

# Charge-conserving electron density averaging for a set of nuclear configurations

Paul G. Mezey

Received: 15 January 2006 / Accepted: 15 February 2006 / Published online: 20 February 2008  
© Springer Science+Business Media, LLC 2008

**Abstract** The Löwdin-Inverse Löwdin Transformation (LILT) technique combined with the Diophantine Density Matrix Purification (DDMP) method is suggested for a charge-preserving density averaging approach for a range of nuclear configurations.

**Keywords** Electron density averaging · Charge invariance · Approximate density matrices · Idempotency · Range of nuclear configurations

## 1 Introduction

Information on molecular properties can be deduced from molecular electron densities, and the Hohenberg-Kohn theorem [1] with its various consequences and extensions [2–6] provide the tools for the theoretical background for such analyses. Density functional methods [7–12], as well as more conventional molecular orbital methods [13–21] form the basis of contemporary modeling approaches suitable to access this information.

Many molecular properties are actually associated not with a single nuclear arrangement, but with a whole range of nuclear configurations. It is often useful to associate

---

P. G. Mezey (✉)

Scientific Modeling and Simulation Laboratory (SMSL), Department of Chemistry and Department of Physics and Physical Oceanography, Memorial University of Newfoundland, St. John's, NL, Canada A1B 3X7  
e-mail: pmezey@mun.ca

P. G. Mezey

Mathematics Department, Institute J.A. Dieudonne, University of Nice Sophia Antipolis, Nice, France

P. G. Mezey

Institute for Advanced Study, Collegium Budapest, Szentháromság u. 2, 1014 Budapest, Hungary  
e-mail: paul.mezey@gmail.com

such ranges of nuclear configurations with domains of potential energy hypersurfaces, for example, with formal catchment regions representing chemical species in the potential surface model, and their various subdomains [22–33]. Whereas nuclear arrangements associated with minimum energy conformations are commonly used as reference when interpreting comparisons between experimental results and the results of theoretical, computational modeling studies, nevertheless, in reality an entire family of nuclear arrangements contributes to the experimentally observed molecular properties. In the extreme case of single molecule experiments, the quantum mechanical nature of molecules implies that models of single, geometrically defined conformations cannot provide proper interpretations, and even in semi-classical representations, the time scale of observation is usually such that necessarily a whole range of semi-classical nuclear arrangements must be considered.

As a consequence, not individual conformations but ranges of molecular electron density functions belonging to a family of nuclear arrangements, and various properties of such ranges are of importance. One of the simplest tools for studying the family of electron densities in such conformational ranges is the statistical approach, generating various statistical moments for the distribution. The most important moment of this distribution is the average electron density for the range.

In earlier studies various approaches have been described for the computation of approximations for average electron densities [34–38]. One of the main problems relates to the fact that simple direct averaging for a finite set of nuclear arrangements may lead to fundamentally wrong results, where the number of local maxima of electron densities might become a multiple of the actual density maxima present in the real molecule, simply as a consequence of averaging functions having local maxima at different locations (corresponding to the differing locations of nuclei in the family of conformations considered). Even in the case of averaging densities of just two different nuclear conformations, the relative nuclear locations differ, and it is likely that a simple, direct averaging leads to twice the number of local density maxima. Whereas this problem appears on a different level if not a finite number but a continuum of nuclear arrangements is used, nevertheless, for proper averaging additional steps are needed [34–38].

One such approach [34,35] is based on the Löwdin-Inverse Löwdin Transformation (LILT) technique [39–42], a transformation that is related to an efficient, analogous tool applied in quantum crystallography [43], where the fuzzy density fragment approach is likely to provide advantages [44]. In the earlier implementation of the LILT approach, the charge conservation problem has been addressed at the level of transformations of individual electron densities to a common nuclear framework corresponding to that of the average density [34,35]. In that approach, the idempotency of each transformed density matrix was ensured separately at the new geometry, and the actual averaging was carried out for the corresponding transformed electron densities. Although in principle this approach also leads to charge conservation, since for each transformed density the total charge is conserved due to the actual way individual idempotency of each of the transformed density matrices is achieved, nevertheless, it appears useful to ensure that a single density matrix is generated that corresponds to the average electron density, and the charge conservation condition is ensured by enforcing the idempotency of this new density matrix. For this latter purpose the recently proposed diophantine density matrix purification method [45] is suggested.

In addition, by having access to this new “averaged” density matrix, some average properties associated with the average density can also be easily computed, in addition to just the actual averaged electron density. This allows for a full exploitation of the advantages of averaging. For this reason, in this contribution an alternative approach, one based on the generation of a density matrix associated with the average density is described. It is expected that by the use of statistical approaches to electron density representations, most notably, by electron density averaging (and possibly by using higher statistical moments such as standard deviation, skewness, and kurtosis), some of the earlier approaches to the analysis of macromolecular electron densities [46–67], molecular shape analysis methods [68–77], originally employed for shape analysis of potential surfaces [33], and more general molecular similarity analysis techniques [78–88] may find extensions and additional applications.

In the discussion below, we shall follow some of the derivation and the notations of the previous approach [34,35] where averaging has been carried out directly for densities.

## 2 Weighted averaging followed by “purification” of transformed density matrices

Consider a set  $K$  of nuclear geometries  $K_i$  within a conformational domain of a molecule:

$$K = \{K_1, K_2, \dots, K_i, \dots, K_m\}. \quad (1)$$

These nuclear geometries  $K_i$  can be considered as elements of a  $(3N-6)$ -dimensional nuclear configuration space  $M$  [24,33] defined for a stoichiometric family of molecules of the given  $N$  nuclei of the molecule. In contrast to the  $3N$ -dimensional mass-weighted Euclidean configuration space  ${}^{3N}E$ , space  $M$  is that of the internal configurations, also called the “reduced nuclear configuration space”, where the Euclidean space redundancies of rigid translations and rigid rotations are eliminated. In fact,  $M$  is the quotient space of the Euclidean configuration space  ${}^{3N}E$ , by the equivalence relation generated by rigid translations and rigid rotations. Note that  $M$  is a metric space with a well-defined metric, a distance function  $d(K_i, K_j)$ , but  $M$  is not a vector space, in fact, no vector space representation is possible for any complete set of internal coordinates of any diatomic or larger molecule [24,33].

The weighted average conformation  $K_{av}$  is taken as the formal average conformation of configurations  $K_1, K_2, \dots, K_i, \dots, K_m$  that minimizes the weighted average configuration-space distance between  $K_{av}$  and the “reference conformations”  $K_1, K_2, \dots, K_i, \dots, K_m$ , that is, as the conformation that minimizes the following sum:

$$s^2(K_{av}, K) = \sum_{i=1,m} w_i d^2(K_{av}, K_i). \quad (2)$$

Here the quantities  $w_i$  are relative weights reflecting the relevance of the individual conformations  $K_i$ , for example, these weights can be chosen as

$$w_i = \exp(-\beta E_i) / \sum_{j=1,m} \exp(-\beta E_j), \quad (3)$$

where

$$\beta = 1/kT \quad (4)$$

for a suitable chosen “temperature”  $T$ , or some alternative quantity that serves as a suitable tool for energy-averaging.

Evidently, for the choice above, the positive weight factors  $w_i > 0$  satisfy the convexity constraint

$$\sum_{i=0,m} w_i = 1. \quad (5)$$

Note that the search for the average conformation involves a multidimensional optimization, the minimization of the quantity  $s^2(K_{av}, K)$  in Eq. 2, and simpler alternative “quasi-average” and “biased average” constructions have also been explored [34–36]. Nevertheless, for simplicity of discussions, here we shall focus on the choice given by Eq. 2.

Consider some  ${}^3N_E$ -space representations

$$\mathbf{x}^{(av)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(m)} \quad (6)$$

of internal configurations  $K_{av}, K_1, K_2, \dots, K_i, \dots, K_m$ . Note that these representations  $\mathbf{x}^{(av)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(i)}, \dots, \mathbf{x}^{(m)}$  are not unique, due to the fact that rigid translations and rotations yield the same internal configurations, nevertheless, for the discussions below this redundancy does not lead to any ambiguity.

We shall assume that the same set of basis functions are used for a standard HF MO computation, attached to the appropriate set of nuclei for each nuclear configuration  $K_1, K_2, \dots, K_m$ . By referring to the Euclidean space representations  $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(m)}$ , the individual AO basis functions  $\varphi_k(\mathbf{r}, \mathbf{x}^{(i)})$  with centers at the nuclear locations corresponding to the nuclear configuration  $\mathbf{x}^{(i)}$  are regarded as components of a formal vector  $\varphi(\mathbf{x}^{(i)})$ . For each set of AO's, the overlap matrix corresponding to the nuclear configuration  $K_i$  is denoted by  $\mathbf{S}(\mathbf{x}^{(i)})$ , and the corresponding first-order density matrix is denoted by  $\mathbf{P}(\mathbf{x}^{(i)})$ .

Our first task is to convert the individual density matrices

$$\mathbf{P}(\mathbf{x}^{(1)}), \mathbf{P}(\mathbf{x}^{(2)}), \dots, \mathbf{P}(\mathbf{x}^{(i)}), \dots, \mathbf{P}(\mathbf{x}^{(m)}), \quad (7)$$

obtained for the various configurations  $K_1, K_2, \dots, K_i, \dots, K_m$  of set  $K$  to a set of approximate, but idempotent density matrices for the average nuclear configuration  $K_{av}$ .

$$\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(1)}]), \mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(2)}]), \dots, \mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}]), \dots, \mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(m)}]). \quad (8)$$

Note that all these approximate density matrices refer to the same nuclear geometry  $\mathbf{x}^{(av)}$ , nevertheless, these matrices are generally different, since they originate from different nuclear arrangements.

For the purpose of generating these transformed density matrices  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}])$ , we shall use the Löwdin transform-inverse Löwdin transform (LILT) method.

The well-known Löwdin transform

$$\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2} \tag{9}$$

of a density matrix  $\mathbf{P}(\mathbf{x}^{(i)})$ , defined in terms of the AO basis set  $\varphi(\mathbf{x}^{(i)})$ , is generated by pre- and post-multiplication by the matrix  $\mathbf{S}(\mathbf{x}^{(i)})^{1/2}$ . Since the overlap matrix  $\mathbf{S}(\mathbf{x}^{(i)})$  is always positive definite, the formal “square-root” matrix  $\mathbf{S}(\mathbf{x}^{(i)})^{1/2}$  always exists and can be computed by standard techniques. The conceptually simplest technique involves the solution of the eigenvalue problem of  $\mathbf{S}(\mathbf{x}^{(i)})$ , followed by the replacement of the eigenvalues with their square roots, and completed by pre- and post multiplications with the orthogonal matrix of eigenvectors.

From the overlap-weighted idempotency property of density matrices,

$$\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})\mathbf{P}(\mathbf{x}^{(i)}) = \mathbf{P}(\mathbf{x}^{(i)}), \tag{10}$$

it follows that, the Löwdin transform  $\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}$  of the density matrix  $\mathbf{P}(\mathbf{x}^{(i)})$  is idempotent with respect to ordinary matrix multiplication:

$$\begin{aligned} &\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2} \\ &= \mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}. \end{aligned} \tag{11}$$

The essence of the Löwdin transform-inverse Löwdin transform (LILT) method is the observation that the AO basis reference of a density matrix can be approximately “shifted” to an AO basis (of the same type) at a new set of atomic coordinates using an interplay between overlap matrices at the two nuclear configurations.

We shall assume that the overlap matrix  $\mathbf{S}(\mathbf{x}^{(av)})$  is determined for the set of new AO basis functions  $\varphi(\mathbf{x}^{(av)})$ , taken simply as the set  $\varphi(\mathbf{x}^{(i)})$  of atomic orbitals placed to the new nuclear location  $\mathbf{x}^{(av)}$ . For each index  $i$ , an inverse Löwdin transform of the idempotent matrix  $\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}$  can be generated by the new overlap matrix  $\mathbf{S}(\mathbf{x}^{(av)})$ . This provides an approximate density matrix  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}])$  for the average nuclear configuration  $\mathbf{x}^{(av)}$ , and can be written as

$$\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}]) = \mathbf{S}(\mathbf{x}^{(av)})^{-1/2}\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{S}(\mathbf{x}^{(av)})^{-1/2}. \tag{12}$$

For each index  $i$ , this new approximate density matrix  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}])$  is idempotent with respect to multiplication involving the actual overlap matrix  $\mathbf{S}(\mathbf{x}^{(av)})$ . Indeed,

$$\begin{aligned} &\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}])\mathbf{S}(\mathbf{x}^{(av)})\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}]) \\ &= \mathbf{S}(\mathbf{x}^{(av)})^{-1/2}\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2} \\ &\quad \times \mathbf{S}(\mathbf{x}^{(av)})^{-1/2}\mathbf{S}(\mathbf{x}^{(av)})\mathbf{S}(\mathbf{x}^{(av)})^{-1/2}\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{S}(\mathbf{x}^{(av)})^{-1/2} \\ &= \mathbf{S}(\mathbf{x}^{(av)})^{-1/2}\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{S}(\mathbf{x}^{(av)})^{-1/2} \\ &= \mathbf{S}(\mathbf{x}^{(av)})^{-1/2}\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{S}(\mathbf{x}^{(av)})^{-1/2} \\ &= \mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}]), \end{aligned} \tag{13}$$

that is,

$$\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}])\mathbf{S}(\mathbf{x}^{(av)})\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}]) = \mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}]), \quad (14)$$

where the facts that

$$\mathbf{S}(\mathbf{x}^{(av)})^{-1/2}\mathbf{S}(\mathbf{x}^{(av)})\mathbf{S}(\mathbf{x}^{(av)})^{-1/2} = \mathbf{I}, \quad (15)$$

the unit matrix, and the idempotency of  $\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}$ ,

$$\begin{aligned} & \mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2} \\ & = \mathbf{S}(\mathbf{x}^{(i)})^{1/2}\mathbf{P}(\mathbf{x}^{(i)})\mathbf{S}(\mathbf{x}^{(i)})^{1/2} \end{aligned} \quad (16)$$

have been exploited.

In earlier studies [34–36] the averaging has focused on electron densities, and individual densities have been used for various conformations to generate an approximate electron density for some intermediate location in configuration space. Whereas those approaches have their advantages, notably in the fact that the averaging actually occurs for a physically observable property, electron density, nevertheless, the assurances that the computed electron density fulfills some necessary conditions (e.g., that it can be represented by a density matrix that fulfills the idempotency condition) were not easy to achieve.

For this reason in this study we follow a different approach: first an approximate density matrix  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(av)}])$  is generated by collecting the “weighted” information from the already transformed density matrices  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}])$ , and then the idempotency condition is directly enforced by a density matrix “purification” method, leading to an improved approximate density matrix  $\mathbf{P}(\mathbf{x}^{(av)})$  at the averaged nuclear configuration. Using this idempotent density matrix  $\mathbf{P}(\mathbf{x}^{(av)})$  a new approximate electron density  $\rho(\mathbf{x}^{(av)})$  (and other properties) can be computed for the nuclear configuration  $\mathbf{K}_{av}$ .

The approximate density matrix  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(av)}])$  is obtained by the analogous (or possibly a different) weighted averaging as that used for the nuclear configurations:

$$\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(av)}]) = \sum_{i=1,m} w_i \mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(i)}]). \quad (17)$$

Note that the trace of each of these matrices occurring in the summation is properly reflecting the electron count, since each individual density matrix is idempotent with respect to the actual  $\mathbf{S}(\mathbf{x}^{(av)})$  overlap matrix (Eq. 14), furthermore, their convex combination in Eq. 17 also preserves this property. Nevertheless, their average may deviate from exact idempotency. For this reason, a density matrix purification procedure is recommended.

It is useful to recall that in the transformations leading to the averaging, the eigenvalues (and using most methods, the eigenvectors) of the overlap matrix  $\mathbf{S}(\mathbf{x}^{(av)})$  have been computed, and eigenvalue methods are necessarily required for the approach. Consequently, the one-step diophantine method (DDMP) of density matrix purification [45] is especially suitable for the actual task at hand.

When applying this approach, one takes the eigenvalue equation for the corresponding approximate density matrix  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(av)}])$ :

$$\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(av)}])^{(avav)} \mathbf{L} = {}^{(avav)} \mathbf{L} {}^{(avav)} \mathbf{A}, \quad (18)$$

where  ${}^{(avav)} \mathbf{L}$  is the matrix of the eigenvectors and  ${}^{(avav)} \mathbf{A}$  is the diagonal matrix of eigenvalues:

$${}^{(avav)} \mathbf{A}_{ij} = {}^{(avav)} a_i \delta_{ij}. \quad (19)$$

Since a typical density matrix  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(av)}])$  is approximate, the  ${}^{(avav)} a_i$  eigenvalues are almost always different from both 0 and 1, furthermore, the density matrix  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(av)}])$  is almost always non-degenerate.

It has been shown in ref. [45] that if the McWeeny density matrix purification method [20] is applied to any approximate density matrix  $\mathbf{P}^{(appr)}$ , then the final, converged (idempotent) density matrix will have the same eigenvectors  $\mathbf{L}$  as those of the original matrix, and only the eigenvalues change during that purification process. This recognition, and the fact that the eigenvalues of the final, converged (idempotent) density matrix must be either 1 or zero, have led to a one-step Diophantine matrix solution, simply by replacing the eigenvalues by the value 1 or 0, whichever is closer to the initial eigenvalue, and use the eigenvector matrix  $\mathbf{L}$  to obtain a matrix equal to the final converged matrix of the McWeeny iterative process.

In our actual case, one obtains the Diophantine integer matrix  ${}^{(av)} \mathbf{A}$  of elements  ${}^{(av)} \mathbf{A}_{ij}$  defined as

$${}^{(av)} \mathbf{A}_{ij} = \text{ent}(0.5 + {}^{(avav)} \mathbf{A}_{ij}), \quad (20)$$

where  $\text{ent}$  is the integer part (entier) function. Using this Diophantine matrix  ${}^{(av)} \mathbf{A}$  of integers, the actual, idempotent density matrix  $\mathbf{P}(\mathbf{x}^{(av)})$  is obtained in a single step as

$$\mathbf{P}(\mathbf{x}^{(av)}) = {}^{(avav)} \mathbf{L} {}^{(av)} \mathbf{A} {}^{(avav)} \mathbf{L}^{-1}. \quad (21)$$

This entire DDMP “purification” process involves only a single additional matrix diagonalization, that of the initial approximate density matrix  $\mathbf{P}(\mathbf{x}^{(av)}, [\mathbf{x}^{(av)}])$  and the matrix multiplication of Eq. 21.

Note that, in the matrix multiplication step (21), the central matrix is not only diagonal, but it also corresponds to identity in one subspace, and to the zero matrix in the complementing subspace, that reduces the complexity of the associated computational task.

Using the averaged density matrix  $\mathbf{P}(\mathbf{x}^{(av)})$ , various “averaged” molecular properties can be computed for the conformational range represented by the family  $K$ , for example, an approximate electron density  $\rho_{\text{appr}}(\mathbf{r}, \mathbf{x}^{(av)})$  can be computed for the average nuclear geometry  $\mathbf{x}^{(av)}$ :

$$\rho_{\text{appr}}(\mathbf{r}, \mathbf{x}^{(av)}) = \sum_{k=1, n} \sum_{j=1, n} P_{kj}(\mathbf{x}^{(av)}) \varphi_k(\mathbf{r}, \mathbf{x}^{(av)}) \varphi_j(\mathbf{r}, \mathbf{x}^{(av)}). \quad (22)$$

This approximate electron density is obtained using the idempotent density matrix approximation  $\mathbf{P}(\mathbf{x}^{(av)})$  where the charge conservation is maintained, and the idempotency provides some quality assurances.

### 3 Conclusions

The approach presented provides a new averaging procedure designed directly for idempotent approximate density matrices corresponding to a “compromise” for a range of nuclear configurations, for example, those belonging to a region of a “catchment region” representing a stable chemical species on a potential energy surface. The new approach avoids some of the ambiguities associated with direct electron density averaging, that has not allowed in general to use simple quality tests, such as idempotency, available for density matrices.

**Acknowledgements** This work was supported by funds from the Institute for Advanced Study, Collegium Budapest, Hungary, a Visiting Professor grant from the Mathematics Department, Institute J.A. Dieudonne, University of Nice Sophia Antipolis, Nice, France, by an operating research grant from the Natural Sciences and Engineering Research Council of Canada, and the Canada Research Chair and CFI projects at the Memorial University of Newfoundland, Canada.

### References

1. P. Hohenberg, W. Kohn, *Phys. Rev.* **136**, B864 (1964)
2. J. Riess, W. Münch, *Theor. Chim. Acta* **58**, 295 (1981)
3. P.G. Mezey, *Mol. Phys.* **96**, 169 (1999)
4. P.G. Mezey, *J. Math. Chem.* **23**, 65 (1998)
5. P.G. Mezey, *J. Chem. Inf. Comp. Sci.* **39**, 224 (1999)
6. P.G. Mezey, *J. Math. Chem.* **30**, 299 (2001)
7. W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965)
8. R.G. Parr, *Proc. Natl. Acad. Sci. USA* **72**, 763 (1975)
9. M. Levy, *Proc. Natl. Acad. Sci. USA* **76**, 6062 (1979)
10. M. Levy, *Bull. Amer. Phys. Soc.* **24**, 626 (1979)
11. M. Levy, *Phys. Rev. A* **26**, 1200 (1982)
12. M. Levy, *Adv. Quant. Chem.* **21**, 69 (1990)
13. D.R. Hartree, *Proc. Cambridge Phil. Soc.* **24**, 111, 426 (1928), *ibid.* **25**, 225, 310 (1929)
14. V. Fock, *Z. Physik* **61**, 126 (1930)
15. C.C. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951), *ibid.* **32**, 179 (1960)
16. G.G. Hall, *Proc. Roy. Soc. London* **A205**, 541 (1951)
17. P.-O. Löwdin, *Phys. Rev.* **97**, 1474 (1955)
18. R.S. Mulliken, *J. Chem. Phys.* **23**, 1833, 1841, 2338, 2343 (1955)
19. R.S. Mulliken, *J. Chem. Phys.* **36**, 3428 (1962)
20. R. McWeeny, *Rev. Mod. Phys.* **32**, 335 (1960)
21. A.J. Coleman, *Rev. Mod. Phys.* **35**, 668 (1963)
22. P.G. Mezey, *Theor. Chim. Acta* **58**, 309 (1981)
23. P.G. Mezey, *Theor. Chim. Acta* **60**, 97 (1981)
24. P.G. Mezey, *Internat. J. Quantum Chem., Quantum Biol. Symp.* **8**, 185 (1981)
25. P.G. Mezey, *Chem. Phys. Lett.* **82**, 100 (1981), **86**, 562 (1982)
26. P.G. Mezey, *Theor. Chim. Acta* **62**, 133 (1982)
27. P.G. Mezey, *Can. J. Chem.* **61**, 956 (1983) (volume dedicated to Prof. H. Gunning)
28. P.G. Mezey, *Theor. Chim. Acta* **63**, 9 (1983)
29. P.G. Mezey, *Internat. J. Quantum Chem.* **26**, 983 (1984)
30. P.G. Mezey, *Theor. Chim. Acta* **67**, 43 (1985)



31. P.G. Mezey, *Theor. Chim. Acta* **67**, 91 (1985)
32. P.G. Mezey, *Theor. Chim. Acta* **67**, 115 (1985)
33. P.G. Mezey, *Potential Energy Hypersurfaces* (Elsevier, Amsterdam, 1987)
34. P.G. Mezey, *J. Comput. Chem.* **19**, 1337 (1998)
35. P.G. Mezey, *Comput. Chem.* **25**, 69 (2001)
36. Zs. Szekeres, P.G. Mezey, *J. Math. Chem.* **30**, 315 (2001)
37. T.E. Exner, P.G. Mezey, *J. Phys. Chem. A* **106**, 5504 (2002)
38. P.G. Mezey, *Theochem*, **727**, 123 (2005) (Special Volume dedicated to Prof. Ramon Carbo-Dorca)
39. P.-O. Löwdin, *J. Chem. Phys.* **18**, 365 (1950)
40. P.-O. Löwdin, *Adv. Phys.* **5**, 1 (1956)
41. F.L. Pilar, *Elementary Quantum Chemistry* (McGraw-Hill, New York, 1968)
42. P.-O. Löwdin, *Adv. Quantum Chem.* **5**, 185 (1970)
43. L. Massa, L. Huang, J. Karle, *Int. J. Quantum Chem., Quantum Chem. Symp.* **29**, 371 (1995)
44. P.G. Mezey, *Adv. Molec. Struct. Res.* **4**, 115 (1998)
45. Zs. Szekeres, P.G. Mezey, *Molec. Phys.* **103**, 1013 (2005) (Special Volume dedicated to Prof. N. Handy)
46. P.D. Walker, P.G. Mezey, *J. Amer. Chem. Soc.* **115**, 12423 (1993)
47. P.D. Walker, P.G. Mezey, *J. Amer. Chem. Soc.* **116**, 12022 (1994)
48. P.D. Walker, P.G. Mezey, *Can. J. Chem.* **72**, 2531 (1994)
49. P.D. Walker, P.G. Mezey, *J. Comput. Chem.* **16**, 1238 (1995)
50. P.D. Walker, P.G. Mezey, *J. Math. Chem.* **17**, 203 (1995)
51. P.G. Mezey, Z. Zimpel, P. Warburton, P.D. Walker, D.G. Irvine, D.G. Dixon, B. Greenberg, *J. Chem. Inf. Comp. Sci.* **36**, 602 (1996)
52. P.G. Mezey, P.D. Walker, *Drug Discov. Today* (Elsevier Trend Journal), **2**, 6 (1997)
53. P.G. Mezey, *Pharma. News* **4**, 29 (1997)
54. P.G. Mezey, *J. Math. Chem.*, **18**, 141 (1995)
55. P.G. Mezey, *Adv. Quantum Chem.* **27**, 163 (1996)
56. P.G. Mezey, in *Computational Chemistry: Reviews and Current Trends*, vol. 1, ed. by J. Leszczynski (World Scientific Publ., Singapore, 1996), pp. 109–137
57. P.G. Mezey, *Int. J. Quantum Chem.* **63**, 39 (1997)
58. P.G. Mezey, *Internat. Rev. Phys. Chem.* **16**, 361 (1997)
59. P.G. Mezey, *Adv. Molec. Struct. Res.* **4**, 115 (1998)
60. P.G. Mezey, *J. Comput. Methods Sci. Eng. (JCMSE)* **1**, 99, (2001)
61. T.E. Exner, P.G. Mezey, *J. Phys. Chem. A* **106**, 5504 (2002)
62. T.E. Exner, P.G. Mezey, *J. Phys. Chem. A* **106**, 11791 (2002)
63. T.E. Exner, P.G. Mezey, *J. Comput. Chem.* **24**, 1980 (2003)
64. T.E. Exner, P.G. Mezey, *J. Phys. Chem.* **108**, 4301 (2004)
65. Zs. Szekeres, T.E. Exner, P.G. Mezey, *Int. J. Quantum Chem.* **104**, 847 (2005)
66. T.E. Exner, P.G. Mezey, *Phys. Chem. Chem. Phys.* **7**, 4061 (2005)
67. Zs. Szekeres, P.G. Mezey, P. Surjan, *Chem. Phys. Lett.* **424**, 420 (2006)
68. P.G. Mezey, *Shape in Chemistry: An Introduction to Molecular Shape and Topology* (VCH Publishers, New York, 1993)
69. P.G. Mezey, *Int. J. Quantum Chem., Quant. Biol. Symp.* **12**, 113 (1986)
70. P.G. Mezey, *J. Amer. Chem. Soc.* **108**, 3976 (1986)
71. P.G. Mezey, *J. Comput. Chem.* **8**, 462 (1987)
72. P.G. Mezey, *Int. J. Quantum Chem., Quant. Biol. Symp.* **14**, 127 (1987)
73. P.G. Mezey, *Struct. Chem.* **6**, 261 (1995)
74. P.G. Mezey, in *Encyclopedia of Computational Chemistry*, vol. 4, ed. by P.v.R. Schleyer, N.L. Allinger, T. Clark, J. Gasteiger, P.A. Kollman, H.F. Schaefer III, P.R. Schreiner (John Wiley & Sons, Chichester, UK, 1998), pp. 2582–2589
75. P.G. Mezey, *Internat. Rev. Phys. Chem.* **16**, 361 (1997)
76. P.G. Mezey, *J. Chem. Inf. Comp. Sci.* **39**, 224 (1999)
77. A. Jokic, Z. Zimpel, P.M. Huang, P.G. Mezey, *SAR QSAR Environ. Res.* **12**, 297 (2001)
78. R. Carbó, L.Leyda, M. Arnau, *Int. J. Quantum Chem.* **17**, 1185 (1980)
79. E.E. Hodgkin, W.G. Richards, *J. Chem. Soc. Chem. Commun.* **1986**, 1342 (1986)
80. R. Carbó, Ll. Domingo, *Int. J. Quantum Chem.* **32**, 517 (1987)
81. S. Leicester, R. Bywater, J.L. Finney, *J. Mol. Graph.* **6**, 104 (1988)

82. M.A. Johnson, G.M. Maggiora (eds.), *Concepts and Applications of Molecular Similarity* (Wiley, New York, 1990)
83. R. Ponec, M. Strnad, *J. Chem. Inf. Comp. Sci.* **32**, 693 (1992)
84. G. Sello, *J. Chem. Inf. Comp. Sci.* **32**, 713 (1992)
85. P.G. Mezey, in *Topics in Current Chemistry*, vol. 173, *Molecular Similarity*, ed. by K. Sen (Springer-Verlag, Heidelberg, 1995), pp. 63–83
86. B.B. Stefanov, J. Cioslowski, *J. Comput. Chem.* **16**, 1394 (1995)
87. P.M. Dean (ed.), *Molecular Similarity in Drug Design* (Chapman & Hall - Blackie Publishers, Glasgow, U.K., 1995)
88. P.G. Mezey, *Adv. Molec. Similar.* **1**, 89 (1996)