

Topological Organic Chemistry. 8. Graph Theory and Topological Indices of Heteronuclear Systems

Harry P. Schultz^{*,†}

Department of Chemistry, University of Miami, Coral Gables, Florida 33124

Emily B. Schultz[‡] and Tor P. Schultz[§]

Department of Forestry and Forest Products Laboratory, Mississippi State University,
Mississippi State, Mississippi 39762

Received April 7, 1994*

Distance matrices of representative heteronuclear molecules acquired from edge, vertex, and valence weighted graphs were used to generate invariant, whole-number topological indices: the MTI' index, determinant, permanent, product of row sums, and the long hafnian. Within a related group of compounds these descriptors were of monotonic series and displayed reasonable quantitative-structural comparisons.

INTRODUCTION

This is the target paper of a series for which the first was meant to be. The current article extends to heteroatom-containing organic substances the studies of the utilities for the generation of topological indices of structure-descriptive homonuclear distance (**D**) matrices reported in earlier papers¹ of this sequence. Sylvester² was the first to investigate graphic representations that reflected early, primitive knowledge of molecular structures. The study of topological indices, mostly of homonuclear (carbon atoms) graphs, has greatly expanded in the past 4 decades, as effectively reviewed by Mihalić et al.³

Barysz et al.⁴ were early investigators of heteronuclear graphs and cited publications on that topic, a course more recently also taken by Katritsky and Gordeeva.⁵ Researchers probing heteronuclear graphs, their matrices, and topological indices redefined in sundry ways the relationships of the edges and the identities of the vertices of molecular graphs. Construction with variously weighted graph values of graph-theoretical **D** matrices of heteronuclear organic compounds was thereby attained.

COMPUTATIONS

Molecular (multi)graphs, with hydrogens suppressed and with the interatomic carbon-carbon edge counts set at unity for tetrahedral, two for trigonal, and three for digonal hybridized carbon atoms, were derived from the representative compounds listed in Tables 1-5. Both the canonical CIP and IUPAC systems of numbering were used in all calculations, although all of the investigated matrix functions were invariant-topological descriptors. Particulars about specific molecules used the IUPAC system of nomenclature. Reference is made to prior papers of this series for each mode of computation for the various functions and indices described in succeeding portions of this paper.

After much study, three variables were utilized for insertion into the **D** matrix that represented the graph of a heteronuclear organic molecule: the scaled or weighted edge (*e*) value which represented the dimension of the carbon-heteroatom or heteroatom-heteroatom bond of a graph, the weighted vertex (*V*) integer which depicted the different elements, and lastly, each vertex valence (degree, *v*), a number that indicated the summation of bonds and/or unshared electron pairs associated with each vertex.

Extensive experimentation dictated that the **D** matrix elements, and therefore also resultant matrix functions, would in this paper be expressed as integers that would reflect and contrast the differences between carbon atoms and heteroatoms. This made for more simple computations and greater ease of comparisons. The adjacency (**A**) matrix was not used in this work because the information conveyed by it was the same for a heteronuclear graph as for its congeneric homonuclear graph. A prior paper¹⁸ has observed the dubiety of **A** matrix usage within the purview of these studies.

Much thought was applied to the calculation of edge weighted (*e*) bond lengths. Initially it was hoped that integer values for bond lengths (nanometers) or bond energies (kilocalories or kilojoules per mole) either directly or as ratios (hence usually not integers) could be utilized, for the **D** matrix functional values derived from such employment would have a great likelihood of uniqueness. However, even with simple, low molecular weight molecules, the resultant matrix functions were already very large. The atomic numbers of atoms joined by an edge were used to calculate the value of an *e* weighted bond (EWB). For two attached heteroatoms, the EWB value was determined by use of the algorithm

$$\text{EWB} = \text{EC} + (\text{sum atomic no. of heteroatoms}) - (2 \times \text{atomic no. of carbon})$$

where the edge count (EC) is 1 for a single, 2 for a double, and 3 for a triple bond. The value of the single nitrogen-oxygen *e* weighted bond in hydroxylamine derivatives, for example, was evaluated from the expression

$$\text{EWB}(\text{N-O}) = 1 + (7 + 8) - (2 \times 6) = 4$$

Usually the heteroatom was attached directly to a carbon atom and the above algorithm reduced to

$$\text{EWB} = \text{EC} + (\text{atomic no. of heteroatom} - \text{atomic no. of carbon})$$

The magnitude of the double carbon-nitrogen *e* weighted

* Present address: P.O. Box 262, Big Horn, WY 82833.

† Department of Forestry.

‡ Forest Products Laboratory.

• Abstract published in *Advance ACS Abstracts*, July 1, 1994.

Table 1. MTI' Indices for Selected Heteronuclear C₃ Graphs Calculated from Various Weighted (Edge, *e*; Vertex, *V*; and/or Valence, *v*) Distance Matrices

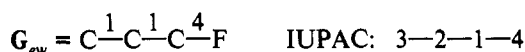
no.	compound	MTI' index $D_N^N(G_w)$							
		G	G_{ew}	G_{Vw}	G_{Vw}	G_{eVw}	G_{eVw}	G_{Vw}	G_{eVw}
1	1-propanamine	34	45	46	62	63	83	86	119
2	N-methylethanamine	32	55	44	64	76	111	100	174
3	trimethylamine	27	54	39	63	78	126	111	222
4	2-propanamine	29	40	39	57	56	80	77	112
5	1-propylidenimine	59	74	86	153	110	194	234	302
6	2-propylidenimine	52	67	76	148	100	193	220	292
7	propanenitrile	92	111	140	320	171	389	512	629
8	1-propanol	40	68	67	92	140	164	200	380
9	methoxyethane	36	88	68	92	168	228	220	548
10	2-propanol	34	62	64	82	128	158	172	356
11	propanal	68	104	140	216	224	340	504	820
12	propanone	60	96	124	204	208	336	460	784
13	propanoic acid	123	221	273	407	511	741	953	1787
14	methyl ethanoate	119	271	271	419	615	955	1027	2331
15	ethyl methanoate	135	319	303	451	711	1067	1123	2635
16	1-fluoropropane	46	97	118	134	277	305	422	1025
17	2-fluoropropane	39	90	99	117	258	294	357	966
18	1,1-difluoropropane	100	256	292	334	808	898	1102	3106
19	2,2-difluoropropane	86	242	254	302	770	890	974	3002
20	1,1,1-trifluoropropane	172	487	532	620	1603	1817	2060	6281
21	propanamide	114	183	228	362	381	593	782	1325
22	propanoyl fluoride	132	267	336	470	727	975	1286	2815
23	N-propylhydroxylamine	87	175	168	215	364	449	458	1016
24	O-propylhydroxylamine	84	191	160	214	375	493	478	1117
25	1-nitropropane	211	460	487	719	1108	1616	1663	4208
26	propyl nitrite	228	579	516	776	1335	2015	1928	5039
	av	86	181	191	273	437	597	673	1592

bond of oximes was calculated as follows:

$$\text{EWB}(\text{C}=\text{N}) = 2 + (7 - 6) = 3$$

In the above expressions all *e* weighted bonds simplify to 1 for a single, 2 for a double, and 3 for a triple bond for **D** matrix construction involving only homonuclear (carbon) bonds. Elements that lie to the left of carbon in the second quantum level of the periodic chart yield zero or negative *e* weighted bond values (Li, -2; Be, -1; B, 0) for single bonded carbon-heteroatom edges; the zero or negative values were incorporated into **D** matrices to yield meaningful matrix functions.

The *e* weighted graph of 1-fluoropropane is



Its *e* weighted **D** matrix is

$$D_4^4(G_{ew}) = \begin{bmatrix} 0 & 1 & 2 & 4 \\ 1 & 0 & 1 & 5 \\ 2 & 1 & 0 & 6 \\ 4 & 5 & 6 & 0 \end{bmatrix}$$

The *e* weighted graph of propyl lithium is

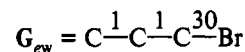


Its *e* weighted **D** matrix is

$$D_4^4(G_{ew}) = \begin{bmatrix} 0 & 1 & 2 & -2 \\ 1 & 0 & 1 & -1 \\ 2 & 1 & 0 & 0 \\ -2 & -1 & 0 & 0 \end{bmatrix}$$

As in the nonweighted **D** matrices, symmetry about the

principal diagonal was preserved in the *e* weighted **D** matrices. The above technique for representing an *e* weighted graph was applicable to elements that exist in higher quantum levels, as is frequently encountered in the halogen family. The *e* weighted graph of 1-bromopropane is

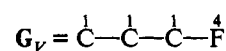


Vertex weights (*V*) were based upon the same data as were edge weights—the atomic numbers of the vertices. The algorithm is

$$V = 1 + (\text{atomic no. of heteroatom} - \text{atomic no. of carbon})$$

Thus, the vertex weights for the seven reactive elements of the second quantum level of the periodic chart are as follows: Li, -2; Be, -1; B, 0; C, 1; N, 2; O, 3; F, 4. The heavier the heteroatom, the greater is the weighted *V* value. In the instance of the homonuclear carbon graphs, the above expression for *V* reduced to unity.

The *V* weight data were recorded and used in a weighting, diagonal matrix which utilized only its principal diagonal to record the *V* data for weighting purposes, as seen in and derived from the *V* weighted graph for 1-fluoropropane



and its weighting matrix, **M**

$$M_4^4(G_V) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

It is interesting to note the similarity of the *V* weighted diagonal matrix to the corresponding unit or identity matrix. Pre-

Table 2. Descriptors for Selected Heteronuclear C₃ Graphs Calculated from the Corresponding Edge (*e*), Vertex (*V*), and Valence (*v*) Weighted Distance Matrices

no.	compound	MTI'	det	index $D_N^V(G_{eVw})$		
				per	PRS	lhaf
1	1-propanamine	119	-512	2 304	25 200	896
2	N-methylethanamine	174	-960	6 912	58 212	2 976
3	trimethylamine	222	-1 536	4 608	48 000	1 536
4	2-propanamine	112	-384	1 200	4 500	408
5	1-propylidenimine	302	-2 160	9 216	124 416	3 528
6	2-propylidenimine	292	-1 440	4 704	64 680	1 632
7	propanenitrile	629	-6 144	25 600	423 360	9 728
8	1-propanol	380	-2 160	9 216	124 416	3 528
9	methoxyethane	548	-6 048	55 296	460 800	24 824
10	2-propanol	356	-1 620	5 292	72 765	1 836
11	propanal	820	-6 912	28 800	476 280	10 944
12	propanone	784	-4 608	15 552	258 048	5 472
13	propanoic acid	1 787	746 496	14 542 848	556 174 080	1 147 008
14	methyl ethanoate	2 331	1 824 768	53 637 120	1 575 415 296	2 666 112
15	ethyl methanoate	2 635	2 737 152	162 570 240	4 324 147 200	5 574 384
16	1-fluoropropane	1 025	-6 144	25 600	423 360	9 728
17	2-fluoropropane	966	-4 608	15 552	258 048	5 472
18	1,1-difluoropropane	3 106	1 966 080	39 714 816	1 608 837 120	3 392 000
19	2,2-difluoropropane	3 002	1 310 720	20 316 160	837 591 040	2 169 600
20	1,1,1-trifluoropropane	6 281	-469 760 000	54 869 884 928	5 227 505 418 240	3 787 456 512
21	propanamide	1 325	196 608	3 735 552	141 523 200	301 312
22	propanoyl fluoride	2 815	1 966 080	39 714 816	1 608 837 120	3 158 272
23	N-propylhydroxylamine	1 016	110 592	3 483 648	138 996 000	373 152
24	O-propylhydroxylamine	1 117	165 888	6 881 280	266 133 504	748 352
25	1-nitropropane	4 208	-38 338 560	9 498 599 424	925 893 918 720	829 992 960
26	propyl nitrite	5 039	-61 931 520	46 430 078 976	4 216 069 440 000	4 901 907 456

multiplying the *V* weighting matrix and the nonweighted *D* matrix of 1-fluoropropane gives the *V* weighted *D* matrix of 1-fluoropropane.⁶

$$M_4^V(G_V) \times D_4^V(G) = D_4^V(G_{Vw})$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 3 \\ 1 & 2 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 3 \\ 4 & 8 & 12 & 0 \end{bmatrix}$$

And similarly multiplying the *V* diagonal matrix with the *e* weighted *D* matrix of 1-fluoropropane yields the *eV* weighted matrix of 1-fluoropropane.

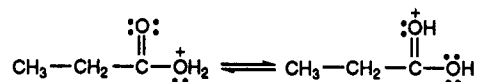
$$M_4^V \times D_4^e(G_{ew}) = D_4^e(G_{eVw})$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 4 \\ 1 & 0 & 1 & 5 \\ 2 & 1 & 0 & 6 \\ 4 & 5 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 2 & 4 \\ 1 & 0 & 1 & 5 \\ 2 & 1 & 0 & 6 \\ 16 & 20 & 24 & 0 \end{bmatrix}$$

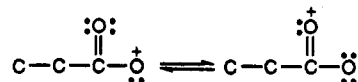
The above matrix multiplications destroyed the original symmetries of nonweighted and *e* weighted *D* matrices. These techniques also extended to higher quantum level elements that might be attached to carbon atoms, such as chlorine and bromine.

The third and last structural variant injected into the *D* matrices of heteronuclear graphs was the valence (degree, *v*) of each vertex. The edge count about each hydrogen-suppressed vertex supplied a portion of the valence information; pairs of unshared electrons, present or absent, furnished the remainder. The unshared electron pairs (••) associated with the heteroatoms that lie to the right of the element carbon in the second and higher quantum levels of the periodic chart were regarded in the same light as were the extra π electrons of alkene carbon atoms surveyed in an earlier paper¹⁸—each electron pair counted as one valence. However, unlike the

carbon π alkene electrons, the unshared heteroatom electron pairs were not counted as edges. This procedure permitted the calculation of topological indices for various charged heteroatom-containing (aminium and oxonium) species, as for example the equilibrium expressions for the protonated propanoic acid (Table 3, column 2; 8/10).

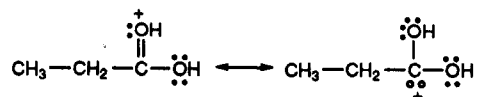


which when hydrogen-suppressed became the following graphs

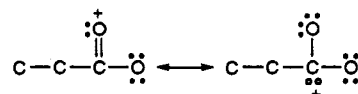


that then yielded weighted *D* matrices that by any mode of computation gave different topological indices.

For those heteroatoms lying to the left of carbon in the periodic chart, each missing electron pair (••) contributed -1 to the total valence of the heteroatom. Thus, all the heteroatoms in the second quantum level connected by a single bond to one carbon atom had the *v* numbers: Li, -2; Be, -1; B, 0; N, 2; O, 3; F, 4. This concept was also extended to various carbonium ion species, for example the canonical structures of the contributing resonance hybrids of the carbonyl oxygen protonated propanoic acid (Table 3, column 2; 8/9),

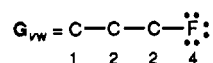


which when hydrogen-suppressed are represented by the following graphs



and yielded different topological indices from the corresponding weighted **D** matrices. Free radicals, with only one electron in the outer valence shell, contributed half a bond to the total valence of the atom (Table 3, column 2; 6).

As with vertex (*V*) weights, the valency (*v*) data were summarized in a square, diagonal matrix, then injected by premultiplication into the appropriate **D** matrix to form a *v* weighted **D** matrix. This operation is illustrated with the graph of 1-fluoropropane



whose diagonal matrix is

$$M_4^4(G_v) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

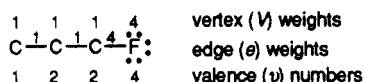
and whose *v* weighted **D** matrix is

$$D_4^4(G_{vw}) = \begin{bmatrix} 0 & 2 & 4 & 2 \\ 2 & 0 & 2 & 4 \\ 2 & 1 & 0 & 3 \\ 4 & 8 & 12 & 0 \end{bmatrix}$$

and whose edge (*e*), vertex (*V*), and valence (*v*) weighted **D** matrix is

$$D_4^4(G_{eVvw}) = \begin{bmatrix} 0 & 2 & 4 & 8 \\ 2 & 0 & 2 & 10 \\ 2 & 1 & 0 & 6 \\ 64 & 80 & 96 & 0 \end{bmatrix}$$

which reflects all the *e*, *V*, and *v* data, as defined in this work, of 1-fluoropropane.



The triple weighted **D** matrix provided a wealth of structural information, perhaps eventually even making possible the attainment of that chimera of all topological indices—total uniqueness. However, the triple weighted matrix was bulky and awkward to construct and use and gave indices of relatively great magnitude. The above operations can be extended to elements in higher quantum levels of the periodic chart, recognizing if needed the expansion of the outer valence shells of some larger elements, although in this paper the maximum number of electrons in any outer shell was restricted to eight.

To this point, means have been presented for constructing **D** matrices of heteronuclear graphs of no weight up to triple weight, with six options lying between the two extremes: $D_N^N(G)$, $D_N^N(G_{ew})$, $D_N^N(G_{Vw})$, $D_N^N(G_{vw})$, $D_N^N(G_{eVw})$, $D_N^N(G_{eVvw})$, and $D_N^N(G_{eVvw})$; only the first two of the listed matrices were symmetrical.

Topological Indices Derived from the Graph-Theoretical Weighted Distance Matrices. Five topological indices are reported on in this paper: the MTI' index, determinant (|det|), permanent (per), product of the row sums (PRS), and the long hafnian (lhaf).

Trinajstić and co-workers^{7,8} described the MTI' index, which is derived from the **D** matrix and the degree vector of a graph

$$\text{MTI}' = \sum_{i=1}^N e_i'$$

where e_i' ($i = 1, 2, \dots, N$) are the elements of the row ($1 \times N$) vector and $v \cdot \mathbf{D} = (e_1', e_2', \dots, e_N')$. In short, the MTI' index was obtained by premultiplying the valence (*v*) vector and the **D** matrix and summing the resulting vector elements. In this report in effect the MTI' index weights the valences of a vertex twice, unlike the other indices portrayed in the paper, once by definition and once as a consequence of weighting the **D** matrix with the *v* diagonal matrix.

Matrix determinants as topological indices were described in the second paper^{1b} of this series. Pettoufrezzo⁶ defined the value of the determinant of a square matrix **A** = (a_{ij}) of order *N* as the sum of the *N!* terms of the form

$$(-1)^k a_{1i_1} a_{2i_2} \dots a_{Ni_N}$$

each term holding only one element from different rows and different columns. The exponent *k* represents the number of interchanges of two elements necessary to place the subscripts in the order 1, 2, ..., *N*.

Matrix permanents and products of the row sums were described in the fourth paper^{1d} of this sequence. The definition of the permanent is similar to that for the determinant, except that all permanent values are positive (if all elements are positive); hence the $(-1)^k$ factor of the determinant expression is omitted. The product of the row sums was an early crude approximation to the permanents.

The hafnians were described in the fifth article^{1e} of this series as the summation over all permutations i_1, i_2, \dots, i_{2N} which satisfy the limitations $i_1 < i_2, i_3 < i_4, \dots, i_{2N-1} < i_{2N}, i_1 < i_3 < i_5 \dots < i_{2N-1}$ for a $N \times N$ matrix.

RESULTS

Table 1 displays and compares all the possible weighted combinations reported in this paper for the MTI' indices of a collectanea of C_3 heteroatom-containing substances; the MTI' index was chosen for this initial survey because it was invariant and relatively easy to calculate. The MTI' indices of the **D** matrices of the nonweighted graphs (Table 1, column 3) and triple weighted (*eVv*) graphs (Table 1, column 10) presented the lowest and highest limits of the MTI' indices of this representative sample of C_3 heteronuclear graphs. In between, single weighted (*e*, *V*, *v*) and double weighted (*eV*, *ev*, *Vv*) matrix values (Table 1, columns 4–9) completed the data. Each one of the columns of data was averaged to quantify the relative order and assumed significance of each set of data; this operation was especially productive in displaying the relative consequences of the information presented by the single and double weighted graphs. It was presumed that the larger the MTI' index was for a given substance, the more information was gathered into the **D** matrix from which the MTI' index was derived, and it was therefore qualitatively and subjectively concluded that, of the single weighted indices, the *v* weighted MTI' values were the most important and, of the double weighted indices, the *Vv* weighted MTI' indices were the most significant.

A few compounds (Table 1: columns 4 and 5, 2, 3, 4, 8, 9; columns 8 and 9, 2, 3, 4) displayed slight inversions from the order of the average values of each column. Were extensive studies to be undertaken of only such slightly deviant types

Table 3. Descriptors Derived from the Edge (*e*), Vertex (*V*), and Valence (*v*) Weighted Distance Matrices of Selected Homonuclear and Heteronuclear Graphs^a

no.	compound	index $D_N^N(G_{eVw})$				
		MTI'	det	per	PRS	lhaf
1	<chem>CH3CH2CH2Li</chem>	35	0	256	-144	128
2	<chem>CH3CH2CH2BeH</chem>	20	16	0	64	8
3	<chem>CH3CH2CH2BH2</chem>	29	0	0	0	0
4	<chem>CH3CH2CH2CH2</chem>	200	-1 440	6 144	82 944	2 352
5	<chem>CH3CH2CH2</chem>	11	0	0	0	2
6	<chem>CH3CH2CH2</chem>	17.75	12	12	54	11
7	<chem>CH3CH2CH2</chem>	23	16	16	72	14
8	<chem>CH3CH2C(OH)OH</chem>	1 325	559 872	10 907 136	417 130 560	862 020
9	<chem>CH3CH2C(OH)2</chem>	1 061	186 624	3 530 304	122 821 920	343 692
10	<chem>CH3CH2C(OH)OH2</chem>	1 502	497 664	9 695 232	370 782 720	769 536
11	<chem>c1ccccc1F</chem>	3 057	0	16 711 121 664	3 730 987 008 000	240 866 946
12	<chem>c1ccccc1</chem>	3 116	0	22 989 837 312	4 672 870 871 040	286 190 496

^a Unshared electrons present, *; electrons absent, ° °.**Table 4.** Descriptors Derived from the Edge (*e*), Vertex (*V*), and Valence (*v*) Weighted Distance Matrices of Selected Carboxylic Acids

no.	compound, acid	index $D_N^N(G_{eVw})$				
		MTI'	det	per	PRS	lhaf
1	methanoic	861	5.4432×10^4	54 432	249 480	6 111
2	ethanoic	1 284	-1.65888×10^5	623 808	7 741 440	228 960
3	fluoroethanoic	3 722	6.3700992×10^7	1 066 106 880	35 978 342 400	35 937 984
4	difluoroethanoic	7 244	-1.631×10^{10}	1 495 125 983 232	126 750 004 346 880	94 351 785 984
5	trifluoroethanoic	11 874	3.4789×10^{12}	2 012 890 647 232 512	433 564 990 434 123 776	5 353 521 020 928
6	chloroethanoic	13 954	9.5551488×10^8	18 391 007 232	1 100 170 874 880	544 621 248
7	dichloroethanoic	35 810	-1.761×10^{12}	243 214 082 113 536	35 588 492 505 907 200	16 906 689 183 744
8	trichloroethanoic	67 434	5.5795×10^{15}	5 925 854 344 321 695 744	2 105 323 955 684 827 987 968	11 327 250 900 713 472
9	bromoethanoic	66 928	1.1347×10^{10}	244 784 332 800	51 472 392 192 000	7 041 700 800
10	iodoethanoic	161 374	4.2807×10^{10}	960 855 146 496	504 501 214 593 024	27 486 311 616
11	propanoic	1 787	7.46496×10^5	14 542 848	556 174 080	1 147 008
12	2-chloropropanoic	17 051	-3.0099×10^9	428 869 168 128	86 110 611 874 560	31 137 219 072
13	3-chloropropanoic	17 684	-4.0132×10^9	622 310 916 096	132 365 565 296 640	45 824 901 120
14	hydroxyethanoic	2 557	2.4634368×10^7	407 089 152	13 450 364 928	14 798 592
15	aminoethanoic	1 924	6.63552×10^6	111 476 736	3 777 269 760	4 793 664
16	zwitterion of 15	2 324	4.42368×10^6	74 317 824	2 518 179 840	3 479 424

of compounds, a different average priority of the germane weight values might easily be ascertained.

Four of the MTI' indices in Table 1 were duplicate values (Table 1: column 3, 1/10; column 5, 3/4; column 6, 8/9; column 9, 6/9). The remainders of the index sets were unique, including values for the rather sophisticated pairs of structural isomers (Table 1: 14/15, 23/24, and 25/26).

The C₃ compounds listed in Table 1 contained the three common heteroatoms—nitrogen, oxygen, and fluorine—of quantum level 2 of the periodic chart, each of which was successively substituted onto a functional carbon atom so as to elevate the carbon atom to its first, second, or third oxidation levels. All of the MTI' indices increased in the same order as do the atomic numbers of the attached heteroatoms (Table 1: 1, 8, 16, for one example); the indices also increased with the increasing elevation of oxidation levels of the functional carbon atom with any heteroatom (Table 1: 1, 5, 7, as an

example). For multiple different heteroatom functions (Table 1: 21, 13, 22), the MTI' indices reflected the increasing totality of the atomic numbers of all attached heteroatoms.

As the heteroatom substituent moved from a primary to a secondary carbon atom, the MTI' index value generally decreased (Table 1: 1/4, 8/10, and 16/17). However, as a heteroatom was inserted from the exterior to the interior of a carbon chain as in the secondary and tertiary amines and the ether (Table 1: 1/2/3, 8/9), the index value consequences of such shifts were variable; usually the MTI' index decreased in the instances of lesser weighted D matrices but increased in most of the more highly weighted matrices.

The experimental results condensed in Table 1 established the background for extending this study to other invariant indices, as summarized in Table 2 for the same sampling of C₃ heteronuclear compounds and in Table 3 for instances of more unusual C₃ organic species. The eight different weighted

Table 5. Descriptors Derived from the Edge (*e*), Vertex (*V*), and Valence (*v*) Weighted Distance Matrices of Representative Alcohols

no.	compound, -ol	bp/°C	index $D_N^N(G_{eVvW})$				
			MTI'	det	per	PRS	lhaf
1	methan	65	84	-81	81	81	81
2	ethan	79	210	432	432	2 520	308
3	1-propan	97	380	-2 160	9 216	124 416	3 528
4	1-butan	117.7	602	10 368	276 480	10 077 696	34 120
5	1-pentan	138	884	-48 384	13 879 872	1 251 478 800	1 162 080
6	1-hexan	158	1 234	221 184	961 482 240	223 094 995 200	18 715 840
7	1-heptan	176	1 660	-995 328	90 956 611 584	54 177 693 696 000	1 235 049 984
8	2-propan	82	356	-1 620	5 292	72 765	1 836
9	2-butan	99.5	548	7 776	154 656	5 322 240	18 390
10	2-pentan	120	802	-36 288	7 637 760	638 668 800	603 936
11	2-hexan	140	1 126	165 888	533 471 616	113 416 535 232	10 408 656
12	2-heptan	160	1 528	-746 496	51 066 167 040	27 807 142 222 800	674 967 168
13	3-pentan	116	766	-36 288	6 646 320	555 716 700	475 848
14	3-hexan	135	1 054	165 888	439 579 008	91 707 033 600	7 616 208
15	3-heptan	157	1 420	-746 496	40 926 477 312	21 679 995 617 280	499 644 288
16	4-heptan	161	1 384	-746 496	37 497 848 832	19 867 943 042 640	424 486 656
17	2-methyl-2-propan	82.4	528	5 184	63 936	2 361 960	9 120
18	2-methyl-2-butan	102	746	-24 192	2 612 736	232 243 200	169 920
19	2-methyl-2-pentan	121.4	1 038	110 592	174 735 360	37 955 174 400	3 505 664
20	2-methyl-2-hexan	143	1 412	497 664	16 746 195 456	9 089 908 581 600	195 478 272
21	3-methyl-3-pentan	122.4	990	110 592	132 408 576	29 568 240 000	2 476 800
22	3-methyl-3-hexan	143	1 316	497 664	10 957 662 720	6 143 456 124 000	107 557 632
23	2-methyl-1-propan	108	574	7 776	129 600	5 140 800	17 982
24	2-methyl-1-butan	127.8	818	-36 288	5 371 056	527 489 820	365 256
25	2-methyl-1-pentan	148	1 132	165 888	360 025 344	86 612 198 400	6 920 688
26	2-methyl-1-hexan	164	1 524	746 496	34 318 964 736	20 502 801 285 120	416 147 328

D matrices used for the MTI' calculations of Table 1 presented an embarrassment of riches for use in the subsequent studies of this effort; therefore only *eVv* weighted **D** matrices were selected for determining the topological descriptors listed in Tables 2–5.

Unfortunately, prior papers of this sequence calculated topological descriptors of the **D** matrices of homonuclear graphs that were automatically *eV* weighted (both were unity), but none were computed from the **D** matrices of *v* weighted graphs. Hence, none of the data presented in this paper were directly comparable to data in former publications of this series. Gutman⁹ has interestingly contrasted the summed and product influences of vertex valences on MTI' and MTI indices.

Table 2 presents the same compounds as those depicted in Table 1, all of whose graphs were used to formulate the corresponding *eVv* weighted **D** matrices, from which were then derived the following invariant topological indices: MTI' (repeated from Table 1), determinant (*det*), permanent (*per*), product of row sums (*PRS*), and long hafnian (*lhaf*). The *eVv* weighted **D** matrices were not symmetrical; hence the short hafnian described in an earlier report¹⁶ was a meaningless function in these experiments. No unusual problems were encountered in the course of these calculations; it appeared that the weighted **D** matrices were utilizable as sources of topological descriptors. Not all the data of Table 2 were unique, for the determinants and the permanents exhibited the same four pairs of duplicate values (Table 2: columns 4 and 5, 5/8, 7/16, 12/17, 18/22).

Interesting generalities were seen in each instance. It was noted that the topological indices, including the absolute values of the determinants, decreased in value when the heteroatom(s) were moved from an exterior to an interior carbon atom. When the oxidation level of the functional carbon atom was elevated, the index value increased, and as the substituent heteroatoms progressed from lowest (nitrogen) to highest (fluorine) atomic numbers, the index numbers at any comparable level of substitution on the carbon atom were enhanced.

Although not monotonic overall, within related short sequences (Table 2: 1, 5, 7; 8, 11, 13; 16, 18, 20, for example)

orderly progressions of the indices were observed. Experiments were summarized with the data of Table 5 that tested the potential for these techniques to yield series of indices displaying monotonicity and quantitative–structural relationships.

Table 3 continued data that recorded the testing of the efficacy of the techniques examined in this paper to yield reasonable topological indices for a variety of unusual C₃ organic molecules, ions, and radicals. No quantitative–structural relationships were sought for the types of compounds listed in Table 3 nor was there evidence of monotonic relationships among the structures.

However, it was satisfying to note that equilibrium structures (Table 3: 8/10) and contributing resonance hybrids (Table 3: 8/9, 11/12), as observed above, possessed different topological indices as well as the contrasts displayed by ions and fragments (Table 3: 5/6/7).

Table 4 was assembled to test the extension of the principles of calculation of topological indices derived from the graph-theoretical weighted **D** matrices espoused in this paper to elements that lie in quantum levels above the second level of the periodic chart. The halogen family was chosen as being a widely encountered instance of this circumstance. The halogen atoms were incorporated into organic carboxylic acids, all common and well-known materials. Additionally, the topological data for glycolic acid and glycine were included (Table 4: 14, 15). Because amino acids are aminium acids, not carboxylic acids, the acidic proton being derived from the nitrogen atom of the amino acid rather than its oxygen atom(s), both sets of topological indices were listed (Table 4: 15, 16). No problems were encountered in calculating the topological indices for the compounds listed in Table 4, but many were of great magnitude.

Although the various sample sets of substituted carboxylic acids possessed too few substances for regression studies involving quantitative–structural relationships, it appeared that there was an air of monotonicity displayed by several small series of related acids (Table 4: 1, 2, 11; 3, 6, 9, 10; 15, 14, 3; and 6, 7, 8, for several instances).

Table 5 presents a list of C₁–C₇ alcohols comprising a variety of types of these useful and common substances. Diastereoisomeric structures were excluded; of the enantiomeric pairs containing one stereocenter (Table 5: 9, 10, 11, 12, 14, 15, 22, 24, 25, 26) a couple of racemic mixtures and the corresponding optically active antipodes had boiling points that differed by no more than 1 °C—probably observer artifacts—from one another. No problems were encountered in calculating the various data, but the determinants, unlike the other indices listed, displayed six sets of replicate values (Table 5: column 5, 9/23, 10/13/24, 11/14/25, 12/15/16/26, 19/21, 20/22). The purpose of gathering the data into Table 5 was to test the ease of computation of various indices described in this paper when applied to one class of compound and to present data for a single functional group that tested the quantitative–structural relations between boiling points and topological indices.

MTI' (D) Index (Table 5, Column 4). This sequence, like all except the determinants of Table 5, was unique and possessed a numbers progression that indicated monotonically related index values. Regression analysis for the boiling points versus the MTI' index values for all the alcohols listed gave the relatively simple equation

$$\text{bp}/^{\circ}\text{C} = 64.85 + 0.0654(\text{MTI}')$$

$$n = 26 \quad r^2 = 0.918 \quad s = 8.741 \quad F = 267.1$$

The other descriptors (Table 5, columns 5–8) were similarly compared by regression analyses for the boiling points versus the respective indices. As in the preceding tables, the signs of the determinant indices were recorded, but only the absolute values were used in the regression calculations.

det|D| Index (Table 5, Column 5).

$$\text{bp}/^{\circ}\text{C} = 22.06 + 3.517(\text{det|D|})^{1/2}$$

$$n = 26 \quad r^2 = 0.906 \quad s = 9.318 \quad F = 232$$

per(D) Index (Table 5, Column 6).

$$\text{bp}/^{\circ}\text{C} = 47.25 + 10.613(\log \text{per(D)})$$

$$n = 26 \quad r^2 = 0.902 \quad s = 9.530 \quad F = 220.8$$

PRS(D) Index (Table 5, Column 7).

$$\text{bp}/^{\circ}\text{C} = 47.84 + 8.376(\log \text{PRS(D)})$$

$$n = 26 \quad r^2 = 0.835 \quad s = 12.37 \quad F = 121.4$$

lhaf(D) Index (Table 5, Column 8).

$$\text{bp}/^{\circ}\text{C} = 42.35 + 13.710(\log \text{lhaf(D)})$$

$$n = 26 \quad r^2 = 0.916 \quad s = 8.830 \quad F = 261.2$$

As seen from the above regression equations of all the data of Table 5, similar results were obtained; the five sets of topological indices were ordered in the following sequence of

decreasing reliability.

$$\text{MTI}' > \text{lhaf} > \text{det} > \text{per} > \text{PRS}$$

SUMMARY

Heteronuclear graphs were represented with weighted edges, vertices, and valences and were successfully and usefully transformed into weighted distance matrices. Invariant and integer matrix functions useful as topological indices—the MTI' index, determinant, permanent, product of row sums, and the long hafnian—were calculated from the information presented by the weighted distance matrices.

Although few of the series of topological indices reported in this article exhibited monotonicity within the totality of the representative groups of compounds, similar sets of substances possessing related heteroatoms at comparable oxidation levels displayed monotonicity, and as illustrated by the data of Table 5 when quantitative–structural comparisons were examined, reasonably acceptable results were obtained.

ACKNOWLEDGMENT

The authors thank Dr. Milan Randić, Drake University, Des Moines, IA, for constructive comments and advice.

REFERENCES AND NOTES

- (a) Schultz, H. P. *Topological Organic Chemistry*. 1. Graph Theory and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **1989**, 29, 227–228. (b) Schultz, H. P.; Schultz, E. B.; Schultz, T. P. *Topological Organic Chemistry*. 2. Graph Theory, Matrix Determinants and Eigenvalues, and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **1990**, 30, 27–29. (c) Schultz, H. P.; Schultz, T. P. *Topological Organic Chemistry*. 3. Graph Theory, Binary and Decimal Adjacency Matrices, and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **1991**, 31, 144–147. (d) Schultz, H. P.; Schultz, E. B.; Schultz, T. P. *Topological Organic Chemistry*. 4. Graph Theory, Matrix Permanents, and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **1992**, 32, 69–72. (e) Schultz, H. P.; Schultz, T. P. *Topological Organic Chemistry*. 5. Graph Theory, Matrix Hafnians and Pfaffians, and Topological Indices of Alkanes. *J. Chem. Inf. Comput. Sci.* **1992**, 32, 364–368. (f) Schultz, H. P.; Schultz, T. P. *Topological Organic Chemistry*. 6. Graph Theory and Molecular Topological Indices of Cycloalkanes. *J. Chem. Inf. Comput. Sci.* **1993**, 33, 240–244. (g) Schultz, H. P.; Schultz, E. B.; Schultz, T. P. *Topological Organic Chemistry*. 7. Graph Theory and Molecular Topological Indices of Unsaturated and Aromatic Hydrocarbons. *J. Chem. Inf. Comput. Sci.* **1993**, 33, 863–867.
- Sylvester, J. J. On an Application of the New Atomic Theory to the Graphical Representation of the Invariants and Covariants of Binary Quantics—With Three Appendices. *Am. J. Math.* **1878**, 1, 64–90.
- Mihalić, Z.; Veljan, D.; Amić, D.; Nikolić, S.; Plavšić, D.; Trinajstić, N. The Distance Matrix in Chemistry. *J. Math. Chem.* **1992**, 11, 223–258, and references cited therein.
- Barysz, M.; Jashari, G.; Lall, R. S.; Srivastava, V. K.; Trinajstić, N. The Distance Matrix of Molecules Containing Heteroatoms. In *Chemical Applications of Topology and Graph Theory*; King, R. B., Ed.; Elsevier: Amsterdam, 1983; pp 222–229, and references cited therein.
- Katrinsky, A. R.; Gordeeva, G. V. Traditional Topological Indices vs Electronic, Geometrical, and Combined Molecular Descriptors in QSAR/QSPR Research. *J. Chem. Inf. Comput. Sci.* **1993**, 33, 835–857, and references cited therein.
- Pettofrezzo, J. A. *Matrices and Transformations*; Prentice-Hall: Englewood Cliffs, NJ, 1966; pp 13–29.
- Müller, W. R.; Szymanski, K.; Knop, K. V.; Trinajstić, N. Molecular Topological Index. *J. Chem. Inf. Comput. Sci.* **1990**, 30, 160–163.
- Mihalić, Z.; Nikolić, S.; Trinajstić, N. Comparative Study of Molecular Descriptors Derived from the Distance Matrix. *J. Chem. Inf. Comput. Sci.* **1992**, 32, 28–37.
- Gutman, I. Some Properties of the Schultz Molecular Topological Index. *J. Chem. Inf. Comput. Sci.*, in press.