = 1.02$

- 2)  $6311+G^{**}$ :  $S = 1.04$
- 3) Bulk:  $S = 0.99$

It is interesting to note here that in benzene the bulk distribution appears as an upper limit value of the entropy with respect to basis set variation. In borazine does not occur the same, the bulk entropy value looks like a lower limit, which might be reached from the large basis set down to the smaller one.

At the same time, comparisons involving more sophisticated distance definitions, precisely presented several years ago [23–26] to compare probability distributions. The generalization due to Rao can be also employed [23], but introducing them here will destroy the simplicity sought for this first approach to shape atomic population definitions and comparisons.

## 5 Conclusions

In molecules possessing a coherent number of atoms, upon defining shape atomic populations as some kind of discrete probability distributions, it is a simple matter to describe several possibilities under which the diverse atomic populations can be compared. Distances, cosines, Shannon entropy and Rao probability distribution distances can be easily employed for this purpose. In all the studied cases, not only the defined measures can be used for comparison, but as a source of new molecular indices as well. The interesting thing above all consists in that, in general, the population comparisons here described can be done without any great computing effort.

**Acknowledgment** This work was elaborated and finished during the author's stage at RIKEN, Kobe, Japan. The hospitality of prof. K. Hirao is highly appreciated.

## References

1. E. Besalú, R. Carbó-Dorca, J. Chem. Theor. Comp. **8**, 854–861 (2012)
2. R. Carbó-Dorca, E. Besalú, J. Math. Chem. **50**, 981–988 (2012)
3. E. Besalú, R. Carbó-Dorca, Softened Electrostatic Molecular Potentials, IQC Technical, Report TR-2012-3
4. R. Carbó-Dorca, Density Functions and Generating Wave Functions Reviews of Modern Quantum Chemistry. **15**, pp. 401–412 (2002) in: *A Celebration of the Contributions of Robert G. Parr*, vol I, ed. by K. Sen.
5. P. Constans, R. Carbó, J. Chem. Inf. Comp. Sci. **35**, 1046–1053 (1995)
6. P. Constans, X. Fradera, Ll. Amat, R. Carbó, Quantum Molecular Similarity Measures (QMSM) and the Atomic Shell Approximation (ASA). in *Proceedings of the 2nd Girona Seminar on Molecular Similarity. July 1995. Advances in Molecular Similarity*. (JAI PRESS INC. Greenwich (Conn.) 1996), vol. 1, p. 187
7. Ll Amat, R. Carbó-Dorca, J. Comput. Chem. **18**, 2023–2039 (1997)
8. Ll Amat, R. Carbó-Dorca, J. Comput. Chem. **20**, 911–920 (1999)
9. Ll Amat, R. Carbó-Dorca, J. Chem. Inf. Comput. Sci. **40**, 1188–1198 (2000)
10. R. Carbó-Dorca, E. Besalú, X. Gironès, Adv. Quant. Chem. **38**, 3–63 (2000)
11. X. Gironès, R. Carbó-Dorca, P.G. Mezey, J. Mol. Graph. Mod. **19**, 343–348 (2001)
12. R. Carbó-Dorca, Ll. Amat, E. Besalú, X. Gironès, D. Robert, in: *Mathematical and Computational Chemistry: Fundamentals of Molecular Similarity. Quantum Molecular Similarity: Theory and applications to the evaluation of molecular properties, biological activity and toxicity* (Kluwer Academic/Plenum Publishers), pp. 187–320 (2001)

13. X. Gironés, Ll Amat, R. Carbó-Dorca, *J. Chem. Inf. Comput. Sci.* **42**, 847–852 (2002)
14. E. Besalú, R. Carbó-Dorca, *J. Math. Chem.* **49**, 1769–1784 (2011)
15. R. Carbó-Dorca, E. Besalú, L.D. Mercado, *J. Comp. Chem.* **32**, 582–599 (2011)
16. R.S. Mulliken, *J. Chem. Phys.* **23**, 1833–2343 (1955)
17. Spartan '10, Version 1.1.0, (Wavefunction Inc., Irvine, CA, 2011)
18. R. Carbó, L. Leyda, M. Arnau, *Int. J. Quant. Chem.* **17**, 1185–1189 (1980)
19. R. Carbó-Dorca, *J. Math. Chem.* **49**, 2109–2115 (2011)
20. R. Carbó-Dorca, E. Besalú, *J. Math. Chem.* **50**, 210–219 (2012)
21. R. Carbó-Dorca, *J. Math. Chem.* **50**, 734–740 (2012)
22. P. Bultinck, R. Carbó-Dorca, *J. Math. Chem.* **36**, 191–200 (2004)
23. J. Burbea, R. Rao, *IEEE Trans. Inf. Theor.* **28**, 489–495 (1982)
24. I.J. Good, *Ann. Math. Stat.* **34**, 911–934 (1963)
25. H. Jeffreys, *Proc. R. Soc. Lond. A* **186**, 453–461 (1946)
26. S. Kullback, R.A. Leibler, *Ann. Math. Stat.* **22**, 79–86 (1951)