Journal of Mathematical Chemistry 5(1990)323–357

THEORY OF PROJECTED PROBABILITIES ON NON-ORTHOGONAL STATES: APPLICATION TO ELECTRONIC POPULATIONS IN MOLECULES

R.S. MANNING* and N. De LEON*†
Department of Chemistry, Yale University, 225 Prospect Street, New Haven, CT 06511, USA

Received 28 November 1989

Abstract

Often it is important to consider the expansion of a quantum state $|\psi\rangle$ in terms of physically meaningful basis states. For example, molecular orbitals can be expressed as linear combinations of atomic orbitals, or vibrational states can be expressed as superpositions of local or normal mode eigenstates. In such expansions, it then becomes desirable to determine how much "character" a quantum state has in one of these basis states. One way of accimplishing this task is to calculate the projected probability of $|\psi\rangle$ on basis state $|j\rangle$. In this paper, we consider this general quantum mechanical problem. If the basis states are orthonormal, then the projected probability of $|\psi\rangle$ on $|j\rangle$ is of course $|\langle\psi|j\rangle|^2$. However, if the basis states are not orthogonal, then this result is no longer valid and one must develop a more general theory to calculate these projected probabilities. An earlier paper used one-dimensional projection operators to initiate this theory and gave closed form results for the case of two non-orthogonal basis states [1]. One- and many-dimensional projection operators, together with linear algebraic techniques, are used to extend this theory to the n non-orthogonal basis state case. Explicit closed form results are given for the two- and three-state cases, and a general algorithm is developed for the case of four or more basis states. Application of the theory is made to atomic populations in three- to six-atom molecules, and comparisons are made to the related work of Mulliken.

1. Introduction

Consider a normalized quantum state $|\psi\rangle$ expressed as a linear combination of several normalized basis states [2]:

$$|\psi\rangle = \sum_{i=1}^{n} \alpha_i |i\rangle.$$

If the basis states $|i\rangle$ are orthogonal, one can interpret α_i^2 as the projected probability of $|\psi\rangle$ on the basis state $|i\rangle$. For non-orthogonal basis sets, however, α_i^2 cannot be

*National Science Foundation Pre-doctoral Fellow. Present address: Center for Applied Mathematics, Cornell University, Ithaca, NY 14853, USA.

*To whom all correspondence should be addressed.

† Present address: Department of Chemistry/Physics/Astronomy, Indiana University Northwest, Gary, IN 46408, USA.

interpreted in this manner. Nevertheless, the notion of a "projected" probability is still useful as a measure of the contribution of each basis state to the overall state $|\psi\rangle$; therefore, we generalize the concept of a projected probability to include both orthogonal and non-orthogonal basis states.

The method developed in this paper involves the projection operator decomposition of $|\psi\rangle$ into components along the basis states. The squares of these components are interpreted as contributions to overall probabilities. The particular projection operator decomposition we outline guarantees that each probability is between zero and one, and that the sum of all of the probabilities is unity.

Projected probabilities can be applied to various physical phenomena. For example, they have often been used to calculate electronic populations in LCAO (Linear Combination of Atomic Orbitals) molecular orbitals [1,3]. In addition, these probabilities can be used to calculate correlation energies [4] and the normal and local mode character of molecular vibrations [5].

Section 2.1 begins the discussion of a theory of projected probabilities by focusing on the case in which $|\psi\rangle$ is expressed in terms of two basis states. This theory is expanded in section 2.2 to include three basis states. Next, a generalization of the theory to the n-state case is offered in section 2.3, followed by a description in section 3 of a computational algorithm which can be used to implement the theoretical results. In section 4, the theory is applied to the problem of electronic populations, including several examples and comparisons to Mulliken populations. Finally, in section 5, we briefly discuss our results.

2. Theory

2.1. THE 2-STATE CASE

Consider a normalized state $|\psi\rangle$ expressed as a linear combination of two normalized basis states [1]

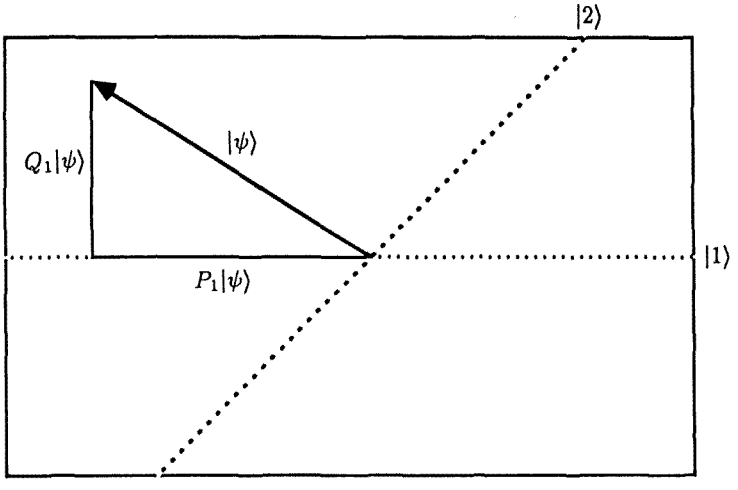
$$|\psi\rangle = \alpha_1|1\rangle + \alpha_2|2\rangle.$$

First, define the projection operators P_i and their orthogonal complements Q_i:

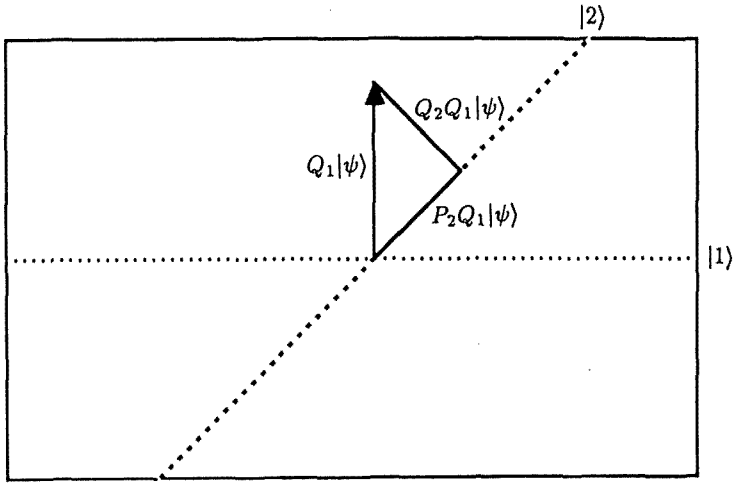
$$P_1 \equiv |1\rangle\langle 1|, \quad P_2 \equiv |2\rangle\langle 2|,$$

$$Q_1 \equiv 1 - P_1, \quad Q_2 \equiv 1 - P_2.$$

Using these operators, we can expand $|\psi\rangle$ as follows (see fig. 1):

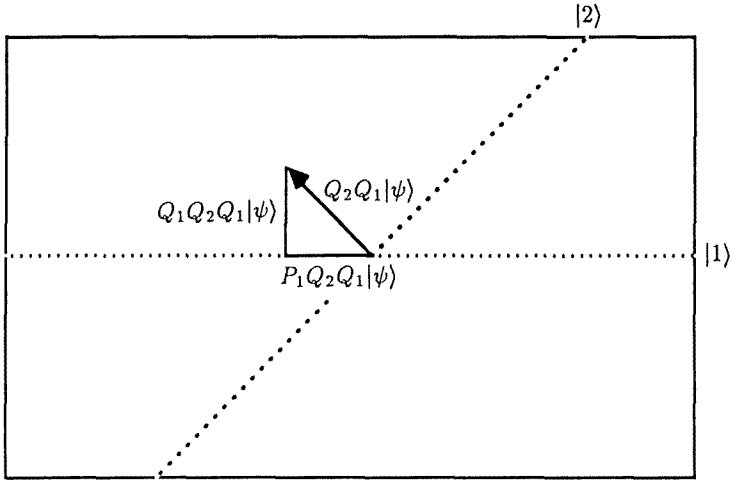


Projection 1

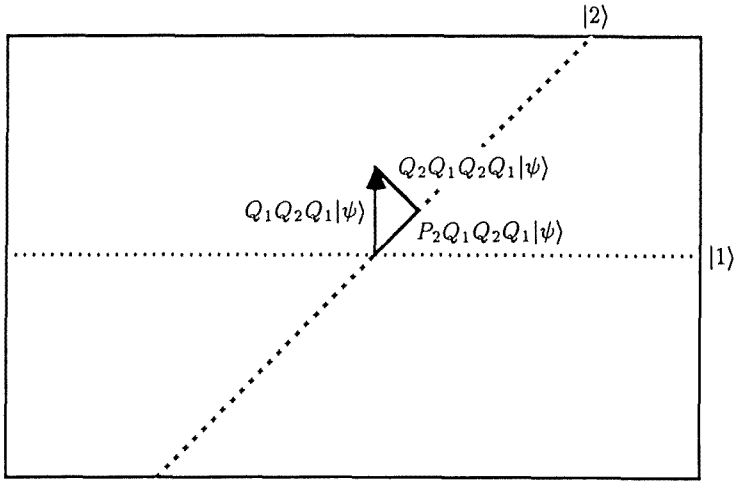


Projection 2

Fig. 1. (caption on following page).



Projection 3



Projection 4

Fig. 1. Projection operator decomposition for two states. The initial vector for each projection is shown with an arrowhead and with its tail at the origin; its parallel component (projection) and perpendicular component (orthogonal projection) are shown without arrowheads. The perpendicular component of a given projection becomes the initial vector for the next projection, as indicated by the labels.

$$
\begin{aligned}
|\psi\rangle &= P_1|\psi\rangle + Q_1|\psi\rangle \\
&= P_1|\psi\rangle + P_2 Q_1|\psi\rangle + Q_2 Q_1|\psi\rangle \\
&= P_1|\psi\rangle + P_2 Q_1|\psi\rangle + P_1 Q_2 Q_1|\psi\rangle + Q_1 Q_2 Q_1|\psi\rangle \\
&= P_1|\psi\rangle + P_2 Q_1|\psi\rangle + P_1 Q_2 Q_1|\psi\rangle + P_2 Q_1 Q_2 Q_1|\psi\rangle + \dots ,
\end{aligned} \tag{1}
$$

since $P_1 + Q_1 = 1$ and $P_2 + Q_2 = 1$. Starting instead with the expansion $|\psi\rangle = P_2|\psi\rangle + Q_2|\psi\rangle$, and then following the above pattern, we find:

$$
|\psi\rangle = P_2|\psi\rangle + P_1 Q_2|\psi\rangle + P_2 Q_1 Q_2|\psi\rangle + P_1 Q_2 Q_1 Q_2|\psi\rangle + \dots . \tag{2}
$$

Clearly, we could begin the decomposition by states with P_1 or P_2. The symmetrized expansion which avoids any initial bias is:

$$
|\psi\rangle = \frac{1}{2}\begin{bmatrix} P_1(1 + Q_2 + Q_2 Q_1 + Q_2 Q_1 Q_2 + Q_2 Q_1 Q_2 Q_1 + \dots)|\psi\rangle + \\ P_2(1 + Q_1 + Q_1 Q_2 + Q_1 Q_2 Q_1 + Q_1 Q_2 Q_1 Q_2 + \dots)|\psi\rangle \end{bmatrix}. \tag{3}
$$

From the expansion in eq. (3), we define the probabilities $(\mathcal{P}_1)_{12}$ and $(\mathcal{P}_2)_{12}$ in terms of "probability operators" $(\hat{P}_1)_{12}$ and $(\hat{P}_2)_{12}$:

$$
\begin{aligned}
(\mathcal{P}_1)_{12} &\equiv \langle\psi|(\hat{P}_1)_{12}|\psi\rangle, \\
(\mathcal{P}_2)_{12} &\equiv \langle\psi|(\hat{P}_2)_{12}|\psi\rangle,
\end{aligned} \tag{4}
$$

where the subscripts inside the parentheses indicate the basis state for which we are calculating the probability, and the subscripts outside the parentheses indicate the entire set of basis states. Although this notation is cumbersome for the 2-state case, it is necessary for the general problem.

All that remains is to define the probability operators. Notice that eq. (3) partitions $|\psi\rangle$ into the $|1\rangle$ direction and in the $|2\rangle$ direction. That is,

$$
\begin{aligned}
\alpha_1|1\rangle &= \tfrac{1}{2}\left[P_1(1 + Q_2 + Q_2 Q_1 + \dots)|\psi\rangle\right], \\
\alpha_2|2\rangle &= \tfrac{1}{2}\left[P_2(1 + Q_1 + Q_1 Q_2 + \dots)|\psi\rangle\right].
\end{aligned}
$$

Thus, each term in eq. (3) which begins with P_1 is a projection of $|\psi\rangle$ onto $|1\rangle$. Hence, each of these terms can be interpreted as a probability amplitude on $|1\rangle$. The sum of the absolute squares of these terms is then the total projected probability on $|1\rangle$. By this reasoning, and by parallel reasoning for $|2\rangle$, we define the probability operators:

$$(\hat{P}_1)_{12} \equiv \tfrac{1}{2}\left[P_1 + Q_2 P_1 Q_2 + Q_1 Q_2 P_1 Q_2 Q_1 + Q_2 Q_1 Q_2 P_1 Q_2 Q_1 Q_2 + \ldots\right],$$

$$(\hat{P}_2)_{12} \equiv \tfrac{1}{2}\left[P_2 + Q_1 P_2 Q_1 + Q_2 Q_1 P_2 Q_1 Q_2 + Q_1 Q_2 Q_1 P_2 Q_1 Q_2 Q_1 + \ldots\right]. \quad (5)$$

It can easily be verified that the sum of the absolute squares of the terms in each line of eq. (1) is unity. Similarly, the sum of the squares of the terms in eq. (2) must be unity. Therefore, the symmetrized probability operators in eq. (5) must also sum to unity, as must the probabilities defined in eq. (4).

The infinite sums in eq. (5) can be recast by noting the relations:

$$Q_2 Q_1 Q_2 |1\rangle = S_{12}^2 Q_2 |1\rangle,$$

$$Q_1 Q_2 Q_1 |2\rangle = S_{12}^2 Q_1 |2\rangle,$$

which imply that alternate terms in eq. (5) beginning with the second term form a convergent geometric series, as do alternate terms beginning with the third term. In this way, eq. (5) simplifies to the finite expressions:

$$(\hat{P}_1)_{12} = (1 - S_{12}^4)^{-1}\left[P_1 + S_{12}^2 P_2 - \frac{1 + S_{12}^2}{2}(P_1 P_2 + P_2 P_1)\right],$$

$$(\hat{P}_2)_{12} = (1 - S_{12}^4)^{-1}\left[P_2 + S_{12}^2 P_1 - \frac{1 + S_{12}^2}{2}(P_1 P_2 + P_2 P_1)\right]. \quad (6)$$

From these expressions, we determine the probabilities [6]:

$$(\mathcal{P}_1)_{12} = \langle\psi|(\hat{P}_1)_{12}|\psi\rangle = \frac{\alpha_1^2 + \alpha_2^2 S_{12}^2}{1 + S_{12}^2} + \alpha_1 \alpha_2 S_{12},$$

$$(\mathcal{P}_2)_{12} = \langle\psi|(\hat{P}_2)_{12}|\psi\rangle = \frac{\alpha_2^2 + \alpha_1^2 S_{12}^2}{1 + S_{12}^2} + \alpha_1 \alpha_2 S_{12}. \quad (7)$$

Note that for $S_{12} = 0$, the expressions in eq. (7) reduce to the usual expressions: $\mathcal{P}_i = \alpha_i^2$.

2.2. THE 3-STATE CASE

We now consider the 3-state expansion of the normalized quantum state $|\psi\rangle$:

$$|\psi\rangle = \alpha_1|1\rangle + \alpha_2|2\rangle + \alpha_3|3\rangle.$$

The advantage of the 2-state projection operator decomposition in eq. (3) is that, once the initial projection is chosen, the alternating pattern of future projections is well determined. If we impose a similar projection operator decomposition on the 3-state case, a similar well-determined pattern does not emerge, since at each step there are two possible vectors for the next projection. To overcome this problem, we expand $|\psi\rangle$ by alternately projecting onto one of the basis vectors and onto the *plane* formed by the remaining two basis vectors; in effect, this creates a new 2-state case, where one "state" is a vector and the other "state" is a plane.

Geometric intuition suggests that to project onto a plane, we project onto any two orthonormal vectors in this plane, and then add vectorially the resulting projections. For example, in order to project onto the plane spanned by $|2\rangle$ and $|3\rangle$, we can project onto the vectors $|2\rangle$ and $(1 - S_{23}^2)^{-1/2} Q_2 |3\rangle$. The elementary projection operators corresponding to each of these vectors are P_2 and $(1 - S_{23}^2)^{-1} Q_2 P_3 Q_2$; hence, the overall plane projection operators are:

$$P_{23} \equiv P_2 + \frac{Q_2 P_3 Q_2}{1 - S_{23}^2} = \frac{P_2 + P_3 - P_2 P_3 - P_3 P_2}{1 - S_{23}^2},$$

$$Q_{23} \equiv 1 - P_{23}.$$

The following properties are easily verified from the definitions given above:

$$P_{23} P_2 = P_2 P_{23} = P_2,$$

$$P_{23} P_3 = P_3 P_{23} = P_3.$$

Geometrically, these properties imply that the projection of a vector $|i\rangle$ onto a plane in which it lies is the vector $|i\rangle$ itself. The obvious corollaries are:

$$Q_{23} P_2 = P_2 Q_{23} = 0,$$

$$Q_{23} P_3 = P_3 Q_{23} = 0.$$

In addition, from the definition of Q_{23}, we can derive the expression:

$$\langle 1 | Q_{23} | 1 \rangle = \frac{1 - S_{12}^2 - S_{13}^2 - S_{23}^2 + 2 S_{12} S_{13} S_{23}}{1 - S_{23}^2} = \frac{\det \mathbf{S}_{123}}{\det \mathbf{S}_{23}}, \tag{8}$$

where the S matrices are the overlap matrices for the subscripted basis states. We show in appendix 1 that $0 < \langle 1 | Q_{23} | 1 \rangle \leq 1$.

We are now positioned to treat the 3-state case. We choose one of the three basis vectors, say $|1\rangle$, and alternately project onto that basis vector and onto the plane formed by the remaining two basis vectors:

$$\begin{aligned}|\psi\rangle &= P_1|\psi\rangle + Q_1|\psi\rangle \\ &= P_1|\psi\rangle + P_{23}Q_1|\psi\rangle + Q_{23}Q_1|\psi\rangle \\ &= P_1|\psi\rangle + P_{23}Q_1|\psi\rangle + P_1Q_{23}Q_1|\psi\rangle + \ldots\ .\end{aligned}$$

As in the 2-state case, we use the symmetrized expansion of $|\psi\rangle$:

$$|\psi\rangle = \frac{1}{2}\begin{bmatrix} P_1(1 + Q_{23} + Q_{23}Q_1 + Q_{23}Q_1Q_{23} + \ldots)|\psi\rangle + \\ P_{23}(1 + Q_1 + Q_1Q_{23} + Q_1Q_{23}Q_1 + \ldots)|\psi\rangle \end{bmatrix}. \tag{9}$$

Despite this symmetrization, eq. (9) still biases the isolated basis state $|1\rangle$ over the others. This bias will be eliminated by an additional symmetrization described later in this section.

The expansion in eq. (9) is the exact expansion used in the 2-state case (cf. eq. (3)), with P_{23} replacing P_2, and Q_{23} replacing Q_2; thus, we define probability operators as in the 2-state case. Since the $|1\rangle$ component of $|\psi\rangle$ has been isolated in eq. (9), we define:

$$\begin{aligned}&(\hat{P}_{11})_{123} \\ &\equiv \tfrac{1}{2}\left[P_1 + Q_{23}P_1Q_{23} + Q_1Q_{23}P_1Q_{23}Q_1 + Q_{23}Q_1Q_{23}P_1Q_{23}Q_1Q_{23} + \ldots\right],\end{aligned} \tag{10}$$

as in eq. (5). The second subscript inside parentheses of the probability operator indicates the isolated basis state in the expansion of $|\psi\rangle$. As before, the first subscript indicates which state's probability we are calculating, and the subscripts outside the parentheses indicate the complete set of basis states.

Given eq. (9), we cannot so simply write down $(\hat{P}_{21})_{123}$ and $(\hat{P}_{31})_{123}$, since the $|2\rangle$ and $|3\rangle$ components of $|\psi\rangle$ have not been separated: instead, there are terms that begin with P_{23} which contain both $|2\rangle$ and $|3\rangle$ components. However, each term that begins with P_{23} is a vector in the two-dimensional space spanned by $|2\rangle$ and $|3\rangle$ and can be expanded using the standard 2-state projection operator decompositon (cf. eq. (3)):

$$P_{23}|\psi\rangle = \frac{1}{2}\begin{bmatrix} P_2(1 + Q_3 + Q_3Q_2 + Q_3Q_2Q_3 + \ldots)P_{23}|\psi\rangle + \\ P_3(1 + Q_2 + Q_2Q_3 + Q_2Q_3Q_2 + \ldots)P_{23}|\psi\rangle \end{bmatrix},$$

$$P_{23}Q_1|\psi\rangle = \frac{1}{2}\begin{bmatrix} P_2(1 + Q_3 + Q_3Q_2 + Q_3Q_2Q_3 + \ldots)P_{23}Q_1|\psi\rangle + \\ P_3(1 + Q_2 + Q_2Q_3 + Q_2Q_3Q_2 + \ldots)P_{23}Q_1|\psi\rangle \end{bmatrix},$$

$\ldots$

From each of these expansions, we extract a contribution to the probability operator $(\hat{P}_{21})_{123}$:

Contribution from $P_{23}|\psi\rangle$:

$$\tfrac{1}{2}\left[P_{23}P_2P_{23} + P_{23}Q_3P_2Q_3P_{23} + P_{23}Q_2Q_3P_2Q_3Q_2P_{23} + \ldots\right]$$

$$= \tfrac{1}{2}P_{23}\left[P_2 + Q_3P_2Q_3 + Q_2Q_3P_2Q_3Q_2 + \ldots\right]P_{23}$$

$$= P_{23}(\hat{P}_2)_{23}P_{23}\,.$$

Contribution from $P_{23}Q_1|\psi\rangle$:

$$\tfrac{1}{2}\left[Q_1P_{23}P_2P_{23}Q_1 + Q_1P_{23}Q_3P_2Q_3P_{23}Q_1 + Q_1P_{23}Q_2Q_3P_2Q_3Q_2P_{23}Q_1 + \ldots\right]$$

$$= Q_1P_{23}(\hat{P}_2)_{23}P_{23}Q_1\,.$$

$$\ldots$$

Collecting all such contributions yields:

$$(\hat{P}_{21})_{123}$$

$$\equiv \frac{1}{2}\left[\begin{array}{l} P_{23}(\hat{P}_2)_{23}P_{23} + Q_1P_{23}(\hat{P}_2)_{23}P_{23}Q_1 + Q_{23}Q_1P_{23}(\hat{P}_2)_{23}P_{23}Q_1Q_{23} + \\ Q_1Q_{23}Q_1P_{23}(\hat{P}_2)_{23}P_{23}Q_1Q_{23}Q_1 + \ldots \end{array}\right]. \quad (11)$$

By symmetry, we can immediately write $(\hat{P}_{31})_{123}$:

$$(\hat{P}_{31})_{123}$$

$$\equiv \frac{1}{2}\left[\begin{array}{l} P_{23}(\hat{P}_3)_{23}P_{23} + Q_1P_{23}(\hat{P}_3)_{23}P_{23}Q_1 + Q_{23}Q_1P_{23}(\hat{P}_3)_{23}P_{23}Q_1Q_{23} + \\ Q_1Q_{23}Q_1P_{23}(\hat{P}_3)_{23}P_{23}Q_1Q_{23}Q_1 + \ldots \end{array}\right]. \quad (12)$$

Next, we simplify the expressions for the probability operators (eqs. (10)–(12)), beginning with the expression for $(\hat{P}_{11})_{123}$ in eq. (19). We observe that:

$$\begin{aligned} Q_{23}Q_1Q_{23}|1\rangle &= Q_{23}(1-P_1)Q_{23}|1\rangle \\ &= \left[1-\langle 1|Q_{23}|1\rangle\right]Q_{23}|1\rangle \\ &= (X_1)_{123}Q_{23}|1\rangle, \end{aligned} \quad (13)$$

where $(X_1)_{123} \equiv 1 - (\det \mathbf{S}_{123}/\det \mathbf{S}_{23})$ from eq. (8). Applying Q_1 to the left-hand side of this equation yields:

$$Q_1 Q_{23} Q_1 Q_{23} |1\rangle = (X_1)_{123} Q_1 Q_{23} |1\rangle. \tag{14}$$

Equations (13) and (14) imply:

$$Q_{23} Q_1 Q_{23} P_1 Q_{23} Q_1 Q_{23} = (X_1)^2_{123} Q_{23} P_1 Q_{23},$$

$$Q_1 Q_{23} Q_1 Q_{23} P_1 Q_{23} Q_1 Q_{23} Q_1 = (X_1)^2_{123} Q_1 Q_{23} P_1 Q_{23} Q_1.$$

Therefore, in the expression for $(\hat{P}_{11})_{123}$ in eq. (10), alternate terms beginning with the second term form a geometric series with geometric ratio $(X_1)^2_{123}$; similarly, alternate terms beginning with the third term form a geometric series with the same ratio. In addition, since $0 < \langle 1|Q_{23}|1\rangle \leq 1$ (see appendix 1), we are guaranteed that $0 \leq (X_1)^2_{123} < 1$, and the convergence of the geometric series is assured:

$$(\hat{P}_{11})_{123} = \frac{1}{2}\left[P_1 + \frac{Q_{23} P_1 Q_{23}}{1-(X_1)^2_{123}} + \frac{Q_1 Q_{23} P_1 Q_{23} Q_1}{1-(X_1)^2_{123}}\right]. \tag{15}$$

We use a similar strategy to simplify the expression for $(\hat{P}_{21})_{123}$ in eq. (11). Note that since $P_{23} P_2 = P_2$ and $P_{23} P_3 = P_3$, and since $(\hat{P}_2)_{23}$ is given as follows (cf. eq. (6)):

$$(\hat{P}_2)_{23} = (1 - S^4_{23})^{-1}\left[P_2 + S^2_{23} P_3 - \frac{1+S^2_{23}}{2}(P_3 P_2 + P_2 P_3)\right],$$

we can make the simplification:

$$P_{23}(\hat{P}_2)_{23} = (\hat{P}_2)_{23} P_{23} = (\hat{P}_2)_{23}. \tag{16}$$

Then, eq. (11) reduces to:

$$(\hat{P}_{21})_{123} = \tfrac{1}{2}[(\hat{P}_2)_{23} + Q_1(\hat{P}_2)_{23} Q_1 + Q_{23} Q_1 (\hat{P}_2)_{23} Q_1 Q_{23} + Q_1 Q_{23} Q_1 (\hat{P}_2)_{23} Q_1 Q_{23} Q_1 + \ldots]. \tag{17}$$

We observe that:

$$Q_{23}Q_1(\hat{P}_2)_{23}Q_1Q_{23} = Q_{23}(1-P_1)(\hat{P}_2)_{23}(1-P_1)Q_{23}$$
$$= Q_{23}P_1(\hat{P}_2)_{23}P_1Q_{23} + Q_{23}(\hat{P}_2)_{23}Q_{23}$$
$$- Q_{23}(\hat{P}_2)_{23}P_1Q_{23} - Q_{23}P_1(\hat{P}_2)_{23}Q_{23}$$
$$= Q_{23}P_1(\hat{P}_2)_{23}P_1Q_{23}, \tag{18}$$

since $Q_{23}(\hat{P}_2)_{23} = (\hat{P}_2)_{23}Q_{23} = 0$ (cf. eq. (16)). Therefore, eq. (17) can be rewritten as:

$$(\hat{P}_{21})_{123} = \tfrac{1}{2}[(\hat{P}_2)_{23} + Q_1(\hat{P}_2)_{23}Q_1 + Q_{23}P_1(\hat{P}_2)_{23}P_1Q_{23}$$
$$+ Q_1Q_{23}P_1(\hat{P}_2)_{23}P_1Q_{23}Q_1 + \ldots].$$

Applying eqs. (13) and (14) to this series, we see that alternate terms beginning with the third term, and alternate terms beginning with the fourth term, form two convergent geometric series with geometric ratio $(X_1)^2_{123}$, leading to the expression:

$$(\hat{P}_{21})_{123} = \frac{1}{2}\left[(\hat{P}_2)_{23} + Q_1(\hat{P}_2)_{23}Q_1 + \frac{Q_{23}P_1(\hat{P}_2)_{23}P_1Q_{23}}{1-(X_1)^2_{123}} + \frac{Q_1Q_{23}P_1(\hat{P}_2)_{23}P_1Q_{23}Q_1}{1-(X_1)^2_{123}}\right]. \tag{19}$$

By symmetry,

$$(\hat{P}_{31})_{123} = \frac{1}{2}\left[(\hat{P}_3)_{23} + Q_1(\hat{P}_3)_{23}Q_1 + \frac{Q_{23}P_1(\hat{P}_3)_{23}P_1Q_{23}}{1-(X_1)^2_{123}} + \frac{Q_{23}P_1(\hat{P}_3)_{23}P_1Q_{23}}{1-(X_1)^2_{123}}\right]. \tag{20}$$

As noted earlier, in order to eliminate the bias created by choosing an isolated basis vector, we must symmetrize our expansion of $|\psi\rangle$:

$$|\psi\rangle = \tfrac{1}{3}\left[|\psi\rangle + |\psi\rangle + |\psi\rangle\right]$$

$$= \frac{1}{3}\left[\begin{array}{l} \frac{1}{2}\left[\begin{array}{l} P_1(1 + Q_{23} + Q_{23}Q_1 + Q_{23}Q_1Q_{23} + \ldots)|\psi\rangle + \\ P_{23}(1 + Q_1 + Q_1Q_{23} + Q_1Q_{23}Q_1 + \ldots)|\psi\rangle \end{array}\right] + \\ \\ \frac{1}{2}\left[\begin{array}{l} P_2(1 + Q_{13} + Q_{13}Q_2 + Q_{13}Q_2Q_{13} + \ldots)|\psi\rangle + \\ P_{13}(1 + Q_2 + Q_2Q_{13} + Q_2Q_{13}Q_2 + \ldots)|\psi\rangle \end{array}\right] + \\ \\ \frac{1}{2}\left[\begin{array}{l} P_3(1 + Q_{12} + Q_{12}Q_3 + Q_{12}Q_3Q_{12} + \ldots)|\psi\rangle + \\ P_{12}(1 + Q_3 + Q_3Q_{12} + Q_3Q_{12}Q_3 + \ldots)|\psi\rangle \end{array}\right] \end{array}\right]. \tag{21}$$

By symmetry, we can write expressions for $(\hat{P}_{12})_{123}$ and $(\hat{P}_{13})_{123}$ from eqs. (19) and (20) and then evaluate $(\hat{P}_1)_{123}$, the operator for the projected probability on $|1\rangle$:

$$(\hat{P}_1)_{123} = \tfrac{1}{3}\left[(\hat{P}_{11})_{123} + (\hat{P}_{12})_{123} + (\hat{P}_{13})_{123}\right],$$

with parallel results for $(\hat{P}_2)_{123}$ and $(\hat{P}_3)_{123}$.

A direct implementation of eqs. (15), (19), and (20) to calculate 3-state probabilities, while straightforward, is significantly more complicated than in the 2-state case and is detailed in appendix 2.

2.3. THE GENERAL n-STATE CASE

To derive the projected probabilities of $|\psi\rangle$ on $n > 3$ basis states, we must generalize the definition of a plane projection operator. An n-dimensional hyperplane projection operator gives the geometric projection onto the vector space spanned by n independent vectors. The definition of a hyperplane projection operator comes directly from this geometric interpretation. Given an orthonormal basis for an n-dimensional vector space, any vector is equal to the vector sum of its projections onto all of the orthonormal basis vectors. Therefore, the hyperplane projection operator is defined as the sum of the elementary projection operators of an orthonormal basis $\{|1'\rangle, |2'\rangle, \ldots, |n'\rangle\}$ for the space spanned by $|1\rangle, \ldots, |n\rangle$ [7]:

$$P_{1\ldots n} \equiv |1'\rangle\langle 1'| + |2'\rangle\langle 2'| + \ldots + |n'\rangle\langle n'| \ . \tag{22}$$

We can easily verify that $P_{1\ldots n}$ is well defined, i.e. it is independent of the choice of orthonormal basis (see appendix 1). Next, the orthogonal complement of $P_{1\ldots n}$ is defined by $Q_{1\ldots n} \equiv 1 - P_{1\ldots n}$. These hyperplane projection operators $P_{1\ldots n}$ and $Q_{1\ldots n}$ obey the same essential algebra as the elementary projection operators P_i and Q_i.

It can be shown (see appendix 1) that given any vector $|\phi\rangle$, which may or may not lie in the space spanned by $\{|1\rangle, |2\rangle, \ldots, |n\rangle\}$, we have:

$$\langle\phi|Q_{1\ldots n}|\phi\rangle = \frac{\det \mathbf{S}_{1\ldots n\phi}}{\det \mathbf{S}_{1\ldots n}}, \tag{23}$$

where the S matrices are the overlap matrices of the subscripted states. We find, in addition, that if $|\phi\rangle$ is normalized, then $0 < \langle\phi|Q_{1\ldots n}|\phi\rangle \leq 1$.

The hyperplane projection operator allows the derivation of the n-state probability operator to be essentially the same as the 3-state derivation. We choose one of the n basis vectors, say $|j\rangle$, to isolate for an initial projection, and then form the hyperplane projection operators $P_{1\ldots j-1,j+1\ldots n}$ and $Q_{1\ldots j-1,j+1\ldots n}$. (For simplicity of notation, abbreviate these operators to P^j and Q^j). Then, we expand $|\psi\rangle$ using these operators (cf. eq. (9)):

$$|\psi\rangle = \frac{1}{2}\begin{bmatrix} P_j(1 + Q^j + Q^jQ_j + Q^jQ_jQ^j + \ldots)|\psi\rangle + \\ P^j(1 + Q_j + Q_jQ^j + Q_jQ^jQ_j + \ldots)|\psi\rangle \end{bmatrix}. \tag{24}$$

From the expansion in eq. (24), we define:

$$(\hat{P}_{jj})_{1\ldots n} \equiv \tfrac{1}{2}\left[P_j + Q^jP_jQ^j + Q_jQ^jP_jQ^jQ_j + \ldots\right], \tag{25}$$

as in eq. (10). Similarly, we define $(\hat{P}_{ij})_{1\ldots n}$ for $i \neq j$ as in the 3-state case, dividing each of the P^j terms in eq. (24) into components in the $|1\rangle, \ldots, |j-1\rangle, |j+1\rangle, \ldots, |n\rangle$ directions. Each of these P^j terms is a vector in an $(n-1)$-dimensional vector space, and we expand it into $(n-2)$-dimensional terms and 1-dimensional terms by the $(n-1)$-dimensional generalization of the expansion in eq. (21). Within this expansion, we expand each $(n-2)$-dimensional term by the $(n-2)$-dimensional generalization of eq. (21). Finally, after $n-1$ such expansions, each original P^j term is written as a *fully symmetrized* expansion in the directions of the $n-1$ basis vectors. Clearly, this expansion will be very complicated, and can not be easily written for large n. However, the expansion can still be used to define probability operators.

Consider any vector $|\chi\rangle$ in the space spanned by $|1\rangle, \ldots, |j-1\rangle, |j+1\rangle, \ldots, |n\rangle$. Expand $|\chi\rangle$ as described in the previous paragraph, and consider the sum of the absolute squares of all terms in this expansion that begin with P_i (i.e. that are projections onto the basis vector $|i\rangle$). *By definition*, this sum must be $(\hat{P}_i)_{1\ldots j-1,j+1\ldots n}$ (which we abbreviate to $(\hat{P}_i)^j$), since this expansion is the standard fully symmetrized projection operator expansion. Note that in the 3-state case, this same simplification occurred, introducing the operators $(\hat{P}_2)_{23}$ and $(\hat{P}_3)_{23}$.

We define a contribution of $P^j|\psi\rangle$ to $(\hat{P}_{ij})_{1\ldots n}$ by considering the sum of the absolute squares of its expansion terms which begin with P_i. This sum is identical to

the sum of absolute squares obtained from the expansion of $|\chi\rangle$ in the previous paragraph, except that each square begins and ends with P^j. Thus, the contribution to $(\hat{P}_{ij})_{1\ldots n}$ is $P^j(\hat{P}_i)^j P^j$. Similarly, the contribution of $P^j Q_j |\psi\rangle$ is $Q_j P^j(\hat{P}_i)^j P^j Q_j$, since each absolute square begins with $Q_j P^j$ and ends with $P^j Q_j$. This leads to the overall expression (cf. eq. (11)):

$$(\hat{P}_{ij})_{1\ldots n} = \frac{1}{2}\left[\begin{array}{l} P^j(\hat{P}_i)^j P^j + Q_j P^j(\hat{P}_i)^j P^j Q_j + Q^j Q_j P^j(\hat{P}_i)^j P^j Q_j Q^j + \\ Q_j Q^j Q_j P^j(\hat{P}_i)^j P^j Q_j Q^j Q_j + Q^j Q_j Q^j Q_j P^j(\hat{P}_i)^j P^j Q_j Q^j Q_j Q^j + \ldots \end{array}\right], \quad i \neq j. \tag{26}$$

Equations (25) and (26) give definitions for n-state probability operators as infinite sums and, as in the 2-state and 3-state cases, we use the properties of hyperplane projection operators to reduce these infinite sums to finite sums. We observe first that:

$$\begin{aligned} Q^j Q_j Q^j |j\rangle &= Q^j(1 - P_j)Q^j |j\rangle \\ &= \left[1 - \langle j| Q^j |j\rangle\right] Q^j |j\rangle \\ &= (X_j)_{1\ldots n} Q^j |j\rangle, \end{aligned} \tag{27}$$

where $(X_j)_{1\ldots n} \equiv 1 - (\det \mathbf{S}_{1\ldots n}/\det \mathbf{S}^j_{1\ldots n})$ from eq. (23), adopting the shorthand notation $\mathbf{S}^j_{1\ldots n}$ for the matrix $\mathbf{S}_{1\ldots j-1, j+1\ldots n}$. This gives the relations:

$$Q^j Q_j Q^j P_j Q^j Q_j Q^j = (X_j)^2_{1\ldots n} Q^j P_j Q^j,$$

$$Q_j Q^j Q_j Q^j P_j Q^j Q_j Q^j Q_j = (X_j)^2_{1\ldots n} Q_j Q^j P_j Q^j Q_j .$$

Hence, alternate terms in eq. (25) beginning with the second term, and alternate terms beginning with the third term, form two convergent geometric series (since $(X_j)^2_{1\ldots n} < 1$ from appendix 1), giving the expression:

$$(\hat{P}_{jj})_{1\ldots n} = \frac{1}{2}\left[P_j + \frac{Q^j P_j Q^j}{1 - (X_j)^2_{1\ldots n}} + \frac{Q_j Q^j P_j Q^j Q_j}{1 - (X_j)^2_{1\ldots n}}\right]. \tag{28}$$

We simplify eq. (26) in a similar manner. First, we apply the properties of appendix 1 and some operator algebra to eqs. (25) and (26) to find:

$$P_{1\ldots n}(\hat{P}_{ij})_{1\ldots n} P_{1\ldots n} = (\hat{P}_{ij})_{1\ldots n} \qquad \forall i, j,$$

which implies

$$P_{1\ldots n}(\hat{P}_i)_{1\ldots n}P_{1\ldots n} = (\hat{P}_i)_{1\ldots n} \qquad \forall i.$$

This result is not surprising since we expect $(\hat{P}_i)_{1\ldots n}$ to be constructed of only the operators $P_1, \ldots, P_n$ so that operating $P_{1\ldots n}$ on $(\hat{P}_i)_{1\ldots n}$ should have no effect. Then, we conclude the parallel $(n-1)$-state result:

$$P^j(\hat{P}_i)^jP^j = (\hat{P}_i)^j, \tag{29}$$

since we could use the same logic with $n-1$ states. Therefore, we can rewrite eq. (26) as:

$$(\hat{P}_{ij})_{1\ldots n} = \frac{1}{2}\left[\begin{array}{l}(\hat{P}_i)^j + Q_j(\hat{P}_i)^jQ_j + Q^jQ_j(\hat{P}_i)^jQ_jQ^j + \\ Q_jQ^jQ_j(\hat{P}_i)^jQ_jQ^jQ_j + Q^jQ_jQ^jQ_j(\hat{P}_i)^jQ_jQ^jQ_jQ^j + \ldots\end{array}\right], \qquad i \neq j. \tag{30}$$

As in eq. (18), we find:

$$Q^jQ_j(\hat{P}_i)^jQ_jQ^j = Q^jP_j(\hat{P}_i)^jP_jQ^j,$$

which leads to the simplification:

$$(\hat{P}_{ij})_{1\ldots n} = \frac{1}{2}\left[\begin{array}{l}(\hat{P}_i)^j + Q_j(\hat{P}_i)Q_j + Q^jP_j(\hat{P}_i)^jP_jQ^j + \\ Q_jQ^jP_j(\hat{P}_i)^jP_jQ^jQ_j + Q^jQ_jQ^jP_j(\hat{P}_i)^jP_jQ^jQ_jQ^j + \ldots\end{array}\right], \qquad i \neq j. \tag{31}$$

Using eq. (27), we see that alternate terms of eq. (31) beginning with the third term, and alternate terms beginning with the fourth term, form two convergent geometric series, so we write the final expression:

$$(\hat{P}_{ij})_{1\ldots n} = \frac{1}{2}\left[(\hat{P}_i)^j + Q_j(\hat{P}_i)^jQ_j + \frac{Q^jP_j(\hat{P}_i)^jP_jQ^j}{1-(X_j)^2_{1\ldots n}} + \frac{Q_jQ^jP_j(\hat{P}_i)^jQ^jQ_j}{1-(X_j)^2_{1\ldots n}}\right], \qquad i \neq j. \tag{32}$$

Finally, as in the 3-state case, in order to symmetrize with respect to the isolated basis vector, we write the full probability operator as:

$$(\hat{P}_i)_{1\ldots n} = \frac{1}{n}\sum_{j=1}^{n}(\hat{P}_{ij})_{1\ldots n}, \tag{33}$$

and determine the probabilities by the equation:

$$(\mathcal{P}_i)_{1\ldots n} = \langle\psi|(\hat{P}_i)_{1\ldots n}|\psi\rangle. \tag{34}$$

The actual computation of eq. (34) is described in section 3.

3. Computational algorithm for n-state probabilities

The derivation of n-state probability operators (section 2.3) is equivalent to the 3-state derivation (section 2.2) with a change of notation. However, the approach to calculating probabilities from these operators must be different. The expression for the 3-state probability operator in eq. (19) includes 2-state probability operators, and since eq. (6) provides a simple closed form for 2-state probability operators, we can calculate a closed form for the 3-state probabilities (see appendix 2). The expression for the n-state probability operator in eq. (32) includes $(n-1)$-state probability operators. However, we do not have a simple closed form for the $(n-1)$-state probability operator, and the complexity of the 3-state probability formula in appendix 2 suggests that no simple closed form exists for probability operators, and hence for probabilities, for $n > 3$. Fortunately, we can still calculate probabilities numerically, because the hyperplane projection operator formalism lends itself to a manageable computational algorithm.

Using the relations $|\psi\rangle = \sum_{l=1}^{n}\alpha_l|l\rangle$, we rewrite eq. (34) as:

$$(\mathcal{P}_i)_{1\ldots n} = \sum_{l=1}^{n}\sum_{m=1}^{n}\alpha_l\,\alpha_m\langle l|(\hat{P}_i)_{1\ldots n}|m\rangle. \tag{35}$$

Inserting the relation given by eq. (33) gives:

$$(\mathcal{P}_i)_{1\ldots n} = \frac{1}{n}\sum_{j=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n}\alpha_l\,\alpha_m\langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle. \tag{36}$$

Hence, to calculate $(\mathcal{P}_i)_{1\ldots n}$, it is sufficient to calculate $\langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle$ for all $j, l, m \le n$.

Note the similarity of the n-state probability operator $(\hat{P}_{jj})_{1\ldots n}$ in eq. (28) to the 3-state probability operator $(\hat{P}_{jj})_{123}$ in eq. (15). Likewise, note the similarity of the n-state probability operator $(\hat{P}_{ij})_{1\ldots n}$ in eq. (32) to the 3-state probability operator $(\hat{P}_{ij})_{123}$ in eq. (19). Therefore, we can calculate $\langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle$ exactly as in the 3-state case (see appendix 2). The results of these calculations are shown in table 2. The terms in

Table 1

Values of $\langle l|(\hat{P}_{ij})|m\rangle$ for the 3-state case

	$\langle l\|(\hat{P}_{jj})_{123}\|m\rangle$	$\langle l\|(\hat{P}_{ij})_{123}\|m\rangle,\ i\neq j$
$l\neq j,\ m\neq j$	$\frac{S_{jl}S_{jm}}{1+(X_j)_{123}}$	$\langle l\|(\hat{P}_i)_{ik}\|m\rangle - \frac{S_{jl}}{2}\langle j\|(\hat{P}_i)_{ik}\|m\rangle - \frac{S_{jm}}{2}\langle l\|(\hat{P}_i)_{ik}\|j\rangle + \frac{S_{jl}S_{jm}}{1+(X_j)_{123}}\langle j\|(\hat{P}_i)_{ik}\|j\rangle$
$l=j,\ m\neq j$	$\frac{S_{jm}}{2}$	$\frac{1}{2}\langle j\|(\hat{P}_i)_{ik}\|m\rangle$
$l\neq j,\ m=j$	$\frac{S_{jl}}{2}$	$\frac{1}{2}\langle l\|(\hat{P}_i)_{ik}\|j\rangle$
$l=j,\ m=j$	$\frac{1}{1+(X_j)_{123}}$	$\frac{1}{1+(X_j)_{123}}\langle j\|(\hat{P}_i)_{ik}\|j\rangle$

Table 2

Values of $\langle l|(\hat{P}_{ij})|m\rangle$ for the n-state case

	$\langle l\|(\hat{P}_{jj})_{1\ldots n}\|m\rangle$	$\langle l\|(\hat{P}_{ij})_{1\ldots n}\|m\rangle,\ i\neq j$
$l\neq j,\ m\neq j$	$\frac{S_{jl}S_{jm}}{1+(X_j)_{1\ldots n}}$	$\langle l\|(\hat{P}_i)^j\|m\rangle - \frac{S_{jl}}{2}\langle j\|(\hat{P}_i)^j\|m\rangle - \frac{S_{jm}}{2}\langle l\|(\hat{P}_i)^j\|j\rangle + \frac{S_{jl}S_{jm}}{1+(X_j)_{1\ldots n}}\langle j\|(\hat{P}_i)^j\|j\rangle$
$l=j,\ m\neq j$	$\frac{S_{jm}}{2}$	$\frac{1}{2}\langle j\|(\hat{P}_i)^j\|m\rangle$
$l\neq j,\ m=j$	$\frac{S_{jl}}{2}$	$\frac{1}{2}\langle l\|(\hat{P}_i)^j\|j\rangle$
$l=j,\ m=j$	$\frac{1}{1+(X_j)_{1\ldots n}}$	$\frac{1}{1+(X_j)_{1\ldots n}}\langle j\|(\hat{P}_i)^j\|j\rangle$

the left column of table 2 are easily computable, but in order to compute the terms in the right column, we introduce a recursive algorithm, since we do not have a closed form expression for the $(n-1)$-dimensional probability operators involved.

Consider some $j \neq i$. Express $|j\rangle$ as follows:

$$|j\rangle = \beta_1|1\rangle + \ldots + \beta_{j-1}|j-1\rangle + \beta_{j+1}|j+1\rangle + \ldots + \beta_n|n\rangle + |R\rangle,$$

where $|R\rangle$ is orthogonal to $|1\rangle, \ldots, |j-1\rangle, |j+1\rangle, \ldots, |n\rangle$. This is equivalent to the matrix equation:

$$|j\rangle = [|1\rangle, \ldots, |j-1\rangle, |j+1\rangle, \ldots, |n\rangle] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{j-1} \\ \beta_{j+1} \\ \vdots \\ \beta_n \end{bmatrix} + |R\rangle.$$

Operating on the left by the column vector $[|1\rangle, \ldots, |j-1\rangle, |j+1\rangle, \ldots, n\rangle]^\dagger$ gives:

$$\begin{bmatrix} S_{1j} \\ \vdots \\ S_{j-1,j} \\ S_{j+1,j} \\ \vdots \\ S_{nj} \end{bmatrix} = \mathbf{S}^j_{1\ldots n} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{j-1} \\ \beta_{j+1} \\ \vdots \\ \beta_n \end{bmatrix},$$

since $|R\rangle$ is orthogonal to $|1\rangle, \ldots, |j-1\rangle, |j+1\rangle, \ldots, |n\rangle$. Hence, we can solve for the β coefficients by inverting the matrix $\mathbf{S}^j_{1\ldots n}$. Given these coefficients, we write:

$$\langle l|(\hat{P}_i)^j|j\rangle = \langle l|(\hat{P}_i)^j|R\rangle + \sum_{k\neq j} \beta_k \langle l|(\hat{P}_i)^j|k\rangle, \qquad \text{for } l \neq j,$$

$$\langle j|(\hat{P}_i)^j|m\rangle = \langle R|(\hat{P}_i)^j|m\rangle + \sum_{k\neq j} \beta_k \langle k|(\hat{P}_i)^j|m\rangle, \qquad \text{for } m \neq j.$$

$$\langle j|(\hat{P}_i)^j|j\rangle = \langle R|(\hat{P}_i)^j|R\rangle + \sum_{k\neq j} \beta_k \langle k|(\hat{P}_i)^j|R\rangle$$

$$+ \sum_{k\neq j} \beta_k \langle R|(\hat{P}_i)^j|k\rangle + \sum_{k_1\neq j} \sum_{k_2\neq j} \beta_{k_1}\beta_{k_2} \langle k_1|(\hat{P}_i)^j|k_2\rangle.$$

Since $(\hat{P}_i)^j$ contains only the operators $P_1, \ldots, P_{j-1}, P_{j+1}, \ldots, P_n$, we have $(\hat{P}_i)^j|R\rangle = 0$. This leads to the equations:

$$\langle l|(\hat{P}_i)^j|j\rangle = \sum_{k\neq j}\beta_k\langle l|(\hat{P}_i)^j|k\rangle, \quad \text{for } l\neq j,$$

$$\langle j|(\hat{P}_i)^j|m\rangle = \sum_{k\neq j}\beta_k\langle k|(\hat{P}_i)^j|m\rangle, \quad \text{for } m\neq j,$$

$$\langle j|(\hat{P}_i)^j|j\rangle = \sum_{k_1\neq j}\sum_{k_2\neq j}\beta_{k_1}\beta_{k_2}\langle k_1|(\hat{P}_i)^j|k_2\rangle. \tag{37}$$

Given the values of $\langle l|(\hat{P}_i)^j|m\rangle$ for all $i, l, m \in \{1, \ldots, j-1, j+1, \ldots, n\}$, we can calculate all of the expressions in eq. (37), from which we can calculate all of the terms $\langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle$ in table 2, and thereby find $\langle l|(\hat{P}_i)_{1\ldots n}|m\rangle$ for all $i, l, m \leq n$. Thus, given the values for $\langle l|(\hat{P}_i)^j|m\rangle$ for all $j \leq n$ and $i, l, m \in \{1, \ldots, j-1, j+1, \ldots, n\}$ (a set of $(n-1)$-dimensional problems), we can calculate the values for $\langle l|(\hat{P}_i)_{1\ldots n}|m\rangle$ for all $i, l, m \leq n$ (an n-dimensional problem).

Let N be the set $\{1, \ldots, n\}$ and M some subset of N. Then, by the same reasoning as above, given the values for $\langle l|(\hat{P}_i)_{M-j}|m\rangle$ for all $j \in M$ and $i, l, m \in M - j$, we can calculate the values for $\langle l|(\hat{P}_i)_M|m\rangle$ for all $i, l, m \in M$.

This establishes a recursive algorithm for calculating an n-state probability. Set up an array $\mathbf{P}(i_{\text{state}}, i_{\text{subset}}, i_{\text{bra}}, i_{\text{ket}})$ in which the indices i_{state}, i_{bra}, and i_{ket} range from 1 to n, and the index i_{subset} ranges over some ordering of the $2n-1$ nonempty subsets of the set $\{1, \ldots, n\}$. A convenient ordering comes from the binary representation of each subset in which the included basis states are designated by a one and the excluded basis states are designated by a zero (for example, the subset $\{1, 4, 5\}$ of the set $\{1, 2, 3, 4, 5, 6\}$ is represented by $100110 = 38 = 2^{6-1} + 2^{6-4} + 2^{6-5}$). We proceed to calculate the entries of the $\mathbf{P}$ array so that:

$$\mathbf{P}(i, i_{\text{subset}}, l, m) = \langle l|(\hat{P}_i)_M|m\rangle,$$

where M is the subset corresponding to the index i_{subset}.

We begin with the null $\mathbf{P}$ array and proceed to calculate it as follows. For each 1-element subset $M = \{i\} \subseteq N$ and any $l, m \leq n$, we set $\mathbf{P}(i, 2^{n-i}, l, m) = S_{il}S_{im}$, since $(\hat{P}_i)_M = P_i$. Notice that we calculate values of the $\mathbf{P}$ array for which the bra or the ket are not in the set M, since they will be needed in the next step of the recursion.

Then, for each 2-element subset $M = \{i_1, i_2\}$, we evaluate $\mathbf{P}(i, 2^{n-i_1} + 2^{n-i_2}, l, m)$ for all sets $\{i \in M, l \in M, m \in M\}$ by referring to table 2, summing over both values of j in M, and dividing by 2. Note that the values for $\langle l|(\hat{P}_i)^j|m\rangle$ are in the $\mathbf{P}$ array since we have already calculated all of the 1-element subset entries. Next, using eq. (37), we calculate $\mathbf{P}(i, 2^{n-i_1} + 2^{n-i_2}, l, m)$ for the sets $\{i \in M, l \in M, m \notin M\}$, $\{i \in M, l \notin M, m \in M\}$, and $\{i \in M, l \notin M, m \notin M, l = m\}$. Note that these are exactly the entries we will need from table 2 when we proceed to 3-element subsets.

Continuing in this manner through subsets of increasing size, we finally compute the values of $\mathbf{P}(i, 2^n - 1, l, m) = \langle l | (\hat{P}_i)_{1 \ldots n} | m \rangle$ for all sets $\{i, l, m\}$. Then, by eq. (35), we calculate the final probability $(\mathcal{P}_i)_{1 \ldots n}$.

A program for this algorithm was written in FORTRAN 77 and was used for the calculations described in section 4. A calculation for seven basis states required approximately one CPU second on a Sun workstation [8].

4. Application – electronic populations of several molecules

In 1955, Mulliken [3] developed a method to calculate probabilities on non-orthogonal basis sets, with a particular application to electronic populations calculated from LCAO (Linear Combination of Atomic Orbitals) molecular orbitals. Mulliken's method consisted essentially of calculating overlap populations $S_{ij}\, \alpha_i\, \alpha_j$ for every pair of basis states $|i\rangle$ and $|j\rangle$, and assigning half of this overlap population to each of the two involved states. Using this method, however, Mulliken occasionally calculated probabilities to be slightly less than zero or greater than one: for example, Mulliken found a population of -0.0005 on H_2 in the $1a_1$ molecular orbital for H_2O [3].

Other researchers have refined Mulliken's method in an effort to avoid these difficulties [9–13]. In addition, a different approach to computing populations has since been proposed, involving the partitioning of the molecular Cartesian space and the integration of electron density within these partitions [14–18]. However, these integration methods often involve more complex computations. Several other approaches to population analysis have also been recently developed [19–21].

Using the algorithm set forth in section 3, we calculate electronic populations for various molecules using projection probabilities. A molecular orbital $|\psi\rangle$ of a given molecule is commonly expressed as a linear combination of the atomic orbitals of its constituent atoms:

$$|\psi\rangle = \sum_{i=1}^{n} \sum_{s=1}^{a(i)} \alpha_{i,s} |i, s\rangle,$$

where the index i ranges over the n atoms in the molecule, and the index s ranges over the $a(i)$ atomic orbitals for atom i. From a collection of such LCAO molecular wave functions [22], we know the values for the $\alpha_{i,s}$ coefficients and for the overlap matrix of the atomic orbitals $|i, s\rangle$. Next, we consolidate the atomic orbitals of a given atom into a single orbital, by defining:

$$|\chi_i\rangle = \sum_{s=1}^{a(i)} \alpha_{i,k} |i, k\rangle.$$

If we now define:

$$\alpha_i^2 = \langle \chi_i \mid \chi_i \rangle,$$

$$|i\rangle = \frac{|\chi_i\rangle}{\alpha_i},$$

then we have:

$$|\psi\rangle = \sum_{i=1}^{n} \alpha_i |i\rangle,$$

where the $|i\rangle$ are normalized. We can calculate the coefficients α_i of the linear combination and the overlap matrix of the basis vectors $|i\rangle$ as follows:

$$\alpha_i = \sqrt{\sum_{s=1}^{a(i)} \sum_{t=1}^{a(i)} \alpha_{i,s} \alpha_{i,t} \langle i, s \mid i, t \rangle},$$

$$\langle i \mid j \rangle = \frac{\langle \chi_i \mid \chi_j \rangle}{\alpha_i \alpha_j},$$

$$= \frac{\sum_{s=1}^{a(i)} \sum_{t=1}^{a(j)} \alpha_{i,s} \alpha_{j,t} \langle i, s \mid j, t \rangle}{\sqrt{\sum_{s=1}^{a(i)} \sum_{t=1}^{a(i)} \alpha_{i,s} \alpha_{i,t} \langle i, s \mid i, t \rangle} \sqrt{\sum_{s=1}^{a(j)} \sum_{t=1}^{a(j)} \alpha_{j,s} \alpha_{j,t} \langle j, s \mid j, t \rangle}}.$$

Given this information, we calculate a probability for each molecular orbital by the methods of section 3. Then, to derive an electronic population, we multiply this probability by 2, since each occupied molecular orbital contains two electrons. Note that the derivation of projected probabilities in section 2 guarantees that these populations lie between zero and two, thus avoiding an imperfection of Mulliken populations. Table 3 lists the populations calculated by this method for fifteen molecules, compared with the Mulliken populations for the same data [23]. The superscripts in the chemical formulas are used to differentiate atoms of the same element within a given molecule.

If we interpret the difference between the electronic population of an atom and its atomic number as a measure of the atom's effective charge in a given molecule, we notice first that the projection operator populations essentially follow the rules dictated by electronegativity considerations: partial positive charges generally reside on carbon atoms, and partial negative charges generally reside on oxygen, nitrogen and fluorine atoms. We note further that for HCN, NNO, and CH_3OH, the projection operator populations predict an alternation of partial positive and negative charges along the

Table 3

Electronic populations

Molecule : Atom	Projection operator population	Mulliken population
H_2O : H	0.8734	0.6142
: O	8.2531	8.7717
CH_4 : H	1.0256	0.8157
: C	5.8975	6.7374
CH_3F : H	1.0312	0.8258
: C	5.7539	6.1395
: F	9.1524	9.3831
CH_2F_2 : H	1.0409	0.8256
: C	5.5990	5.6725
: F	9.1596	9.3381
BH_3 : H	1.0784	0.9751
: B	4.7649	5.0746
NH_3 : H	0.8965	0.7114
: N	7.3104	7.8659
CO_2 : C	5.7224	5.4133
: O	8.1388	8.2933
H_2O_2 : H	0.8756	0.5839
: O	8.1244	8.4161
HCN : H	0.8983	0.7168
: C	6.1678	6.2302
: N	6.9339	7.0530
CF_2 : C	5.8114	5.4942
: F	9.0943	9.2529
N^1N^2O : N^1	7.0325	6.9687
: N^2	6.8252	6.7373
: O	8.1423	8.2940
FNO : N	6.7564	65409
: O	7.9247	8.0708
: F	9.3189	9.3883
H_2CO : H	1.0447	0.8545
: C	5.8748	5.9848
: O	8.0358	8.3062
$CH^1O^1O^2H^2$: H^1	1.0377	0.8071
: H^2	0.8578	0.5730
: C	5.8654	5.6970
: O^1	8.0847	8.3755
: O^2	8.1543	8.5473
$CH^1_3OH^2$: H^1	1.0316	0.8296
: H^2	0.8865	0.6173
: C	5.8537	6.2728
: O	8.1650	8.6212

molecule, whereas the Mulliken populations do not predict such an alternation. This alternation of partial charges conforms to the separation of charge guidelines used in determining resonance structures: that since partial positive or negative charges repel, they naturally seek to maximize the distance between them. This separation of charge guidelines is by no means an absolute rule; rather, it is one of several competing factors which may determine charge distribution. Hence, the fact that the separation of charge rule does not hold for the projection operator populations of FNO can be explained by the presence of the very electronegative fluorine atom counteracting the effect of the separation of charge rule.

In addition, in many molecules in which a hydrogen atom is bonded to a carbon atom (CH_4, CH_3OH, CHOOH, CH_3F, CH_2F_2, H_2CO), the projection operator population for hydrogen is slightly greater than one, whereas the Mulliken population is less than one. This surprising prediction of a small partial negative charge on hydrogen actually ensures that the alternation of charge rule holds in these molecules for the projection operator populations, where it does not hold for the Mulliken populations. In addition, this partial negative charge on hydrogen agrees with the findings of recent Bader population calculations [24]. Thus, the above results support the application of the concept of projection operator probabilities to electronic charge.

5. Discussion

We have developed a formalism relying on the theory of projection operators to compute probabilities for a linear combination of quantum states. We note that some earlier work on population analysis also utilized projection operators, though in a distinct manner [25]. The development of this projection operator formalism led to an iterative technique for computing probabilities, which easily lent itself to incorporation into a computer algorithm.

We proceeded to apply this theory of quantum probabilities to the problem of atomic populations in molecules. The results were often in agreement with the Mulliken population analysis, but in several cases differed significantly. Further studies of these differences, and comparisons with other recent population analysis procedures, are in order.

Appendix 1

PROPERTIES OF HYPERPLANE PROJECTION OPERATORS

The hyperplane projection operator $P_{1\ldots n}$ is defined as:

$$P_{1\ldots n} \equiv |1'\rangle\langle 1'| + |2'\rangle\langle 2'| + \ldots + |n'\rangle\langle n'|,$$

where $\{|1'\rangle, |2'\rangle, \ldots, |n'\rangle\}$ is an orthonormal basis for the space spanned by $|1\rangle, |2\rangle, \ldots, |n\rangle$, as described in section 2.3.

We first need to verify that this operator is well defined; i.e. it is independent of the choice of orthonormal basis. Take a second orthonormal basis $\{|1''\rangle, \ldots, |n''\rangle\}$ for the space spanned by $|1\rangle, \ldots, |n\rangle$, and write it in terms of the single-primed orthonormal basis:

$$[|1''\rangle, \ldots, |n''\rangle] = [|1'\rangle, \ldots, |n'\rangle]\mathbf{B}, \tag{A1.1}$$

for some $n \times n$ real matrix $\mathbf{B}$. Equivalently, we can write:

$$\begin{bmatrix} \langle 1''| \\ \vdots \\ \langle n''| \end{bmatrix} = \mathbf{B}^{\dagger} \begin{bmatrix} \langle 1'| \\ \vdots \\ \langle n'| \end{bmatrix}. \tag{A1.2}$$

Then, if we multiply eq. (A1.1) on the left-hand side by eq. (A1.2), we find that $\mathbf{I} = \mathbf{B}^{\dagger}\mathbf{I}\mathbf{B}$, which implies $\mathbf{B}^{\dagger} = \mathbf{B}^{-1}$. If we next multiply eq. (A1.2) on the left-hand side by eq. (A1.1), we find that:

$$\begin{aligned} |1''\rangle\langle 1''| + \ldots + |n''\rangle\langle n''| &= [|1'\rangle, \ldots, |n'\rangle]\,\mathbf{B}\,\mathbf{B}^{\dagger}[|1'\rangle, \ldots, |n'\rangle]^{\dagger} \\ &= [|1'\rangle, \ldots, |n'\rangle]\,\mathbf{B}\,\mathbf{B}^{-1}[|1'\rangle, \ldots, |n'\rangle]^{\dagger} \\ &= |1'\rangle\langle 1'| + \ldots + |n'\rangle\langle n'|. \end{aligned}$$

Next, having verified that $P_{1\ldots n}$ is well defined, we define its orthogonal complement by:

$$Q_{1\ldots n} = 1 - P_{1\ldots n}.$$

We can easily see that these operators obey the defining property of projection operators (that $(P_{1\ldots n})^2 = P_{1\ldots n}$ and $(Q_{1\ldots n})^2 = Q_{1\ldots n}$). Thus, we can manipulate these hyperplane projection operators in the same manner as elementary projection operators.

In addition, we can easily prove the following properties from the definition of the hyperplane projection operators:

(1) If $|\chi\rangle$ is in the span of $|1\rangle, \ldots, |n\rangle$, then $P_{1\ldots n}|\chi\rangle = |\chi\rangle$, and $Q_{1\ldots n}|\chi\rangle = 0$.

(2) If $|\chi\rangle$ is orthogonal to $|1\rangle, \ldots, |n\rangle$, then $P_{1\ldots n}|\chi\rangle = 0$, and $Q_{1\ldots n}|\chi\rangle = |\chi\rangle$.

From these properties, we can conclude that if $j \in \{1, \ldots, n\}$, then $P_{1\ldots n}P_j = P_j$. In fact, we could easily show the general property that if $\{i_1, \ldots, i_j\} \subseteq \{1, \ldots, n\}$, then $P_{1\ldots n}P_{i_1\ldots i_j} = P_{i_1\ldots i_j}$.

So far, the properties stated for hyperplane projection operators are intuitive if we consider $P_{1\ldots n}$ as the geometric projection onto the space spanned by $|1\rangle, \ldots, |n\rangle$. Next, we derive a non-intuitive but important property of these hyperplane projection operators. Given any vector $|\phi\rangle$, which may or may not lie in the span of $|1\rangle, \ldots, |n\rangle$, the following expression holds:

$$\langle\phi|Q_{1\ldots n}|\phi\rangle = \frac{\det \mathbf{S}_{1\ldots n\phi}}{\det \mathbf{S}_{1\ldots n}}, \tag{A1.3}$$

where the S matrices are the overlap matrices for the subscripted states. We prove this by inducton on n.

It is easy to verify that eq. (A1.3) holds for $n = 1$, in which case $Q_{1\ldots n} = Q_1$ and $\mathbf{S}_{1\ldots n}$ is the trivial matrix [1].

Now, assume eq. (A1.3) holds for $n-1$ and prove it for n. Note that if $\{|1'\rangle, \ldots, |n-1'\rangle\}$ is an orthonormal basis for the space spanned by $|1\rangle, \ldots, |n-1\rangle$, and if we define:

$$|n'\rangle = \frac{Q_{1\ldots n-1}|n\rangle}{\|Q_{1\ldots n-1}|n\rangle\|} = \frac{Q_{1\ldots n}|n\rangle}{\langle n|Q_{1\ldots n-1}|n\rangle^{1/2}},$$

then $\{|1'\rangle, \ldots, |n'\rangle\}$ is an orthonormal basis for the space spanned by $|1\rangle, \ldots, |n\rangle$. Thus, we write:

$$P_{1\ldots n} = P_{1\ldots n-1} + \frac{Q_{1\ldots n-1}P_nQ_{1\ldots n-1}}{\langle n|Q_{1\ldots n-1}|n\rangle},$$

$$Q_{1\ldots n} = Q_{1\ldots n-1} - \frac{Q_{1\ldots n-1}P_nQ_{1\ldots n-1}}{\langle n|Q_{1\ldots n-1}|n\rangle}.$$

Therefore,

$$\langle\phi|Q_{1\ldots n}|\phi\rangle = \langle\phi|Q_{1\ldots n-1}|\phi\rangle - \frac{\langle\phi|Q_{1\ldots n-1}|n\rangle^2}{\langle n|Q_{1\ldots n-1}|n\rangle}. \tag{A1.4}$$

Choose an orthonormal basis $\{|1^*\rangle, \ldots, |n^*\rangle, |\phi^*\rangle\}$ for the space spanned by $|1\rangle, \ldots, |n\rangle, |\phi\rangle$ so that in the coordinate system defined by this basis we have:

$$|2\rangle = (a_{21}, a_{22}, 0, 0, \ldots, 0),$$
$$\vdots$$
$$|n\rangle = (a_{n1}, a_{n2}, \ldots, a_{nn}, 0),$$
$$|\phi\rangle = (a_{\phi 1}, a_{\phi 2}, \ldots, a_{\phi n}, a_{\phi\phi}). \tag{A1.5}$$

In particular,

$$|\phi\rangle = a_{\phi 1}|1^*\rangle + a_{\phi 2}|2^*\rangle + \ldots + a_{\phi n}|n^*\rangle + a_{\phi\phi}|\phi^*\rangle.$$

We know that $Q_{1\ldots n-1}|1^*\rangle = \ldots = Q_{1\ldots n-1}|n-1^*\rangle = 0$, since $|1^*\rangle, \ldots, |n-1^*\rangle$ lie in the space spanned by $|1\rangle, \ldots, |n-1\rangle$. Hence, we can further write:

$$Q_{1\ldots n-1}|\phi\rangle = a_{\phi n} Q_{1\ldots n-1}|n^*\rangle + a_{\phi\phi} Q_{1\ldots n-1}|\phi^*\rangle.$$

Since $|\phi^*\rangle$ is perpendicular to each of $|1\rangle, \ldots, |n-1\rangle$, we know that $Q_{1\ldots n-1}|\phi^*\rangle = |\phi^*\rangle$, which leads to the simplification:

$$Q_{1\ldots n-1}|\phi\rangle = a_{\phi n} Q_{1\ldots n-1}|n^*\rangle + a_{\phi\phi}|\phi^*\rangle.$$

Now, since $|n\rangle$ has coordinate 0 in the $|\phi^*\rangle$ direction, $\langle n|\phi^*\rangle = 0$, and we have:

$$\langle n|Q_{1\ldots n-1}|\phi\rangle = a_{\phi n}\langle n|Q_{1\ldots n-1}|n^*\rangle. \tag{A1.6}$$

Referring back to eq. (A1.5), we find the expression:

$$|n\rangle = a_{n1}|1^*\rangle + a_{n2}|2^*\rangle + \ldots + a_{nn}|n^*\rangle.$$

Again, since $Q_{1\ldots n-1}|1^*\rangle = \cdots = Q_{1\ldots n}|n-1^*\rangle = 0$, we have:

$$Q_{1\ldots n-1}|n\rangle = a_{nn} Q_{1\ldots n-1}|n^*\rangle,$$
$$\langle n|Q_{1\ldots n-1}|n\rangle = a_{nn}\langle n|Q_{1\ldots n-1}|n^*\rangle.$$

Since a_{nn} must be nonzero in order to preserve the linear independence of the basis vectors $|1\rangle, \ldots, |n\rangle$, we can rewrite eq. (A1.6) as:

$$\langle n \,|\, Q_{1\ldots n-1} \,|\, \phi \rangle = \frac{a_{\phi n}}{a_{nn}} \langle n \,|\, Q_{1\ldots n-1} \,|\, n \rangle.$$

The results of the above section allow us to simplify eq. (A1.4) to:

$$\langle \phi \,|\, Q_{1\ldots n} \,|\, \phi \rangle = \langle \phi \,|\, Q_{1\ldots n-1} \,|\, \phi \rangle - \frac{a_{\phi n}^2}{a_{nn}^2} \langle n \,|\, Q_{1\ldots n-1} \,|\, n \rangle.$$

However, by induction, we have assumed eq. (A1.3) for $n-1$ states, so we can write:

$$\langle \phi \,|\, Q_{1\ldots n} \,|\, \phi \rangle = \frac{\det \mathbf{S}_{1\ldots n-1,\phi}}{\det \mathbf{S}_{1\ldots n-1}} - \frac{a_{\phi n}^2}{a_{nn}^2} \frac{\det \mathbf{S}_{1\ldots n}}{a_{nn}^2 \det \mathbf{S}_{1\ldots n-1}}. \tag{A1.7}$$

Next, we evaluate the determinants in eq. (A1.7). We show in detail the evaluations of $\det \mathbf{S}_{1\ldots n}$, since the same basic method can be used for the other determinants.

We know from eq. (A1.5) that:

$$\begin{bmatrix} \langle 1 | \\ \vdots \\ \langle n | \end{bmatrix} = \mathbf{A} \begin{bmatrix} \langle 1^* | \\ \vdots \\ \langle n^* | \end{bmatrix},$$

$$[|1\rangle, \ldots, |n\rangle] = [|1^*\rangle, \ldots | n^*\rangle]\mathbf{A}^\dagger,$$

where $\mathbf{A}$ is the square matrix:

$$\begin{bmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ a_{21} & a_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{bmatrix}.$$

Multiplying the above vector equations (in the order given) yields:

$$\mathbf{S}_{1\ldots n} = \mathbf{A}\mathbf{I}\mathbf{A}^\dagger = \mathbf{A}\mathbf{A}^\dagger.$$

Therefore,

$$\det \mathbf{S}_{1\ldots n} = \det(\mathbf{A}\mathbf{A}^\dagger) = (\det \mathbf{A})(\det \mathbf{A}^\dagger) = (\det \mathbf{A})^2,$$

by simple rules of determinants. Since **A** is a triangular matrix, its determinant is simply the product of its diagonal elements, so that:

$$\det \mathbf{S}_{1\ldots n,\phi} = a_{11}^2 a_{22}^2 \ldots a_{nn}^2 a_{\phi\phi}^2 .$$

In order to calculate $\det \mathbf{S}_{1\ldots n-1,\phi}$, we need to set up a new coordinate system. We already know that $\{|1^*\rangle, \ldots, |n-1^*\rangle\}$ is an orthonormal basis for the space spanned by $|1\rangle, \ldots, |n-1\rangle$. Then, we can expand this basis to a basis for the space spanned by $|1\rangle, \ldots, |n-1\rangle, |\phi\rangle$, in which we have

$$\begin{aligned}
|1\rangle &= (a_{11}, 0, 0, \ldots, 0), \\
|2\rangle &= (a_{21}, a_{22}, 0, 0, \ldots, 0), \\
&\vdots \\
|n-1\rangle &= (a_{n-1,1}, a_{n-1,2}, \ldots, a_{n-1,n-1}, 0), \\
|\phi\rangle &= (a_{\phi 1}, a_{\phi 2}, \ldots, a_{\phi,n-1}, b_{\phi\phi}).
\end{aligned} \tag{A1.8}$$

From this representation, we use the same method we used for $\det \mathbf{S}_{1\ldots n}$ to show that:

$$\det \mathbf{S}_{1\ldots n-1,\phi} = a_{11}^2 a_{22}^2 \ldots a_{n-1,n-1}^2 b_{\phi\phi}^2 .$$

Now, if we equate the expressions we obtain for $\langle\phi|\,\phi\rangle$ from eq. (A1.5) and eq. (A1.8), we find:

$$b_{\phi\phi}^2 = a_{\phi n}^2 + a_{\phi\phi}^2 .$$

Thus, we obtain the expression:

$$\det \mathbf{S}_{1\ldots n-1,\phi} = a_{11}^2 a_{22}^2 \ldots a_{n-1,n-1}^2 (a_{\phi n}^2 + a_{\phi\phi}^2).$$

Now, we can plug these results into eq. (A1.7) as follows:

$$\langle\phi|Q_{1\ldots n}|\phi\rangle = \frac{\det \mathbf{S}_{1\ldots n-1,\phi}}{\det \mathbf{S}_{1\ldots n-1}} - \frac{a^2_{\phi n}}{a^2_{nn}} \frac{\det \mathbf{S}_{1\ldots n}}{\det \mathbf{S}_{1\ldots n-1}}$$

$$= \frac{a^2_{11}a^2_{22}\ldots a^2_{n-1,n-1}(a^2_{\phi n} + a^2_{\phi\phi})}{a^2_{11}a^2_{22}\ldots a^2_{n-1,n-1}} - \frac{a^2_{\phi n}}{a^2_{nn}} \frac{a^2_{11}a^2_{22}\ldots a^2_{nn}}{a^2_{11}a^2_{22}\ldots a^2_{n-1,n-1}}$$

$$= a^2_{\phi n} + a^2_{\phi\phi} - \frac{a^2_{\phi n}}{a^2_{nn}} a^2_{nn}$$

$$= a^2_{\phi\phi}$$

$$= \frac{\det \mathbf{S}_{1\ldots n,\phi}}{\det \mathbf{S}_{1\ldots n}} .$$

Thus, we have proven eq. (A1.3) by induction.

Notice that we have shown:

$$\langle\phi|Q_{1\ldots n}|\phi\rangle = \frac{\det \mathbf{S}_{1\ldots n,\phi}}{\det \mathbf{S}_{1\ldots n}} = a^2_{\phi\phi} .$$

In order for $|1\rangle, \ldots, |n\rangle, |\phi\rangle$ to be linearly independent, we must have $a_{\phi\phi} \neq 0$ (cf. eq. (A1.5)). In addition, if we assume that $|\phi\rangle$ is normalized (as we generally do), then we must have $a^2_{\phi\phi} \leq 1$. Thus, if $|1\rangle, \ldots, |n\rangle, |\phi\rangle$ are linearly independent and normalized, then $0 < \langle\phi|Q_{1\ldots n}|\phi\rangle \leq 1$.

Appendix 2

CALCULATING PROBABILITIES FROM PROBABILITY OPERATORS

This appendix details the essential algebra for calculating a formula for probabilities given the probability operators derived in sections 2.2 (for the 3-state case) and 2.3 (for the n-state case). Note first that the 3-state probability operators in eqs. (15) and (19) are, as expected, identical to the n-state probability operators in eqs. (28) and (32) evaluated at $n = 3$. Hence, we can work out the algebra for the general n-state probabilities, and then insert the requirement that $n = 3$ to arrive at the specific 3-state results.

Consider the n-state probability operator formulas from eqs. (28) and (32):

$$(\hat{P}_{jj})_{1\ldots n} = \frac{1}{2}\left[P_j + \frac{Q^j P_j Q^j}{1-(X_j)^2_{1\ldots n}} + \frac{Q_j Q^j P_j Q^j Q_j}{1-(X_j)^2_{1\ldots n}}\right],$$

$$(\hat{P}_{ij})_{1\ldots n} = \frac{1}{2}\left[(\hat{P}_i)^j + Q_j(\hat{P}_i)^j Q_j + \frac{Q^j P_j (\hat{P}_i)^j Q^j}{1-(X_j)^2_{1\ldots n}} + \frac{Q_j Q^j P_j (\hat{P}_i)^j P_j Q^j Q_j}{1-(X_j)^2_{1\ldots n}}\right], \quad i \neq j,$$

together with the expression from eq. (36):

$$(\mathcal{P}_i)_{1\ldots n} = \frac{1}{n}\sum_{j=1}^{n}\sum_{l=1}^{n}\sum_{m=1}^{n}\alpha_l \alpha_m \langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle.$$

In order to compute the probabilities $(\mathcal{P}_i)_{1\ldots n}$, it is sufficient to calculate $\langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle$ for all $j, l, m \leq n$. First, we calculate $\langle l|(\hat{P}_{jj})_{1\ldots n}|m\rangle$ from the expression in eq. (28):

Case 1: $l, m \neq j$ $(Q^j|l\rangle = Q^j|m\rangle = 0)$.

$$\begin{aligned}
\langle l|(\hat{P}_{jj})_{1\ldots n}|m\rangle &= \frac{1}{2}\left[S_{jl}S_{jm} + \frac{\langle l|Q_j Q^j P_j Q^j Q_j|m\rangle}{1-(X_j)^2_{1\ldots n}}\right] \\
&= \frac{1}{2}\left[S_{jl}S_{jm} + \frac{\langle l|P_j Q^j P_j Q^j P_j|m\rangle}{1-(X_j)^2_{1\ldots n}}\right] \\
&= \frac{1}{2}\left[S_{jl}S_{jm} + S_{jl}S_{jm}\frac{\langle j|Q^j P_j Q^j|j\rangle}{1-(X_j)^2_{1\ldots n}}\right] \\
&= \frac{1}{2}\left[S_{jl}S_{jm} + S_{jl}S_{jm}\frac{(1-(X_j)^2_{1\ldots n}}{1-(X_j)^2_{1\ldots n}}\right] \\
&= \frac{1}{2}\left[S_{jl}S_{jm} + S_{jl}S_{jm}\frac{1-(X_j)_{1\ldots n}}{1+(X_j)_{1\ldots n}}\right] \\
&= \frac{1}{2}\left[\frac{2S_{jl}S_{jm}}{1+(X_j)_{1\ldots n}}\right] \\
&= \frac{S_{jl}S_{jm}}{1+(X_j)_{1\ldots n}}.
\end{aligned}$$

Case 2: $l = j,\ m \neq j\ \ (Q_j \mid l\rangle = Q^j \mid m\rangle = 0).$

$$\langle l \mid (\hat{P}_{jj})_{1\ldots n} \mid m\rangle = \tfrac{1}{2} S_{jm}\,.$$

Case 3: $l \neq j,\ m = j\ \ (Q^j \mid l\rangle = Q_j \mid m\rangle = 0).$

$$\langle l \mid (\hat{P}_{jj})_{1\ldots n} \mid m\rangle = \tfrac{1}{2} S_{jl}\,.$$

Case 4: $l, m = j\ \ (Q_j \mid l\rangle = Q_j \mid m\rangle = 0).$

$$\begin{aligned}
\langle l \mid (\hat{P}_{jj})_{1\ldots n} m\rangle &= \frac{1}{2}\left[1 + \frac{\langle j \mid Q^j P_j Q^j \mid j\rangle}{1 - (X_j)^2_{1\ldots n}}\right] \\
&= \frac{1}{2}\left[1 + \frac{(1 - (X_j)_{1\ldots n})^2}{1 - (X_j)^2_{1\ldots n}}\right] \\
&= \frac{1}{2}\left[1 + \frac{1 - (X_j)_{1\ldots n}}{1 + (X_j)_{1\ldots n}}\right] \\
&= \frac{1}{2}\left[\frac{2}{1 + (X_j)_{1\ldots n}}\right] \\
&= \frac{1}{1 + (X_j)_{1\ldots n}}\,.
\end{aligned}$$

Next, we calculate $\langle l \mid (\hat{P}_{ij})_{1\ldots n} \mid m\rangle$ for $i \neq j$ from the expression in eq. (32):

Case 1: $l, m \neq j\ \ (Q^j \mid l\rangle = Q^j \mid m\rangle = 0).$

$$\begin{aligned}
&\langle l \mid (\hat{P}_{ij})_{1\ldots n} \mid m\rangle \\
&\quad = \frac{1}{2}\left[\langle l \mid (\hat{P}_i)^j \mid m\rangle + \langle l \mid Q_j (\hat{P}_i)^j Q_j \mid m\rangle + \frac{\langle l \mid Q_j Q^j P_j (\hat{P}_i)^j P_j Q^j Q_j \mid m\rangle}{1 - (X_j)^2_{1\ldots n}}\right] \\
&\quad = \frac{1}{2}\left[\langle l \mid (\hat{P}_i)^j \mid m\rangle + \langle l \mid Q_j (\hat{P}_i)^j Q_j \mid m\rangle + S_{jl} S_{jm} \frac{\langle j \mid Q^j P_j (\hat{P}_i)^j P_j Q^j \mid j\rangle}{1 - (X_j)^2_{1\ldots n}}\right]
\end{aligned}$$

$$= \frac{1}{2}\left[\langle l|(\hat{P}_i)^j|m\rangle + \langle l|Q_j(\hat{P}_i)^j Q_j|m\rangle + S_{jl}S_{jm}\frac{\langle j|(\hat{P}_i)^j|j\rangle(1-(X_j)_{1\ldots n})^2}{2-(X_j)^2_{1\ldots n}}\right]$$

$$= \frac{1}{2}\left[\begin{array}{l}\langle l|(\hat{P}_i)^j m\rangle + \langle l|(\hat{P}_i)^j|m\rangle - S_{jl}\langle j|(\hat{P}_i)^j|m\rangle - S_{jm}\langle l|(\hat{P}_i)^j|j\rangle + \\ S_{jl}S_{jm}\langle j|(\hat{P}_i)^j|j\rangle + S_{jl}S_{jm}\dfrac{\langle j|(\hat{P}_i)^j|j\rangle(1-(X_j)_{1\ldots n}}{1+(X_j)_{1\ldots n}}\end{array}\right]$$

$$= \langle l|(\hat{P}_i)^j|m\rangle - \frac{S_{jl}}{2}\langle j|(\hat{P}_i)^j|m\rangle - \frac{S_{jm}}{2}\langle l|(\hat{P}_i)^j|j\rangle + \frac{S_{jl}S_{jm}\langle j|(\hat{P}_i)^j|j\rangle}{1+(X_j)_{1\ldots n}}.$$

Case 2: $l = j,\ m \neq j\ \ (Q_j|l\rangle = Q^j|m\rangle = 0)$.

$$\langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle = \frac{\langle j|(\hat{P}_i)^j|m\rangle}{2}.$$

Case 3: $l \neq j,\ m = j\ \ (Q^j|l\rangle = Q_j|m\rangle = 0)$.

$$\langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle = \frac{\langle l|(\hat{P}_i)^j|j\rangle}{2}.$$

Case 4: $l, m = j\ \ (Q_j|l\rangle = Q_j|m\rangle = 0)$.

$$\langle l|(\hat{P}_{ij})_{1\ldots n}|m\rangle = \frac{1}{2}\left[\langle j|(\hat{P}_i)^j|j\rangle + \frac{\langle j|Q^j P_j(\hat{P}_i)^j P_j Q^j|j\rangle}{1-(X_j)^2_{1\ldots n}}\right]$$

$$= \frac{1}{2}\left[\langle j|(\hat{P}_i)^j|j\rangle + \frac{\langle j|(\hat{P}_i)^j|j\rangle(1-(X_j)_{1\ldots n})^2}{1-(X_j)^2_{1\ldots n}}\right]$$

$$= \frac{1}{2}\left[\langle j|(\hat{P}_i)^j|j\rangle + \frac{\langle j|(\hat{P}_i)^j|j\rangle(1-(X_j)_{1\ldots n})}{1+(X_j)_{1\ldots n}}\right]$$

$$= \frac{1}{2}\left[\frac{2\langle j|(\hat{P}_i)^j|j\rangle}{1+(X_j)_{1\ldots n}}\right]$$

$$= \frac{\langle j|(\hat{P}_i)^j|j\rangle}{1+(X_j)_{1\ldots n}}.$$

From these results, we can easily deduce the specific 3-state results by substituting $(X_j)_{123}$ for $(X_j)_{1\ldots n}$ and $(\hat{P}_i)_{ik}$ for $(\hat{P}_i)^j$.

Note that in the 3-state case, the values for $\langle l|(\hat{P}_{ij})_{123}|m\rangle$ involve the 2-state probability operator $(\hat{P}_i)_{ik}$. A closed-form expression was found for $(\hat{P}_i)_{ik}$ in eq. (6):

$$(\hat{P}_i)_{ik} = (1 - S_{ik}^4)^{-1}\left[P_i + S_{ik}^2 P_k \frac{1 + S_{ik}^2}{2}(P_i P_k + P_k P_i)\right],$$

from which we can verify the following relations:

$$\langle i|(\hat{P}_i)_{ik}|i\rangle = \frac{1}{1 + S_{ik}^2},$$

$$\langle i|(\hat{P}_i)_{ik}|k\rangle = \frac{S_{ik}}{2},$$

$$\langle k|(\hat{P}_i)_{ik}|k\rangle = \frac{S_{ik}^2}{1 + S_{ik}^2},$$

$$\langle j|(\hat{P}_i)_{ik}|i\rangle = \frac{2S_{ij} + S_{ik}(S_{ij}S_{ik} - S_{jk})}{2(1 + S_{ik}^2)} = \frac{2S_{ij} + S_{ik}D_{jk}}{2(1 + S_{ik}^2)},$$

$$\langle j|(\hat{P}_i)_{ik}|k\rangle = \frac{S_{ik}(S_{ij} + S_{ik}S_{jk})}{2(1 + S_{ik}^2)},$$

$$\langle j|(\hat{P}_i)_{ik}|j\rangle = \frac{S_{ij}^2 + S_{ik}^2 S_{jk}^2 - S_{ij}S_{ik}S_{jk} - S_{ij}S_{ik}^3 S_{jk}}{1 - S_{ik}^4} = \frac{S_{ij}D_{ij} + S_{ik}^2 S_{jk}D_{jk}}{1 - S_{ik}^4},$$

where $D_{ij} \equiv S_{ik}S_{jk} - S_{ij}$. These relations, when substituted into the expressions for $\langle l|(\hat{P}_{ij})_{123}|m\rangle$ and eq. (36), give the final expression:

$$(\mathcal{P}_i)_{123} = \frac{1}{3}\sum_{j=1}^{3}(\mathcal{P}_j)_{123},$$

where

$$(\mathcal{P}_{jj})_{123} = \frac{\alpha_j^2 + S_{ij}^2\,\alpha_i^2 + S_{jk}^2\,\alpha_k^2 + 2S_{ij}\,S_{jk}\,\alpha_i\,\alpha_k}{1 + (X_j)_{123}} + S_{ij}\,\alpha_i\,\alpha_j + S_{jk}\,\alpha_j\,\alpha_k\,,$$

$$(\mathcal{P}_{ij})_{123}$$

$$= \frac{1}{1+S_{ik}^2}\left[\begin{array}{l}\Big(Y_j(S_{ij}D_{ij} + S_{ik}^2 S_{jk}D_{jk})\Big)\alpha_j^2 + \\ \Big(1 - S_{ij}^2 + \tfrac{1}{2}S_{ij}S_{ik}D_{jk} + Y_j S_{ij}^2(S_{ij}D_{ij} + S_{ik}^2 S_{jk}D_{jk})\Big)\alpha_i^2 + \\ \Big(S_{ik}^2 - \tfrac{1}{2}S_{ij}S_{ik}S_{jk} - \tfrac{1}{2}S_{ik}^2 S_{jk}^2 + Y_j S_{jk}^2(S_{ij}D_{ij} + S_{ik}^2 S_{jk}D_{jk})\Big)\alpha_k^2 + \\ \Big(S_{ij} + \tfrac{1}{2}S_{ik}D_{jk}\Big)\alpha_i\,\alpha_j + \Big(\tfrac{1}{2}S_{ij}S_{ik} + \tfrac{1}{2}S_{ik}^2 S_{jk}\Big)\alpha_j\,\alpha_k - \\ \left(\begin{array}{l}D_{ik}(1+S_{ik}^2) + \tfrac{1}{2}S_{ik}(S_{ij}^2 + S_{jk}^2) + \\ 2Y_j S_{ij}S_{jk}(S_{ij}D_{ij} + S_{ik}^2 S_{jk}D_{jk}\end{array}\right)\alpha_i\,\alpha_k\end{array}\right]$$

for $i \neq j$, where

$$Y_j = \frac{1}{(1-S_{ik}^2)(1+(X_j)_{123})}\,.$$

References

[1] N. De Leon and S.P. Neshyba, Chem. Phys. Lett. 151(1988)296.

[2] For the purpose of this paper, the coefficients α_i and the overlaps $S_{ij} = \langle i\,|\,j\rangle$ are assumed to be real. The results could easily be extended to complex coefficients and overlaps.

[3] R.S. Mulliken, J. Chem. Phys. 23(1955)1833.

[4] C. Hollister and O. Sinanŏglu, J. Amer. Chem. Soc. 88(1966)13.

[5] S.P. Neshyba and N. De Leon, J. Chem. Phys. 99(1989)7772.

[6] It has been noted that this formula provides a nonzero result for $\mathcal{P}_1$ even when $|\psi\rangle$ is orthogonal to $|1\rangle$, in apparent contradiction with "orthodox" quantum mechanical intuition. This apparent paradox can be resolved by noting that if a zero result for $\mathcal{P}_1$ were derived in agreement with the orthodox intuition, we would have to conclude that $\mathcal{P}_2 = 1$, even though $|\psi\rangle$ is not equal to $|2\rangle$.

[7] A. Messiah, *Quantum Mechanics*, Vol. 1 (North-Holland, Amsterdam, 1960), p. 262.

[8] A copy of the program to calculate n-state probabilities is available from the authors upon request.

[9] P.O. Löwdin, Adv. Chem. Phys. 2(1959)207.

[10] E.R. Davidson, J. Chem. Phys. 46(1967)3320.

[11] R.E. Christoffersen and K.A. Baker, Chem. Phys. Lett. 8(1971)4.

[12] I. Mayer, Chem. Phys. Lett. 97(1983)270.

[13] An additional reference list is available in the paper by K.E. Edgecombe and R.J. Boyd, J. Chem. Soc. Faraday Trans. 2, 83(1987)1307.

[14] P. Politzer and R.R. Harris, J. Amer. Chem. Soc. 92(1970)6451.

[15] R.F.W. Bader, Acc. Chem. Res. 18(1985)9.
[16] F.W. Biegler-König, R.F.W. Bader and T.-H. Tang, J. Comp. Chem. 3(1982)317.
[17] F.L. Hirshfield, Isr. J. Chem. 16(1977)198.
[18] An additional reference list of integration procedures is available in the paper by J.B. Collins and A. Streitwieser, Jr., J. Comp. Chem. 1(1980)81.
[19] A.E. Reed, R.B. Weinstock and F. Weinhold, J. Chem. Phys. 83(1985)735.
[20] L.E. Chirlian and M.M. Frabcl, J. Comp. Chem. 8(1987)894.
[21] J. Cioslowski, J. Amer. Chem. Soc. 111(1989)8333.
[22] L.C. Snyder and H. Basch, *Molecular Wave Functions and Properties* (Wiley, New York, 1972).
[23] The Mulliken populations are taken from ref. [11], the source of the molecular orbitals, and were not calculated by consolidating atomic orbitals as described above; instead, a population was calculated for each atomic orbital, and then these populations were accumulated by atom.
[24] K.B. Wiberg, R.F.W. Bader and C.D.H. Lai, J. Amer. Chem. Soc. 109(1987)1004.
[25] See, for example, K.R. Roby, Mol. Phys. 27(1974)81.