

Predicting densities of nitrocubanes using partial orders

Guillermo Restrepo · Douglas J. Klein

Received: 12 April 2011 / Accepted: 26 April 2011 / Published online: 6 May 2011
© Springer Science+Business Media, LLC 2011

Abstract The nitrocubane substitution network shows the pattern by which single nitro-substitutions starting from cubane lead ultimately to octanitrocubane. Of the 22 member species, densities have been experimentally measured for only 8, and we utilise these to estimate the densities of the other 14 members. The network is viewed as a partially ordered set (or *poset*), which is used for posetic interpolative methods, which in turn differ from conventional QSAR/QSPR procedures based directly on molecular structures. The posetic methods utilised are the “cluster-expansion” and the “splinoid” ones, which turn out to be stable procedures as assessed by the leave-one-out cross-validation technique.

Keywords Nitrocubanes · Density estimation · Posets · Cluster expansion · Splinoid

1 Introduction

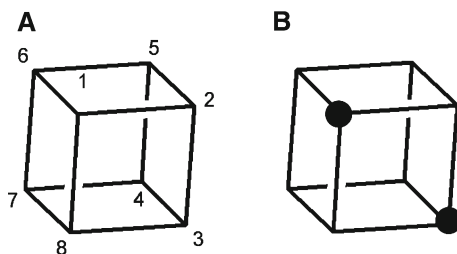
Because of its exceptional structure, strain, and symmetry, cubane (Fig. 1a) is a landmark in organic chemistry [1]. After its synthesis by Eaton and Cole [2] in 1964, a wealth of research has been done to that extent that in 1992, it was possible to state that a wide range of combinations of functional groups permuted to different positions on the cubane skeleton could be achieved by fairly straight-forward chemistry [1].

Electronic supplementary material The online version of this article (doi:[10.1007/s10910-011-9834-3](https://doi.org/10.1007/s10910-011-9834-3)) contains supplementary material, which is available to authorized users.

G. Restrepo · D. J. Klein
Department of Marine Sciences, Texas A&M University, Galveston, TX, USA

G. Restrepo (✉)
Laboratorio de Química Teórica, Universidad de Pamplona, Pamplona, Colombia
e-mail: guillermorestrepo@gmail.com

Fig. 1 Three-dimensional molecular structure of *cubane* with hydrogens suppressed: **a** its IUPAC numbering [3]; and **b** an example representation, for 1,3-disubstituted cubane



Cubane's density (1.29 g cm^{-3}) and heat of formation ($+159 \text{ kcal mol}^{-1}$) are higher than the corresponding values of other hydrocarbons of similar molecular weight. In addition, it is kinetically stable (decomposition only above 220°C). These facts led scientists at the US Army Armament Research and Development Center to propose Octanitrocubane (ONC) as a potential new explosive [4]. It is known that among the different aspects affecting the "effectiveness" of an explosive, the energetics of the decomposition reaction and the number of moles and molecular weight of the gaseous products are critical factors, as well as the density. The more moles of an explosive that can be packed into the limited volume of a shell the better [5]. It has been determined that for a given explosive, the detonation pressure is proportional to the square of its density, so great effort has been made to obtain the highest density form of any particular explosive [4]. Even a small increase in density may have a major impact on performance [6].

The enthalpy of formation is considered to predict explosive performance [7], and, in general, it is possible to state that of two substances with same composition, the one with the higher magnitude (negative) heat of formation will probably release more energy. Other things being equal, the one with higher density will probably be a superior explosive [6]. The typical carbon-based molecules contain one or more nitro groups, with few exceptions, *e.g.* the perchlorates, azides, & peroxides. In the detonation of those nitro-compounds, the nitro group provides (at least in part) oxygen to convert the molecule's hydrogen to water and its carbon to CO_2 or CO . In a perfect explosion the nitrogen atom of the nitro group would [6] be converted N_2 . With such considerations, the race for the synthesis of nitro-cubanes began.

By 2000, all nitro-cubanes were synthesized [8], with the octa-substituted derivative (ONC) the most challenging. Although the latest and most sophisticated calculations predict a density [9] of 2.135 or 2.123 g cm^{-3} for the most stable polymorph of ONC, the current synthesis of this substance leads to a crystal form [9] of 1.979 g cm^{-3} . So far, as Eaton points out, "there is no magic recipe for obtaining" the densest polymorph [4]. However, the current ONC polymorph has a very high detonation heat, and very good detonation properties, which make it a potential important powerhouse, but with a disadvantage that it is not so stable as to be immediately adopted as an explosive material [8]. However, as ONC has no hydrogen, no water forms when it burns; therefore, if used as propellant, it would leave little or no visible smoke (steam) in the plume behind the rocket [4], which is an advantage for some military purposes.

All in all, from the 21 nitrocubane derivatives the density has been experimentally determined, normally by x-ray diffraction studies, for seven of them. Owing to the

importance of this property, to the lack of information, and to the implicit danger of handling these substances in a laboratory, it is of special interest to estimate the densities of the other cubane nitroderivatives [6]. Here we report density estimates using so-called “posetic” interpolative methods, which are briefly described in the following sections.

2 Substitution reaction networks

Progressive reaction networks in which substituents are introduced step-by-step on a molecular skeleton abound in chemistry [10, 11]. In mathematical terms, those networks correspond to a “partial ordering”. That is, if molecular structure ξ can be obtained by substitutive reaction from structure ζ , then we write $\zeta \succeq \xi$, and identify \succeq as a *partial ordering* (which satisfies: first, $\xi \succeq \xi$; second, if $\zeta \succeq \xi$ and $\xi \succeq \zeta$, then $\xi = \zeta$; and third, if $\zeta \succeq \xi$ and $\xi \succeq \eta$, then $\zeta \succeq \eta$) [12]. Such a partially ordered set or *poset* may be depicted by a directed graph called a *Hasse diagram*, which relates pairs satisfying the inequality minimally, *i.e.* in our case relates the pairs obtained through a single substitution. In Fig. 2 we show this diagram for the cubane nitro-substitution reaction network, of which a sizable fragment has earlier been given in an incidental fashion to depict the organisation for synthesis [5]. We use an iconic representation of substituted cubanes as exemplified in Fig. 1b, where the presence of a substituent, such as a nitro group, is indicated by the replacement of a site (a vertex on the figure) by a black. Each of these cubane derivatives, including cubane, then appears as depicted in Fig. 2. Besides showing substitution reaction paths (as sequences of arrows strictly in the downward direction), a Hasse diagram can be used to interpolate properties of the substances appearing therein. Different interpolative posetic methodologies have been developed and several examples of applications have been reported [13–15]. This kind of predictive strategy, based upon the structure of the reaction network rather than on the structure of single molecules, which is customary in QSPR/QSAR studies, has been called quantitative superstructure/activity (property) relationship QSSAR/QSSPR [13–15], where the prefix “super” accounts for the holistic viewpoint of the whole network.

3 Posetic interpolative methods

Three posetic interpolative techniques have been described: Average-poset [11], cluster-expansion [16–19], and splinoid [13, 20]. In a Hasse diagram of a substitution-reaction network, each substance (except the top one) can be obtained from the chemicals on the previous level that send incoming arrows to the substance in question. By the same token that substance generates others located at the next level which are those receiving the on-going arrows from the substance studied (except for the bottom one).

The *average interpolation* method computes the property $Z(\xi)$ of a substance ξ as the average of averages [11] of preceding & succeeding members in the Hasse diagram. That is, one takes: first, the average of those compounds ζ directly leading to ξ ($\zeta \succeq \xi$) in the network; second, the average of those ζ directly following ξ ($\xi \succeq \zeta$); and third, averages these two averages. Although of utter simplicity, the method has the

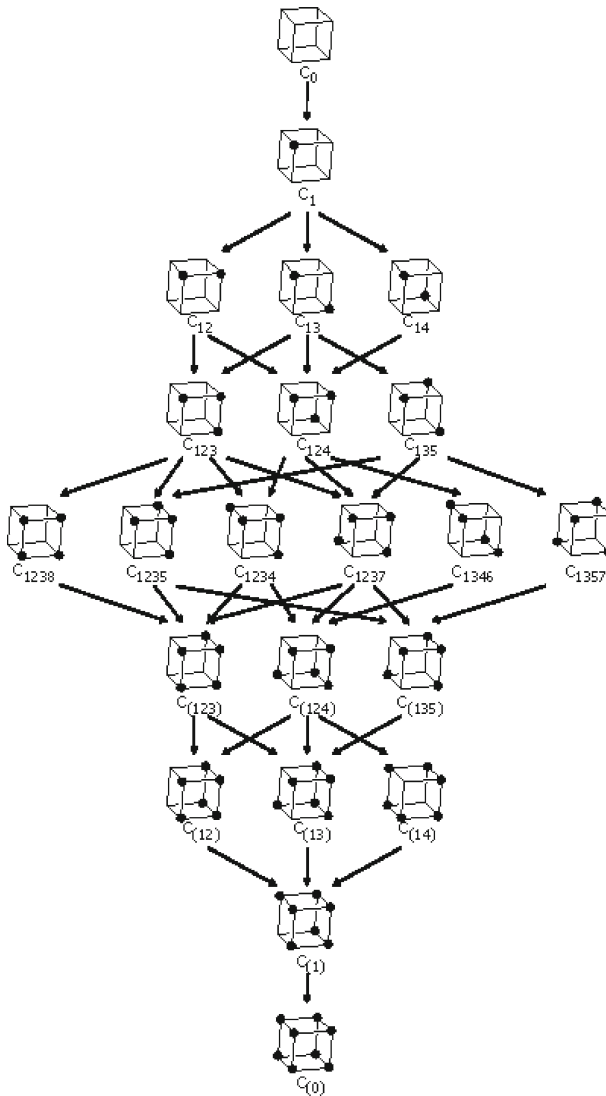


Fig. 2 Substitution reaction network of nitrocubanes, each of which is more abbreviatedly labelled by c_s , with s either the string of sites nitro-substituted or, if s is parentetic, those sites not nitro-substituted

limitation of needing knowledge of properties of all the nearest neighbours of the substance to be interpolated, and this is frequently not the case, *e.g.* as with nitrocubane densities.

In the *cluster expansion* method the property $Z(\xi)$ of the substance ξ is developed in terms of characteristics $z(\zeta)$ of all earlier ζ . This expansion makes use of the number $n(\zeta, \xi)$ of ways in which configurational arrangements $C' \in \zeta$ occur as substructures in a configuration $C \in \xi$. Then

$$Z(\xi) = \sum_{\zeta}^{\geq \xi} n(\zeta, \xi) \cdot z(\zeta) \quad (1)$$

The parameters $z(\zeta)$ are obtained by a fitting procedure, *e.g.* least-squares, where the cluster expansion may be conveniently truncated to a limited sequence of non-zero cluster terms $z(\zeta)$, and so applied when the earlier terms alone offer a good approximation for the property Z . For the cubane system we have found that a reasonable truncation [21] is the one based upon configurations c_0, c_1, c_{12}, c_{13} and c_{14} , that is the first three levels of the Hasse diagram (of Fig. 2) in the downward direction. For example, the configuration c_{1234} utilises just the $n(\zeta, \xi)$ -terms: $n(c_0, c_{1234}) = 1, n(c_1, c_{1234}) = 4, n(c_{12}, c_{1234}) = 3, n(c_{13}, c_{1234}) = 2,$ and $n(c_{14}, c_{1234}) = 1$. Hence, a $5 \times k$ matrix $\mathbf{n} = (n(\zeta, \xi))_{5 \times k}$ can be built where $\zeta \in \{c_0, c_1, c_{12}, c_{13}, c_{14}\}$ and ξ is one of the k species for which the property $Z(\xi)$ is known. If the known values of the property are gathered in a vector k , then through the use of the least-squares fitting procedure the parameters $z(\zeta)$, collected in the vector \vec{z} , may be computed as:

$$\vec{z} = (\mathbf{nn}^T)^{-1} \mathbf{n}k \quad (2)$$

Then, these $z(\zeta)$ are used in Eq. (1) to estimate the unknown values of the property. In contrast to the average-poset method, the cluster-expansion method has the advantage of not requiring property values for the nearest neighbours of the substance for which one wants to predict the property. A requirement for obtaining \vec{z} as intimated is that \mathbf{nn}^T be non-singular. Although often it is non-singular, the singular case sometimes occurs (*e.g.* in the present cubane case if the known set k were just a certain 5-member subset of the actual set of known values). The singular case can also be neatly solved using the generalised inverse [22] of \mathbf{nn}^T in place of the inverse.

In the *splinoid* method [20], for each connection $\zeta \succ \xi$ in the Hasse diagram, a real variable $x_{\zeta \succ \xi}$ ranging from 0 to 1 is attached ($x_{\zeta \succ \xi} = 0$ at ζ and $= 1$ at ξ). Following the usual spline interpolation method [23] a low-degree polynomial (normally a cubic one) is defined for each $x_{\zeta \succ \xi}$. Each ζ of the Hasse diagram is identified by a value α_ζ and a slope β_ζ . The splinoid fit is such that each polynomial with endpoint ζ yields the common value $= \alpha_\zeta$ at ζ , thence giving continuity, and further it is required that if $\zeta \in K$, then α_ζ is the known value $Z(\xi)$. Here K is the set of chemicals for which Z is known. Additionally, the slopes at ζ all match to β_ζ , which guarantees that the polynomials of connected arrows in the Hasse diagram fit together smoothly. These splines are further smoothed, for their “curvature” or “stress” is minimised, whence all the coefficients in the polynomials are determined, in terms of the set K of known values, if there is a sufficient number of these. Then granted this determination, the unknown values collected in U may be expressed [20] solely in terms of the values of the elements in the set K . The resultant formula entails an adjacency matrix \mathbf{A} for the Hasse diagram, as well as its oriented adjacency matrix \mathbf{S} whose elements $S_{\zeta \geq \xi}$ are:

$$S_{\zeta \geq \xi} = \begin{cases} 1, & \text{if } \xi \geq \zeta \\ -1, & \text{if } \zeta \geq \xi \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

The in-degree and out-degree of ζ are denoted by $d_{\geq \zeta}$ and $d_{\zeta \geq}$, respectively. Further matrices \mathbf{D} and $\mathbf{\Delta}$ are defined:

$$\mathbf{D} = \text{diag} [d_{\zeta \geq} - d_{\geq \zeta}] \quad (4)$$

$$\mathbf{\Delta} = \text{diag} [d_{\geq \zeta} + d_{\zeta \geq}] \quad (5)$$

Yet further if P is the set of substances in the Hasse diagram, the matrix \mathbf{U} is defined as the $|U| \times |P|$ submatrix of the unity matrix \mathbf{I} with rows indexed by the elements of U . Likewise, the matrix \mathbf{K} is defined as the $|K| \times |P|$ submatrix of \mathbf{I} with rows indexed by the elements of K . Based upon these matrices a new matrix \mathbf{M} is derived:

$$\mathbf{M} = 2(\mathbf{\Delta} - \mathbf{A}) - 3(\mathbf{D} - \mathbf{S})(\mathbf{A} + 2\mathbf{\Delta})^{-1}(\mathbf{D} + \mathbf{S}) \quad (6)$$

Finally, the vector \vec{u} containing the unknown values of the property Z is computed from the vector \vec{k} of known values as follows:

$$\vec{u} = -(\mathbf{UMU}^T)^{-1}(\mathbf{UMK}^T)\vec{k} \quad (7)$$

From Eq. (7) it can be seen that the calculation of the unknown properties is performed by using all the topological structure of the reaction network; unlike the average and cluster expansion methods that consider adjacent members of the Hasse diagram and early or late ones, respectively (these two methods may be considered local network approaches). In fact in the splinoid method the distribution of the known values is not bound to a particular trend and technically the minimum number of known values to apply the method is one, regardless of the number of substances in the Hasse diagram. Certainly, the performance of the method deteriorates, as any interpolative method, if the number of known values is low. However, we have found that such a deterioration keeps small also for large sets of unknown values [20]. As well as in the cluster expansion method (Eq. (2)), the calculation of the unknown values depends on a matrix that must be invertible; in this case \mathbf{UMU}^T . Although in all cases where the splinoid method has been applied \mathbf{UMU}^T has been invertible, for cases where it should result singular, $(\mathbf{UMU}^T)^{-1}$ could be computed as a generalised inverse.

In this work we applied the cluster expansion and the splinoid methods to the prediction of the densities of nitrocubanes, the average method could not be applied due to the few known experimental densities. The *stability of the models* was assessed through the leave-one-out (LOO) cross-validation technique [24]. In order to compare the results with a non-posetetic method, we contrasted our results with those obtained with the ACD/Labs software [25].

4 Results and discussion

When the calculations for this research started we knew only six experimental values for nitrocubanes (including cubane), they are listed in Table 1; later we found two additional values (Table 1). In the following we show the results based upon the initial

Table 1 Experimental and predicted densities of nitrocubanes

	Densities [g cm ⁻³]					ACD
	Exp.	CE(6)	CE(8)	S(6)	S(8)	
c ₀	1.29	1.290	1.276	1.29	1.29	1.577
c ₁	1.474*	1.519	1.486	1.500	1.474	1.77
c ₁₂	NA	1.747	1.678	1.645	1.635	2.01
c ₁₃	NA	1.683	1.647	1.644	1.633	2.01
c ₁₄	1.66	1.660	1.660	1.66	1.66	2.01
c ₁₂₃	NA	1.910	1.820	1.750	1.749	2.20
c ₁₂₄	NA	1.823	1.802	1.756	1.757	2.20
c ₁₃₅	1.738*	1.783	1.757	1.740	1.738	2.20
c ₁₂₃₅	NA	2.008	1.912	1.848	1.849	2.36
c ₁₂₃₈	NA	2.072	1.943	1.852	1.853	2.36
c ₁₂₃₇	NA	1.921	1.893	1.850	1.851	2.36
c ₁₂₃₄	NA	1.985	1.925	1.853	1.855	2.36
c ₁₃₅₇	1.814	1.817	1.817	1.814	1.814	2.36
c ₁₃₄₆	NA	1.898	1.907	1.855	1.857	2.36
c ₍₁₂₃₎	NA	2.082	1.998	1.948	1.949	2.50
c ₍₁₂₄₎	NA	1.995	1.980	1.949	1.949	2.50
c ₍₁₃₅₎	1.96	1.954	1.935	1.96	1.96	2.50
c ₍₁₂₎	NA	2.090	2.035	2.014	2.014	2.61
c ₍₁₃₎	NA	2.026	2.003	2.020	2.020	2.61
c ₍₁₄₎	NA	2.003	2.017	2.013	2.013	2.61
c ₍₁₎	2.028	2.034	2.022	2.028	2.028	2.71
c ₍₀₎	1.979	1.976	1.990	1.979	1.979	2.80

CE(*n*) and S(*n*) stand for cluster expansion and splinoid methods using *n* known values, respectively. Values marked with * were not considered in CE(6) and S(6) calculations. ACD stands for densities predicted with the ACD\Labs software and NA for not available datum

six values and then we compare them with those obtained by using the eight known densities.

Cluster expansion: The truncation used for this expansion considered the configurations c₀, c₁, c₁₂, c₁₃ and c₁₄. The matrix **n** is the following:

$$\mathbf{n} = \begin{matrix} & & c_0 & c_{14} & c_{1357} & c_{(135)} & c_{(1)} & c_{(0)} \\ \begin{matrix} c_0 \\ c_1 \\ c_{12} \\ c_{13} \\ c_{14} \end{matrix} & \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 4 & 5 & 7 & 8 \\ 0 & 0 & 0 & 3 & 9 & 12 \\ 0 & 0 & 6 & 6 & 9 & 12 \\ 0 & 1 & 0 & 1 & 3 & 4 \end{bmatrix} & \end{matrix} \tag{8}$$

The fitting parameters obtained by linear-squares are: $z(c_0) = 1.29$, $z(c_1) = 0.2292$, $z(c_{12}) = -0.0012$, $z(c_{13}) = -0.065$ and $z(c_{14}) = -0.0885$. The predicted

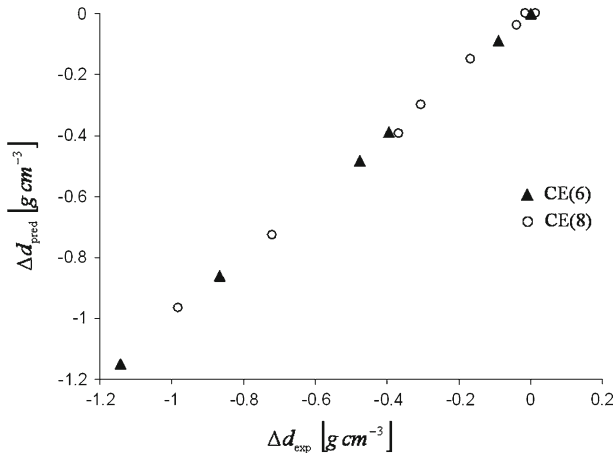


Fig. 3 Experimental versus predicted cluster expansion densities (with *six* CE(6) and *eight* CE(8) experimental values) of nitrocubanes with cubane (c_0) and mononitrocubane (c_1) contributions subtracted out

densities, CE(6), are shown in Table 1. The fitting parameters show that c_0 and c_1 dominate in estimating the densities, with the largest contribution to the model given by the cubane term (c_0). Second in importance is the monosubstituted term (c_1). The disubstituted terms c_{12} , c_{13} , c_{14} are much smaller, but still needed for a good fitting. To assess the contribution of these latter configurations, we subtracted out the contributions of c_0 and c_1 as follows:

$$\Delta d_{\text{exp}}(\xi) = d_{\text{exp}}(\xi) - z(c_0) - n(c_1, \xi) \cdot z(c_1) \quad (9)$$

$$\Delta d_{\text{pred}}(\xi) = d_{\text{pred}}(\xi) - z(c_0) - n(c_1, \xi) \cdot z(c_1) \quad (10)$$

The plot of $\Delta d_{\text{exp}}(\xi)$ versus $\Delta d_{\text{pred}}(\xi)$ is shown in Fig. 3, where can be seen that the configurations c_{12} , c_{13} and c_{14} are still important for keeping a low value of s ($s = 0.004$) and a high correlation ($r = 1$). There is a linear dependence amongst the $n(\zeta, \xi)$ for $\zeta \in \{c_0, c_1, c_{12}, c_{13}, c_{14}\}$ when the ξ are restricted to certain sets S which may be larger in size than $\{c_0, c_1, c_{12}, c_{13}, c_{14}\}$. Indeed, this occurs for the sets S which arise in the LOO procedure, namely when $\xi = c_0$ and c_{14} are left out from $\{c_0, c_{14}, c_{1357}, c_{(135)}, c_{(1)}, c_{(0)}\}$. In these cases the fits should be (and were) done using the generalised inverse [22]. The LOO statistics obtained were: $r = 0.941$ and $s = 0.381$.

The calculations performed with the eight experimental densities, CE(8), yielded the results shown in Table 1. The plot of $\Delta d_{\text{exp}}(\xi)$ versus $\Delta d_{\text{pred}}(\xi)$ is shown in Fig. 3 where, as well as in CE(6), the configurations c_{12} , c_{13} and c_{14} are important for keeping a low value of s ($s = 0.014$) and a high correlation ($r = 0.999$). In the LOO procedure it was found a linear dependence amongst the $n(\zeta, \xi)$ when c_{14} was left out from $\{c_0, c_1, c_{14}, c_{1357}, c_{(135)}, c_{(1)}, c_{(0)}\}$ for which case the fitting was done using the generalised inverse. The LOO statistics for the CE(8) methodology yielded the following statistics: $r = 0.992$ and $s = 0.034$.

Table 2 Square e^2 and average square $\langle se^2 \rangle$ of cluster expansion results for c_1 and c_{135} based upon six known experimental densities

	CE(6)	S(6)	ACD	
c_1	2.025×10^{-3}	6.76×10^{-4}	8.762×10^{-2}	e^2
c_{135}	2.025×10^{-3}	4×10^{-6}	2.134×10^{-1}	
	2.025×10^{-3}	3.4×10^{-4}	1.505×10^{-1}	$\langle se^2 \rangle$

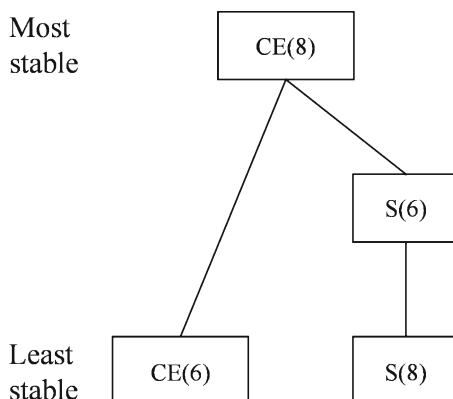
Splinoïd: The matrices **A**, **S**, **D** and **Δ** are shown in the supporting material. The predicted densities obtained based upon the six experimental values, S(6), are shown in Table 1. The LOO cross-validation statistics obtained were: $r = 0.936$ and $s = 0.097$. The calculations performed with the eight experimental values (Table 1) yielded the following LOO statistics: $r = 0.883$ and $s = 0.115$.

To compare the results of our predictions with another method, we looked for the densities calculated with the ACD/Labs software [25] (the predictive method by default in the data base SciFinder of the American Chemical Society). Unlike the posetic methodologies here described, ACD/Labs software calculates the density of a given single molecule based upon its molecular structure. The ACD densities are shown in Table 1.

To analyse the performance of these methods for predicting the experimental densities of c_1 and c_{135} , we calculated the square (of the error), $e^2 = (d_{\text{exp}} - d_{\text{pred}})^2$, of such a prediction (Table 2) and also the average square $\langle se^2 \rangle = \sum (d_{\text{exp}} - d_{\text{pred}})^2$.

According to the information on Table 2 the best method for predicting the experimental densities of c_1 and c_{135} is S(6), followed by CE(6). The least convenient method of the three here considered is ACD. From Table 2 it can be concluded that splinoïd results, S(6), are ten times better than cluster expansion ones, CE(6), and thousand times better than ACD ones, regarding c_1 and c_{135} predictions.

Finally, the relative stability of the cluster expansion and splinoïd results was evaluated by comparing their LOO statistics. By a pair-wise comparison of their r and s values [26] we obtained the poset depicted in Fig. 4. Hence, the cluster expansion method based on eight known values is the most stable one and either the cluster expansion based upon six known densities or the splinoïd based on eight values is the least stable method of the posetic ones considered here.

Fig. 4 Ranking of posetic methodologies based upon their statistics r and s obtained in the density prediction of nitrocubanes

In general, it can be said that the results here obtained show the advantage and stability of the interpolative posetic methods for the prediction of nitrocubane densities. Posetic results are better than those obtained by the ACD/Labs software, this latter a method that does not consider the reaction network but isolated molecules for its calculation.

In this study we analysed the case of the nitrocubanes based upon the substitution reaction network of the cubane, which can be extended to the study of other cubanelike systems governed by the substitution pattern considered in the poset. Thence, one can study, for instance, amino, methyl and chlorosubstituted cubanes with the procedure here developed. For the cluster expansion method of any other substitution cubanelike system, the same Eq. (2) can be applied, with \vec{k} being the vector of known properties: the term to change depending on the case to study. Likewise, when using the splinoid method, Eq. (7) holds, being the unique difference \vec{k} and the matrices \mathbf{U} and \mathbf{K} , which are simply submatrices of the unity matrix. This kind of generality is difficult to attain by the conventional QSAR/QSPR methodologies which normally require the recalculation of the molecular descriptors once the set of study is changed.

Cubanelike structures abound in chemistry and they are not only restricted to organic chemistry, there are different examples in inorganic chemistry and the generality of the methods here described makes possible to foresee future applications of these strategies to those cases.

Acknowledgments Acknowledgement of support (*via* grant BD-0894) is made to the Welch Foundation of Houston, Texas and to the Universidad de Pamplona, Colombia. G. Restrepo specially thanks P.E. Eaton for suggesting key manuscripts for having the experimental densities used in this work.

References

1. P.E. Eaton, *Angew. Chem. Int. Ed. Engl.* **31**, 1421 (1992)
2. P.E. Eaton, T.W. Cole Jr., *J. Am. Chem. Soc.* **86**, 3157 (1964)
3. *IUPAC: A Guide to IUPAC Nomenclature of Organic Compounds (Recommendations 1993)*; Blackwell scientific publications: 1993. The labels 4 and 8 were interchanged as corrected In: H.A. Favre, K.-H. Hellwich, G.P. Moss, W.H. Powell, J.G. Traynham, *Pure Appl. Chem.* **71**, 1327 (1999)
4. P.E. Eaton, R.L. Gilardi, M.-X. Zhang, *Adv. Mater.* **12**, 1143 (2000)
5. K.A. Lukin, J. Li, P.E. Eaton, N. Kanomate, J. Hain, E. Punzalan, R. Gilardi, *J. Am. Chem. Soc.* **119**, 9591 (1997)
6. P.E. Eaton, M.-X. Zhang, R. Gilardi, N. Gelber, S. Iyer, R. Surapaneni, *Propell. Explos. Pyrot* **27**, 1 (2002)
7. M.-H. Liu, C. Chen, Y.-S. Hong, C.-W. Liu, *Theor. Chem. Acc.* **113**, 35 (2005)
8. C. Zhang, *J. Phys. Chem. A* **110**, 14029 (2006)
9. M.-X. Zhang, P.E. Eaton, R. Gilardi, *Angew. Chem. Int. Ed.* **39**, 401 (2000)
10. D.J. Klein, L. Bytautas, *MATCH Commun. Math. Comput. Chem.* **42**, 261 (2000)
11. T. Ivanciuc, D.J. Klein, *J. Chem. Inf. Comput. Sci.* **44**, 610 (2004)
12. W.J. Trotter, *Combinatorics and Partially Ordered Sets Dimension Theory* (The Johns Hopkins university press, Baltimore, 1992)
13. T. Ivanciuc, O. Ivanciuc, D.J. Klein, *J. Chem. Inf. Model.* **45**, 870 (2005)
14. T. Ivanciuc, O. Ivanciuc, D.J. Klein, *Mol. Divers.* **10**, 133 (2006)
15. T. Ivanciuc, O. Ivanciuc, D.J. Klein, *Int. J. Mol. Sci.* **7**, 358 (2006)
16. D.J. Klein, *Int. J. Quantum Chem., Quantum Chem. Symp.* **20**, 153 (1986)
17. T.G. Schmalz, T. Živković, D.J. Klein, *MATH/CHEM/COMP 1987, Proc. Int. Course Conf. Interfaces Math., Chem., Comput. Sci.* 1987, **54**, 173 (1987)
18. D.J. Klein, T.G. Schmalz, L. Bytautas, *SAR QSAR Environ. Res.* **10**, 131 (1999)

19. T. Ivanciuc, D.J. Klein, O. Ivanciuc, J. Math. Chem. **41**, 355 (2007)
20. T. Došlić, D.J. Klein, J. Comput. Appl. Math. **177**, 175 (2002)
21. G. Restrepo, B. Vijayakumar, D.J. Klein, Paper in preparation
22. A. Ben-Israel, T.N.E. Greville, *Generalized Inverses, Theory and Applications* (Wiley, New York, 1974)
23. D. Kincaid, W. Cheney, *Numerical Analysis* (Brooks/Cole, Pacific Grove, 1990)
24. R. Todeschini, V. Consonni, *Handbook of Molecular Descriptors* (Weinheim, Wiley-VCH, 2000)
25. G.O. Spessard, J. Chem. Inf. Comput. Sci. **38**, 1250 (1998)
26. G. Restrepo, M. Weckert, R. Brüggemann, S. Gerstmann, H. Frank, Environ. Sci. Technol. **42**, 2925 (2008)