J Math Chem (2009) 45:607–626
DOI 10.1007/s10910-007-9345-4

ORIGINAL PAPER

On molecular similarity in communication theory of the chemical bond

Roman F. Nalewajski

Received: 16 November 2007 / Accepted: 27 November 2007 / Published online: 26 January 2008

Abstract The molecular similarity problem is addressed within the recently proposed information-theoretic (IT) approach to molecular electronic structure. In this Communication Theory of the chemical bond the direct criteria of the overall similarity are formulated in terms of the conditional-entropy (average noise, IT-covalency) and mutual-information (information flow, IT-ionicity) descriptors of the compared molecular communication systems, and the associated variational principles for their maximum entropy/information resemblance are formulated and discussed. Implications for molecular similarity from the parallel and sequential arrangements of the compared information channels of molecules or their fragments are investigated and tested on illustrative π-electron systems.

Keywords Bond covalency/ionicity · Chemical bonding · Communication theory · Information theory · Molecular communication systems · Parallel information channels · Similarity of molecules · Sequential information cascades · Theory of chemical bonds

1 Introduction

The concepts and methods of Information Theory (IT) [1–4] have been recently applied to diverse problems in the theory of electronic structure [5]. For example, the "stockholder" principle of Hirshfeld [6], for the local partition of the molecular electron density into pieces attributed to Atoms-in-Molecules (AIM), has been justified and extended [5,7–12], the entropy-displacement and information-distance densities have been used as novel diagnostic tools for exploring chemical bonds and electron

---

R. F. Nalewajski (✉)
Department of Theoretical Chemistry, Jagiellonian University, R. Ingardena 3, 30-060 Cracow, Poland
e-mail: nalewajs@chemia.uj.edu.pl

localization in molecular systems [5,6,13–15], the integral information-distance has been applied to quantify the Hammond similarity postulate of the theory of chemical reactivity [16], and the local, thermodynamic-like description of the electron equilibria in molecules has been proposed [17].

Within this IT-perspective on molecular electronic structure the AIM-resolved communication theory of the chemical bond has been developed for both the molecular system as a whole and its constituent fragments [5,18–26]. Its formulations in the orbital and local descriptions have also been proposed [27–31]. In this approach the entropy/information indices of the overall chemical bond multiplicity and its covalent/ionic composition reflect the promotion of the constituent free-atoms of the system promolecule [6] to their respective valence-states in the molecule, due to the presence of the remaining AIM. In the communication theory a molecule is interpreted as an information system, in which the molecular or "promolecular" electron probabilities are propagated ("scattered") *via* the network of chemical bonds connecting the system constituent atoms. The bond *entropy-covalency* (conditional entropy) descriptor of such a molecular channel measures its average communication "noise", i.e., the extra uncertainty in the distribution of the system valence electrons due to their delocalization *via* the network of the occupied Molecular Orbitals (MO). Accordingly, the overall *information-ionicity* (mutual information) index of all bonds in the system under consideration measures the amount of information flowing through the molecular information network, i.e., a fraction of the input information content which has survived the dissipation due to the bond covalency, i.e., due to the channel communication noise.

This communication analysis thus generates the information-flow perspective on classical issues in the theory of molecular electronic structure, dealing with the entropic origin and composition of chemical bonds. One now probes the bond covalency and ionicity through the information scattering in the molecular information system: the overall IT-covalency reflects the extra "noise" in the molecular communication channel due to the bond-formation process. It effectively lowers the information content of the final (output) probabilities of AIM, compared to the initial (input) probabilities, which measures the system overall IT-ionicity. In other words, the covalent component in the communication theory reflects the *delocalization* aspect of the valence electrons, via the system of chemical bonds generated by the occupied MO, while the ionic component describes the *localization* facet of the molecular electronic structure. These complementary entropy/information indices supplement the bond-order measures designed in the MO theory, e.g., [32–40]. The IT description gives a transparent account of the competition between the covalent and ionic components of the chemical bond, which also accords with the chemical intuitive expectations and predictions from the MO theory [5]. In several model systems the IT indices of chemical bonds have been shown to give rise to the dichotomous covalent and ionic contributions [21], which conserve the overall bond-order [10].

In chemistry one often compares the bonding patterns of molecules or their fragments and alternative, hypothetical or real, bond distributions of the given molecular system, e.g., the valence structures of the Valence-Bond (VB) theory [41] or the electron configurations in the Configuration Interaction (CI) approach to the electron correlation problem. Therefore, it is of interest to examine how such a comparison can be effected in the communication theory of chemical bonds. Several specific questions then naturally arise in this context. For example, can the entropy/information concepts and tools of IT be used to facilitate such "similarity" comparisons of molecular systems and their hypothetical VB-structures? Does the overall closeness of the IT-ionicity (or IT-covalency) descriptors of the two compared communication systems indeed imply the similarity of their information systems and chemical behavior? Can the consecutive and/or parallel connections [4,30] of the compared channels be used to diagnose their information and chemical resemblance? It is the main purpose of this work to address such and related issues, and to develop the relevant criteria for such applications of the communication theory in chemistry. Unless specified otherwise, in what follows the entropic quantities are measured, in bits, which correspond to the base 2 of the logarithmic measure of information [2]; the bold symbol $\mathbf{X}$ denotes the square or rectangular matrix, the bold-italic $\boldsymbol{X}$ stands for the row vector, while italic X corresponds to the scalar quantity.

2 Overall criteria of molecular information similarity

A comparison of the bonding structure, reflecting the bond connectivity pattern in molecular systems, can be carried out on both the global and local (diatom-resolved) levels. The former emphasizes the overall indicators of all bonds in two systems and their summary composition, while the latter focuses on distributions of bonds between the specified pairs of atoms. In this work we adopt the global IT approach, by examining the overall IT descriptors of the compared molecular communication systems.

Let us first briefly comment on the simplest, *one*-electron approach, in which one explores the information similarity in the probabilities of finding an electron on the systems constituent AIM. For the common set of atomic events in both systems, $i = 1, 2, \ldots, N$, the information similarity of two probability vectors $\boldsymbol{p} = \{p_i\}$ and $\boldsymbol{q} = \{q_j\}$, in the two compared molecular systems, respectively, is directly reflected by the information-distance index [relative (or cross) entropy, missing information, entropy deficiency] of Kullback and Leibler [3],

$$\Delta S(\boldsymbol{p}|\boldsymbol{q}) = \sum\nolimits_i p_i \log(p_i/q_i), \tag{1}$$

which has been successfully used in justifying the stockholder-rule of partitioning electron distributions [5–12], probing the electron distributions in molecules [5,13,14], and quantifying the Hammond postulate [16]. However, this *one*-electron concept is not discriminatory enough in many chemical applications comparing the bonding patterns of molecules and their fragments, since the origins of the chemical-bond phenomenon are rooted in the *two*-electron probabilities of simultaneously finding two electrons on specified atoms, $\mathbf{P} = \{P_{i,j} \equiv P(a_i, b_j)\}$. They are embodied in the conditional probabilities of finding one electron on atom j, when the other electron is known to have been located on aton j, $\mathbf{d} = \{d(j|i) = P_{i,j}/p_i\}$, where $p_i = \sum_j P_{i,j} \equiv P(a_i)$, which determine the communication networks of molecular information channels [5–31] (Fig. 1), and implied by the valence structures of VB theory [5,41]. Indeed, the criterion of Eq. 1 fails completely in the Hückel theory, when comparing the

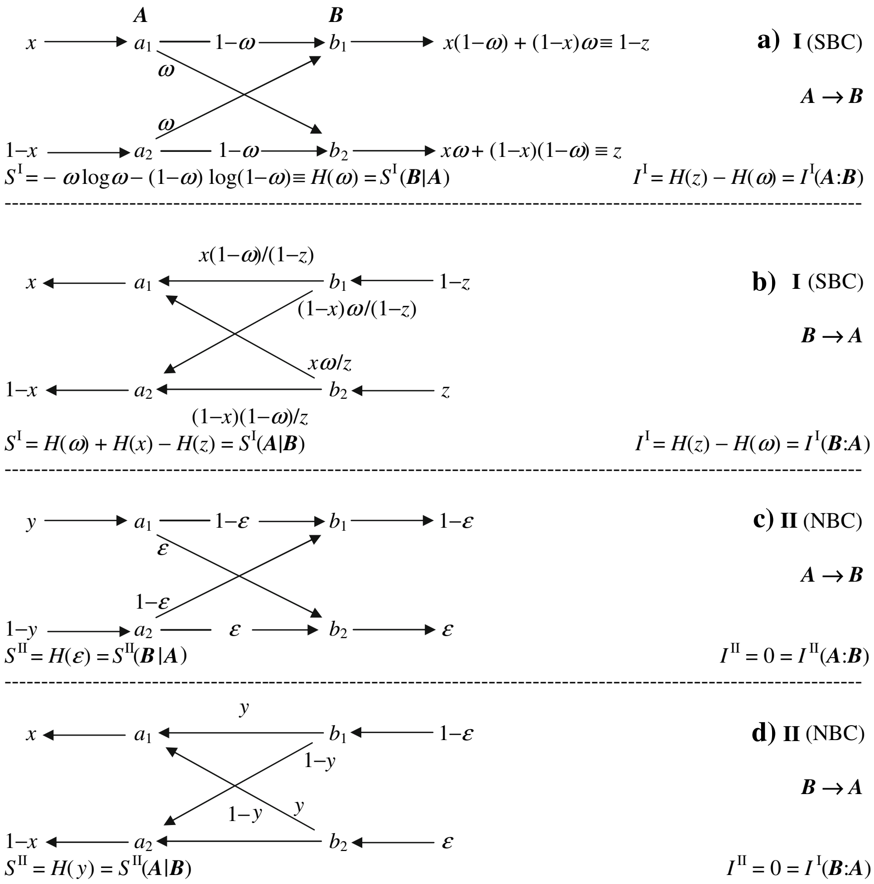


Fig. 1 The $\boldsymbol{A} \rightarrow \boldsymbol{B}$ binary channels defined by the conditional probability matrix $\mathbf{d} \equiv \mathbf{P}(\boldsymbol{B}|\boldsymbol{A}) = \{P(b_j|a_i) = P(a_i, b_j)/P(a_i)\}$: symmetric (SBC, Panel **a**) and non-symmetric (NBC, Panel **c**), and their conditional entropy (IT-covalency) and mutual information (IT-ionicity) descriptors, reported at the bottom of each diagram. In Panels **b** and **d** the corresponding reverse $\boldsymbol{B} \rightarrow \boldsymbol{A}$ channels are defined, which reproduce the input and output probabilities of their forward analogs. The conditional probabilities of the reverse channel, $\mathbf{d}^{rev.} \equiv \mathbf{P}(\boldsymbol{A}|\boldsymbol{B}) = \{P(a_i|b_j) = P(a_i, b_j)/P(b_j)\}$, are derived from the joint probabilities $P(a_i,\ b_j)$ implied by the forward channel: $P(a_i, b_j) = P(a_i)P(b_j|a_i)$. The vanishing I^{II} index in both NBC channels reflects the statistical independence of the input and output probabilities. It also shows that the flow of information in these information systems is purely IT-covalent, while SBC gives rise to a generally non-vanishing IT-ionic component

π-electron systems of alternant hydrocarbons and/or their fragments, due to the known equalization of the carbon π-electron densities in this approximation.

The average-noise (IT-covalency) in the molecular channel is measured by the conditional entropy [4] of the probabilities $\boldsymbol{B} = \{p_j = P(b_j)\}$ for the *molecular* "output" events $\{b_j\}$, given the probabilities $\boldsymbol{A} = \{p_i = P(a_i)\}$ of the *molecular* "input" events $\{a_i\}$, both determined by the system ground-state *one*-electron probabilities $\boldsymbol{p} = \boldsymbol{A} = \boldsymbol{B}$ [5,18–25],

$$S(\boldsymbol{B}|\boldsymbol{A}) = -\sum_i \sum_j P_{i,j} \log d(j|i) = S(\boldsymbol{A}|\boldsymbol{B}) \equiv S, \tag{2}$$

equal to the conditional entropy $S(\boldsymbol{A}|\boldsymbol{B})$, of the molecular input $\boldsymbol{A}$ given the molecular output $\boldsymbol{B}$, since $\mathbf{P} = \mathbf{P}^{\mathrm{T}}$. The complementary index of the IT-ionicity is given by the mutual information in the *promolecular* input probabilities $\boldsymbol{A}^0 = \boldsymbol{p}^0 = \{p_i^0\}$, for the valence electrons of the disconnected (separate, non-bonded) free-atomic fragments, and the *molecular* output distribution $\boldsymbol{p}$:

$$\begin{aligned} I(\boldsymbol{A}^0 : \boldsymbol{B}) &= \sum_i \sum_j P_{i,j} \log[P_{i,j}/(p_i^0 p_j)] = \sum_i \sum_j P_{i,j} \log[d(i|j)/p_i^0] \\ &= H(\boldsymbol{p}^0) - S(\boldsymbol{A}|\boldsymbol{B}) \equiv I, \end{aligned} \tag{3}$$

where the Shannon [2] entropy of the promolecular input $\boldsymbol{A}^0$

$$H(\boldsymbol{p}^0) = -\sum_i p_i^0 \log p_i^0. \tag{4}$$

Together these two overall bond components give rise to the total IT-index of all bonds in then system:

$$N(\boldsymbol{A}^0; \boldsymbol{B}) = I(\boldsymbol{A}^0 : \boldsymbol{B}) + S(\boldsymbol{A}|\boldsymbol{B}) = H(\boldsymbol{p}^0) = H(\boldsymbol{A}^0) \equiv N. \tag{5}$$

Illustrative examples of the binary information systems are given in Fig. 1, where the corresponding conditional entropy (IT-covalency) and mutual information (IT-ionicity) data are also reported.

It should be observed that the amount of information (3) also measures the *two*-electron information-distance $\Delta S(\mathbf{P}|\mathbf{P}^0)$ (Eq. 1), between the molecular *two*-electron probabilities $\mathbf{P}$, of the dependent (statistically correlated) two-electron events in atomic resolution, and the reference distribution $\mathbf{P}^0 = \{P_{i,j}^0 = p_i^0 p_j\}$ characterizing the independent (statistically uncorrelated) atomic events in the molecule and its promolecular reference, respectively.

It is natural to expect that the two electronic structures of a given molecular system, which generate close values of the overall IT-entropy/information descriptors of their respective information channels, i.e., comparable values of the overall IT covalency and ionicity, should be on average chemically similar, as reflecting a comparable number of all chemical bonds and their similar composition. Accordingly, substantial differences in these global bond-indicators should imply marked differences in chemical behaviour and/or geometric structures of the systems compared. For example, the total predictions [5] from the communication theory of the overall π-bond multiplicities $N = 2.585$ and $N = 3.000$ for the π-electron systems in benzene and cyclohexatriene, respectively, both of practically pure-covalent character, reflect different overall (covalent) multiplicity of π-bonds in both systems. Its diminished value in the aromatic system, in accordance with the contemporary views on the relative roles and competition between the σ and π bonds in benzene [42], is responsible for all important changes in the geometric structure and chemical reactivity of both these molecules.

One should realize, however, severe limitations of such a crude, global similarity criteria, of the minimum deviations between the overall bond components of Eqs. 2 and 3 for the information channels of systems I and II,

$$\min(S^{\mathrm{I}} - S^{\mathrm{II}})^2 \quad \text{and} \quad \min(I^{\mathrm{I}} - I^{\mathrm{II}})^2. \tag{6}$$

They may still admit a notable difference in the local distribution of bonds, which in turn generally imply different geometric structures and reactivity patterns of the compared molecular systems.

Another question, which should be also addressed at this point, is how effective are these global principles in enforcing a similarity between the two communication systems, when one (fixed) system is treated as the reference and the other (trial, test) system is modified variationally to satisfy the above minimum principles. We shall examine this "inverse" issue in the illustrative problem of comparing two binary systems shown in Fig. 1, both involving two events in their respective inputs and outputs. In Panels a and b of the figure the Symmetric Binary Channels (SBC) in the forward and reverse directions, respectively, are summarized for the common set of the input/outut probabilities defined in Panel a. Panels c and d of the figure represent the corresponding Non-symmetric Binary Chanels (NBC). As witnessed by the reported mutual information index $I^{\mathrm{I}} = H(z) - H(\omega)$, where $H(\xi) = -\xi \log\xi - (1-\xi)\log(1-\xi)$, stands for the binary entropy function, $0 \le \xi \le 1$, the mutually dependent input and output probabilities of SBC give rise to a generally non-vanishing amount of information flowing through this communication system. In the NBC case this index identically vanishes due to the statistical independence of the probability vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ in these information systems, $I^{\mathrm{II}} = 0$, thus signifying a total IT-covalent (noise) dissipation of the initial (input) information.

Let us now regard the SBC forward channel (Fig. 1a) as the model *reference* system I with its NBC analog (Fig. 1c) then providing a variationally modified *test* channel II. Therefore, the crossover probability ω of the reference network plays a role of the fixed parameter, while the optimum probability ε determining the trial communications is to be determined from the variational principles of Eq. 6. Our goal is to check, whether applying these information-similarity criteria gives rise to the optimum trial system II = I. The Euler equation for this unknown, which results from the least deviation of conditional entropies [in natural units (nats), when log = ln] of these two systems reads:

$$\frac{\partial[H(\omega) - H(\varepsilon)]^2\}}{\partial\varepsilon} = 0 \quad \Rightarrow \quad [H(\varepsilon) - H(\omega)]\ln\left(\frac{1-\varepsilon}{\varepsilon}\right) = 0. \tag{7}$$

It is satisfied, when at least one of these two factors vanishes: $H(\varepsilon) = H(\omega)$, i.e., $\varepsilon = (\omega, 1-\omega)$, or $1-\varepsilon = \varepsilon$, and hence $\varepsilon = {}^1\!/_2$. The second solution demonstrates, that the covalent-similarity criterion correctly identifies the symmetric limit of NBC as one of the systems least deviating from the assumed SBC reference.

The remaining optimum NBC networks, however, which exhibit the same conditional entropy content as the SBC channel, are seen to be generally different from the assumed reference. The elements of each row in their conditional probability

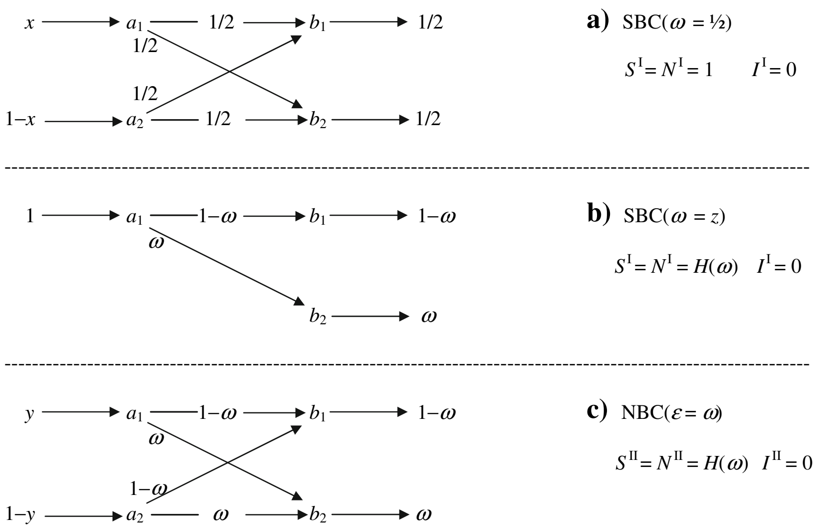


Fig. 2 Specific symmetric and non-symmetric binary channels representing optimum solutions of the global information-similarity criteria

matrices are permuted compared to the SBC matrix. One further observes, that these solutions do admit the identity of the trial and reference information systems. More specifically, Eq. 7 remains fulfilled, when both factors in the Euler equation vanish simultaneously: $\varepsilon = \omega = {}^1\!/\!_2$. This solution indeed corresponds to the symmetric communication system, NBC($\varepsilon = {}^1\!/\!_2$) = SBC($\omega = {}^1\!/\!_2$) (Fig. 2a), of the maximum IT-covalency $S^{\mathrm{I}} = S^{\mathrm{II}} = 1$ and the vanishing IT-ionicity. In the communication theory such network represents the symmetric, purely covalent bond, e.g., the σ-bond in hydrogen molecule or the π-bond in ethylene [5,18–22].

A similar conclusion follows from the second criterion of Eq. 6, of the least deviation between the IT-ionicities of two systems. Since the mutual information descriptor of the NBC channel identically vanishes one now selects this communication network as the reference and searches for the optimum SBC exhibiting the closest information ionicity. The corresponding Euler equation,

$$\frac{\partial[H(z) - H(\omega)]^2}{\partial\omega} = 0 \quad \Rightarrow \quad [H(\omega) - H(z)]\ln\left(\frac{1-\omega}{\omega}\right) = 0, \tag{8}$$

again gives $\omega = (z, 1-z)$ or $\omega = {}^1\!/\!_2$. As we have already observed above, the SBC($\omega = {}^1\!/\!_2$) indeed gives rise to $I^{\mathrm{I}} = I^{\mathrm{II}} = 0$, while the first solution $\omega = z$ implies $x = 1$, so that only one input is used in sending the signals throughout the SBC (Fig. 2b). Such non-symmetric use of SBC also generates the vanishing IT ionicity, as in the reference NBC system. Finally, the vanishing of the two factors in Eq. 8, when $\omega = z = {}^1\!/\!_2$, again leads to the purely covalent SBC, which may correspond to the symmetric σ-bond in H_2 or π-bond in C_2H_4.

It follows from the conservation of the overall bond order in communication theory (see Eq. 5), that for the fixed input probabilities only one bond component is independent. Therefore, the trial channel exhibiting the maximum closeness in the IT-covalent component to the reference system must also minimize the deviation in its IT-ionic component. It should be stressed that the requirement that both variational principles of Eq. 6 be simultaneously satisfied does not eliminate multiple solutions. For example, it follows from Eqs. 7 and 8 that $\varepsilon = \omega = z$ again implies a non-symmetric use of SBC, for $x = 1$ (Fig. 2b), which produces the same output probabilities as in the associated NBC of Fig. 2c.

Now, to summarize this short demonstration, we observe that the first similarity criterion of Eq. 6 gives rise to multiple solutions, which contain the identity case. Thus, although not uniquely identifying the full agreement between these two model communication networks, this criterion does not miss this perfect matching case either.

We finally observe that the simultaneous observance of the both variational principles of Eq. 6 also implies the least deviations in the overall bond index of Eq. 5:

$$\min(N^{\mathrm{I}} - N^{\mathrm{II}})^2, \tag{9}$$

where $N^{\mathrm{I}} = S^{\mathrm{I}} + I^{\mathrm{I}}$ and $N^{\mathrm{II}} = S^{\mathrm{II}} + I^{\mathrm{II}}$.

3 Implications from the combined information systems

Let us next examine possible implications for the molecular similarity issue from the complex information systems, which include the two compared communication networks as constituent parts. For example, one may envisage a use of the *sequential* cascade or the *parallel* arrangement of these two channels, which separately represent the probability scattering in the two molecules or their fragments. Such combined networks have recently been discussed in the context of the intermediate orbital transformations [30] and the local resolution of electron probabilities [31].

3.1 Parallel arrangement

First, let us briefly summarize the combination rules for the conditional-entropy and mutual-information descriptors of the complex system involving the parallel arrangement of two separate channels of Fig. 3a, which is shown in Fig. 3b [30]. The disconnected sub-channels I and II, which represent the compared molecules/fragments, are generated by the *intra*-subsystem conditional probabilities $\mathbf{d}^{\mathrm{I}} = \mathbf{P}_{\mathrm{I}}(\boldsymbol{B}^{\mathrm{I}}|\boldsymbol{A}^{\mathrm{I}})$ and $\mathbf{d}^{\mathrm{II}} = \mathbf{P}_{\mathrm{II}}(\boldsymbol{B}^{\mathrm{II}}|\boldsymbol{A}^{\mathrm{II}})$, respectively. They are subsequently combined in the parallel manner into a single complex channel defined by the effective conditional probabilities $\mathbf{d}^{\mathrm{III}} = \mathbf{P}_{\mathrm{III}}(\boldsymbol{B}|\boldsymbol{A})$, with the input group probabilities of two subsystems in the combined channel, $\boldsymbol{P}_G = (P_{\mathrm{I}}, P_{\mathrm{II}}) = (\lambda, 1 - \lambda)$, reflecting an overall participation of each subsystem in this two-lane complex. It should be stressed, that in the parallel arrangement the numbers of inputs/outputs of the compared channels can differ from one another.

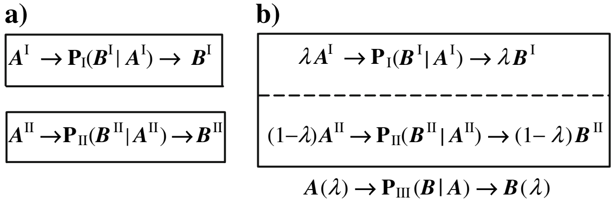


Fig. 3 Two compared (disconnected) information channels (Panel **a**) and their parallel arrangement (Panel **b**). Separate channels are characterized by the conditional probabilities $\mathbf{P}_{\mathrm{I}}(\boldsymbol{B}^{\mathrm{I}}|\boldsymbol{A}^{\mathrm{I}})$ and $\mathbf{P}_{\mathrm{II}}(\boldsymbol{B}^{\mathrm{II}}|\boldsymbol{A}^{\mathrm{II}})$, respectively, while the complex information system is described by the probability-scattering matrix $\mathbf{P}_{\mathrm{III}}(\boldsymbol{B}|\boldsymbol{A})$. The inputs(outputs) of the parallel system combine those of separate subsystems. The normalized input probabilities $\boldsymbol{A}^{\mathrm{I}} = \{A_i^{\mathrm{I}}\}$ and $\boldsymbol{A}^{\mathrm{II}} = \{A_j^{\mathrm{II}}\}$ of the separate sub-channels represent the conditional (*intra*-group) probabilities in the combined system: $\{A_i^{\mathrm{I}} \equiv P(i|\mathrm{I})\}$ and $\{A_j^{\mathrm{II}} \equiv P(j|\mathrm{II})\}$, exhibiting the relevant normalizations: $\sum_i P(i|\mathrm{I}) = \sum_j P(j|\mathrm{II}) = 1$. The absolute values of these probabilities in the complex system are obtained by multiplying these conditional probabilities by the appropriate group probability of the component channel in the parallel complex as a whole: $\boldsymbol{A}(\lambda) = [\{A_i = P_{\mathrm{I}}P(i|\mathrm{I}) = \lambda A_i^{\mathrm{I}}\}, \{A_j = P_{\mathrm{II}}P(j|\mathrm{II}) = (1-\lambda)A_j^{\mathrm{II}}\}$. They give rise to the parallel channel output probabilities $\boldsymbol{B}(\lambda) = \boldsymbol{A}(\lambda)\mathbf{P}_{\mathrm{III}}(\boldsymbol{B}|\boldsymbol{A})$

The normalized input and output probabilities of such parallel channel thus read:

$$\begin{aligned}\boldsymbol{A}(\lambda) = [\lambda\boldsymbol{A}^{\mathrm{I}}, (1-\lambda)\boldsymbol{A}^{\mathrm{II}}], \quad \boldsymbol{B}(\lambda) &= [\lambda\boldsymbol{A}^{\mathrm{I}}\mathbf{P}_{\mathrm{I}}(\boldsymbol{B}^{\mathrm{I}}|\boldsymbol{A}^{\mathrm{I}}), (1-\lambda)\boldsymbol{A}^{\mathrm{II}}\mathbf{P}_{\mathrm{II}}(\boldsymbol{B}^{\mathrm{II}}|\boldsymbol{A}^{\mathrm{II}})]\\ &= [\lambda\boldsymbol{B}^{\mathrm{I}}, (1-\lambda)\boldsymbol{B}^{\mathrm{II}}].\end{aligned} \tag{10}$$

Therefore, the conditional probabilities $\mathbf{P}_{\mathrm{III}}(\boldsymbol{B}|\boldsymbol{A})$ of the combined channel, which transform $\boldsymbol{A}(\lambda)$ into $\boldsymbol{B}(\lambda)$,

$$\boldsymbol{A}(\lambda)\mathbf{P}_{\mathrm{III}}(\boldsymbol{B}|\boldsymbol{A}) = \boldsymbol{B}(\lambda), \tag{11}$$

assume the block-diagonal form:

$$\mathbf{P}_{\mathrm{III}}(\boldsymbol{B}|\boldsymbol{A}) = \begin{bmatrix} \mathbf{P}_{\mathrm{I}}(\boldsymbol{B}^{\mathrm{I}}|\boldsymbol{A}^{\mathrm{I}}) & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\mathrm{II}}(\boldsymbol{B}^{\mathrm{II}}|\boldsymbol{A}^{\mathrm{II}}) \end{bmatrix}. \tag{12}$$

The grouping rules for the IT-covalent and IT-ionic bond components in such complex information system are in the spirit of those for the Shannon entropy of input-probabilities:

$$\begin{aligned} H(\boldsymbol{A}(\lambda)) &= -\sum_{\alpha=\mathrm{I,II}}\sum_{k\in\alpha} P_\alpha P(k|\alpha)\log[P_\alpha P(k|\alpha)] \\ &= -\sum_{\alpha=\mathrm{I,II}} P_\alpha \log P_\alpha - \sum_{\alpha=\mathrm{I,II}} P_\alpha\left[\sum_{k\in\alpha} P(k|\alpha)\log P(k|\alpha)\right] \\ &\equiv H(\boldsymbol{P}_G) + [\lambda H(\boldsymbol{A}^{\mathrm{I}}) + (1-\lambda)H(\boldsymbol{A}^{\mathrm{II}})]. \end{aligned} \tag{13}$$

Here, the first term $H(\boldsymbol{P}_G) = H(\lambda)$ represents the group-uncertainty, while the rest measures the $\boldsymbol{P}_G$-weighted mean value of the intra-group entropies for each separate subsystem. The combination formulas for the system conditional entropy and mutual-information similarly read:

$$\begin{aligned} S(\boldsymbol{B}(\lambda)|\boldsymbol{A}(\lambda)) &= -\sum_{\alpha=\mathrm{I,II}} P_\alpha(\lambda) \sum_{k,l\in\alpha} P(k|\alpha)P_\alpha(l|k)\log P_\alpha(l|k) \\ &= \sum_{\alpha=\mathrm{I,II}} P_\alpha(\lambda) S(\boldsymbol{B}^\alpha \big| \boldsymbol{A}^\alpha) \equiv S(\lambda). \end{aligned} \tag{14}$$

$$\begin{aligned} I(\boldsymbol{A}(\lambda):\boldsymbol{B}(\lambda)) &= \sum_{\alpha=\mathrm{I,II}} P_\alpha(\lambda) \sum_{k,l\in\alpha} P(k|\alpha)P_\alpha(l|k)\log \frac{P_\alpha(l|k)}{P_\alpha(\lambda)P(l|\alpha)} \\ &= \sum_{\alpha=\mathrm{I,II}} P_\alpha(\lambda) I(\boldsymbol{A}^\alpha : \boldsymbol{B}^\alpha) + H(\lambda) \equiv I(\lambda). \end{aligned} \tag{15}$$

Hence the grouping rule for the total bond index:

$$\begin{aligned} N(\boldsymbol{A}(\lambda);\boldsymbol{B}(\lambda)) &= S(\boldsymbol{B}(\lambda)|\boldsymbol{A}(\lambda)) + I(\boldsymbol{A}(\lambda):\boldsymbol{B}(\lambda)) \\ &= H(\boldsymbol{P}_G(\lambda)) + \sum_{\alpha=\mathrm{I,II}} P_\alpha(\lambda)[S(\boldsymbol{B}^\alpha \big| \boldsymbol{A}^\alpha) + I(\boldsymbol{A}^\alpha : \boldsymbol{B}^\alpha)] \\ &\equiv H(\lambda) + \sum_{\alpha=\mathrm{I,II}} P_\alpha(\lambda) N(\boldsymbol{A}^\alpha; \boldsymbol{B}^\alpha) \equiv N(\lambda). \end{aligned} \tag{16}$$

Expressing these bond-indices in terms of the separate subsystem quantities, $S(\boldsymbol{B}^\alpha|\boldsymbol{A}^\alpha) = S^\alpha$, $I(\boldsymbol{A}^\alpha : \boldsymbol{B}^\alpha) = I^\alpha$ and $N(\boldsymbol{A}^\alpha; \boldsymbol{B}^\alpha) = N^\alpha$, $\alpha = \mathrm{I, II}$, then gives the following functions of the input parameter λ:

$$\begin{aligned} &S(\lambda) = S^{\mathrm{II}} + \lambda(S^{\mathrm{I}} - S^{\mathrm{II}}), \quad I(\lambda) = H(\lambda) + \lambda I^{\mathrm{I}} + (1-\lambda) I^{\mathrm{II}}, \\ &N(\lambda) = H(\lambda) + \lambda N^{\mathrm{I}} + (1-\lambda) N^{\mathrm{II}}. \end{aligned} \tag{17}$$

The average noise (IT-covalency) of the parallel channel of Fig. 3b is thus linearly dependent upon the input parameter λ. Therefore, its derivative

$$\frac{\partial S(\lambda)}{\partial \lambda} = S^{\mathrm{I}} - S^{\mathrm{II}}, \tag{18}$$

vanishes only when the conditional entropies of the two compared molecular information systems are equal.

The related condition for the optimum value of the input parameter $\lambda = \lambda^{\max}$, for which the mutual information index (in nats) reaches the maximum value, reads:

$$\left.\frac{\partial I(\lambda)}{\partial \lambda}\right|_{\lambda^{\max}} = I^{\mathrm{I}} - I^{\mathrm{II}} + \left.\frac{\partial H(\lambda)}{\partial \lambda}\right|_{\lambda^{\max}} = I^{\mathrm{I}} - I^{\mathrm{II}} + \ln\left(\frac{1-\lambda^{\max}}{\lambda^{\max}}\right) = 0, \tag{19}$$

and hence,

$$\lambda^{max} = [1 + \exp(I^{\mathrm{II}} - I^{\mathrm{I}})]^{-1} \equiv (1 + \exp\Delta I)^{-1}. \tag{20}$$

Therefore, when manipulating the input signal in the parallel information complex build from the two compared information systems, one reaches the maximum mutual information for the input probabilities $\boldsymbol{P}_G^{max} = (\lambda^{max}, 1 - \lambda^{max})$. The corresponding value $I(\lambda^{max})$ then marks the parallel channel information *capacity* [4,5] (in nats):

$$C(\text{parallel}) = I(\lambda^{\max}) = \ln(1 + \exp\Delta I) - \Delta I[1 + \exp(-\Delta I)]^{-1}, \tag{21}$$

reaching the maximum value ln2 [nats] = 1 [bit] for $\Delta I = 0$. It follows from this optimum solution, that for the equal IT-ionicities of the compared information systems, when $\Delta I = 0$, the parallel complex gives rise to equal input-shares of both subsystems in the whole system: $\lambda^{max}(\Delta I = 0) = 1/2$. Moreover, $\lambda^{max}(\Delta I < 0) > 1/2$ and $\lambda^{max}(\Delta I > 0) < 1/2$, signify a non-symmetric use of the parallel system.

Therefore, for a series of trial channels II the degree of their IT-ionic similarity to the fixed reference communication system I can be measured by a magnitude of the deviation between the current value of λ^{max} from the capacity value for the perfect match, $\lambda^{max} = 1/2$, thus giving rise to the following variational principle:

$$\min(\lambda^{max} - 1/2)^2. \tag{22}$$

One similarly derives the optimum value of $\lambda = \lambda^*$, which maximizes the total IT bond-multiplicity $N(\lambda)$ (Eq. 17),

$$\lambda^* = [1 + \exp(N^{\mathrm{II}} - N^{\mathrm{I}})]^{-1} \equiv (1 + \exp\Delta N)^{-1}. \tag{23}$$

Again, for equal total IT bond-orders in the compared systems, $\lambda^*(\Delta N = 0) = 1/2$, and the increasing deviation of this input parameter for the trial system II and the fixed reference I signifies a decreasing resemblance between these channels in terms of the total bond index.

3.2 Sequential arrangement

In the consecutive information-cascade [4,30] of the two compared information systems, trial and reference, when the outputs of the elementary channel I constitute the inputs of channel II,

$$\boldsymbol{A} \to \mathbf{P}_{\mathrm{I}}(\boldsymbol{B}|\boldsymbol{A}) \to \boldsymbol{B} \to \mathbf{P}_{\mathrm{II}}(\boldsymbol{C}|\boldsymbol{B}) \to \boldsymbol{C}, \tag{24}$$

the role of the preceding step in the series is limited to shaping the input probabilities of the next step. This feature of the sequential information system is illustrated in Fig. 4 for the simplest case of two sub-channels $\mathbf{P}_{\mathrm{I}}(\boldsymbol{B}|\boldsymbol{A}) \equiv \{P_{\mathrm{I}}(j|i)\}$ and $\mathbf{P}_{\mathrm{II}}(\boldsymbol{C}|\boldsymbol{B}) \equiv \{P_{\mathrm{II}}(k|j)\}$ in the series. Here, $\boldsymbol{A} = \{A_i\}$, $\boldsymbol{B} = \{B_j\}$, and $\boldsymbol{C} = \{C_k\}$ group the input, intermediate, and output probabilities, respectively. Therefore, the

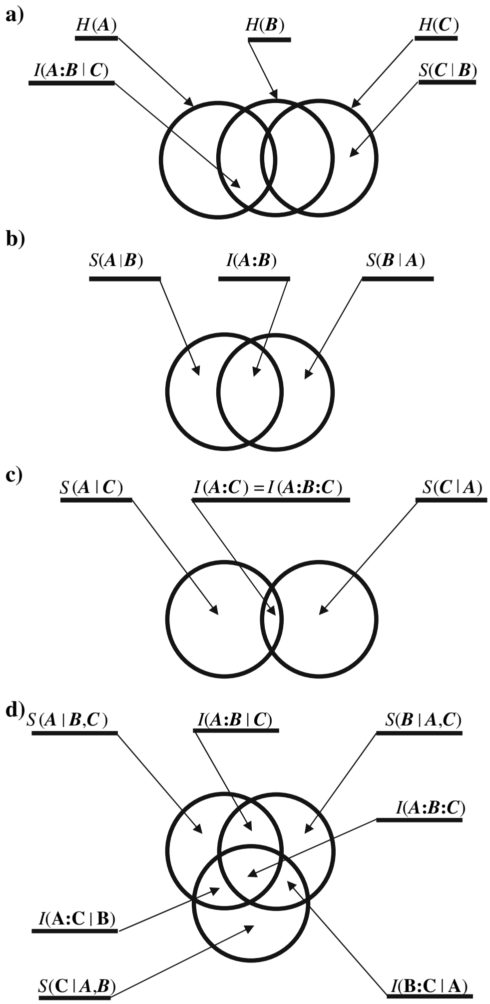


Fig. 4 Dependent probability distributions $\boldsymbol{A} = \{A_i = P(a_i)\}$, $\boldsymbol{B} = \{B_j = P(b_j)\}$, and $\boldsymbol{C} = \{C_k = P(c_k)\}$ of the two-step sequential information channel, $\boldsymbol{A} \rightarrow \mathbf{P}_{\rm I}(\boldsymbol{B}|\boldsymbol{A}) \rightarrow \boldsymbol{B} \rightarrow P_{\rm II}(\boldsymbol{C}|\boldsymbol{B}) \rightarrow \boldsymbol{C}$, which consists of the consecutive arrangement of two sub-channels $\mathbf{P}_{\rm I}(\boldsymbol{B}|\boldsymbol{A})$ and $\mathbf{P}_{\rm II}(\boldsymbol{C}|\boldsymbol{B})$ giving rise to the resultant channel $\mathbf{P}_{\rm III}(\boldsymbol{C}|\boldsymbol{A}) = \mathbf{P}_{\rm I}(\boldsymbol{B}|\boldsymbol{A})\mathbf{P}_{\rm II}(\boldsymbol{C}|\boldsymbol{A})$. These 3 channels give rise to the overlapping Shannon entropies $H(\boldsymbol{A})$, $H(\boldsymbol{B})$, and $H(\boldsymbol{C})$ of Panels (**a**–**c**), which depict their mutual arrangement for the specific case of the sequential arrangement of the two sub-channels compared, when the output of the first stage constitutes the input of the second stage in the series. Therefore, a dependence $\boldsymbol{C}(\boldsymbol{A}) = \boldsymbol{C}(\boldsymbol{B}(\boldsymbol{A}))$, as schematically shown in Panel **a**. This diagram shows that the overlap region between $H(\boldsymbol{A})$ and $H(\boldsymbol{C})$, representing the mutual information $I(\boldsymbol{A}{:}\boldsymbol{C}) = I(\boldsymbol{A}{:}\boldsymbol{B}{:}\boldsymbol{C}) + I(\boldsymbol{A}{:}\boldsymbol{C}|\boldsymbol{B})$ (Panels **a**, **c**), is completely contained in $I(\boldsymbol{A}{:}\boldsymbol{B})$ (Panels **a**, **b**), thus implying the vanishing mutual information in $\boldsymbol{A}$ and $\boldsymbol{C}$, conditional on $\boldsymbol{B}{:}I(\boldsymbol{A}{:}\boldsymbol{C}|\boldsymbol{B}) = 0$. In other words, for the sequential arrangement of two subchannels the mutual information in the peripheral probabilities is equal to that in three probability distributions: $I(\boldsymbol{A}{:}\boldsymbol{C}) = I(\boldsymbol{A}{:}\boldsymbol{B}{:}\boldsymbol{C})$. This is not the case for general dependencies between three probability vectors, which correspond to the mutual arrangement of three subsystem entropies shown in Panel **d**. The diagram **b** and **c** show a gradual loss of information (increase in entropy) at each step, as reflected by the difference between conditional-entropies $S(\boldsymbol{A}|\boldsymbol{C}) > S(\boldsymbol{A}|\boldsymbol{B})$. It implies the associated lowering of the amount of information flowing through the cascade, measured by the corresponding mutual-information quantities, $I(\boldsymbol{A}{:}\boldsymbol{B}) > I(\boldsymbol{A}{:}\boldsymbol{C})$, given by the fractions of the initial amount of information $H(\boldsymbol{A})$ at the cascade input

effective output probabilities of the cascade as a whole is determined by the intermediate set of probabilities $\boldsymbol{B} = \boldsymbol{A}\mathbf{P}_{\mathrm{I}}(\boldsymbol{B}|\boldsymbol{A})$ of the common sets of "events" $\{b_j\}$, which determine both the output of I and the input of II:

$$\boldsymbol{C} = \boldsymbol{B}\mathbf{P}_{\mathrm{II}}(\boldsymbol{C}|\boldsymbol{B}) = \boldsymbol{A}[\mathbf{P}_{\mathrm{I}}(\boldsymbol{B}|\boldsymbol{A})\mathbf{P}_{\mathrm{II}}(\boldsymbol{C}|\boldsymbol{B})] \equiv \boldsymbol{A}\mathbf{P}_{\mathrm{III}}(\boldsymbol{C}|\boldsymbol{A}), \tag{25}$$

with the resultant conditional probabilities $\mathbf{P}_{\mathrm{III}}(\boldsymbol{C}|\boldsymbol{A}) \equiv \{P_{\mathrm{III}}(k|i)\}$ of communications between the inputs $\{a_i\}$ of I and the outputs $\{c_j\}$ of II.

In Fig. 4 the mutual dependence between probability vectors $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ is reflected by the overlap between the corresponding circles representing the associated Shannon entropies $H(\boldsymbol{A})$, $H(\boldsymbol{B})$ and $H(\boldsymbol{C})$. Their mutual arrangement in the consecutive cascade is depicted in Panel a, while Panel d illustrates the conditional-entropy (S) and mutual-information (I) quantities, which appear in a general case of three dependent probability distributions [4]. In the sequential cascade the mutual-information $I(\boldsymbol{A}{:}\boldsymbol{C})$, reflected by the overlap between $H(\boldsymbol{A})$ and $H(\boldsymbol{C})$, is totally contained in the mutual information $I(\boldsymbol{A}{:}\boldsymbol{B})$, represented by the overlap between $H(\boldsymbol{A})$ and $H(\boldsymbol{B})$, since the whole dependence of $\boldsymbol{C}$ on $\boldsymbol{A}$ is due to $\boldsymbol{B}$.

Therefore, the mutual information in the peripheral distributions must be equal to the mutual information in all three probability vectors: $I(\boldsymbol{A}:\boldsymbol{C}) = I(\boldsymbol{A}:\boldsymbol{B}:\boldsymbol{C}) \leq I(\boldsymbol{A}{:}\boldsymbol{B}) < H(\boldsymbol{A})$. The preceding inequalities express a successive loss of information due to consecutive steps of the cascade. This effect can be alternatively represented by inequalities between the missing amounts of information at each step provided by the associated conditional-entropy indices, which are also shown in the figure: $S(\boldsymbol{A}|\boldsymbol{C}) = H(\boldsymbol{A}) - I(\boldsymbol{A}{:}\boldsymbol{C}) > S(\boldsymbol{A}|\boldsymbol{B}) = H(\boldsymbol{A}) - I(\boldsymbol{A}{:}\boldsymbol{B})$, where $S(\boldsymbol{A}|\boldsymbol{C})$ measures the information loss for the cascade as a whole, while $S(\boldsymbol{A}|\boldsymbol{B})$ reflects the information lost in its first stage. These inequalities follow from the elementary information quantities for a general inter-dependencies between the three probability distributions of Fig. 4d:

$$S(\boldsymbol{A}|\boldsymbol{C}) = S(\boldsymbol{A}|\boldsymbol{B}) + I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C}) - I(\boldsymbol{A}{:}\boldsymbol{C}|\boldsymbol{B}) = S(\boldsymbol{A}|\boldsymbol{B}) + I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C}), \tag{26}$$

since for the sequential arrangement of the two compared sub-channels the explicit dependence between peripheral probabilities, which does not result from their dependence on $\boldsymbol{B}$, $I(\boldsymbol{A}{:}\boldsymbol{C}|\boldsymbol{B})$, must identically vanish. The preceding equation expresses the so called *stage-additivity* of the information loss (conditional entropy) in the sequential cascade [30]. Thus, the mutual information $I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C})$, reflecting a dependence of $\boldsymbol{A}$ on $\boldsymbol{B}$, which cannot be attributed to $\boldsymbol{A}[\boldsymbol{C}(\boldsymbol{B})]$, represents the extra loss of the initial information in the second sub-channel, $I(\boldsymbol{A}{:}\boldsymbol{B})$.

As we have already emphasized in Section 2, the overall conditional entropy (S, IT-covalency) and mutual-information (I, IT-ionicity) descriptors of a given communication channel describe how the channel input information of Eq. 5 is partitioned at its output into the dissipated and surviving parts of the initial amount of information (N, total bond index). Therefore, in terms of these descriptors of information channels the compared molecular systems are globally "similar" when these proportions are comparable,

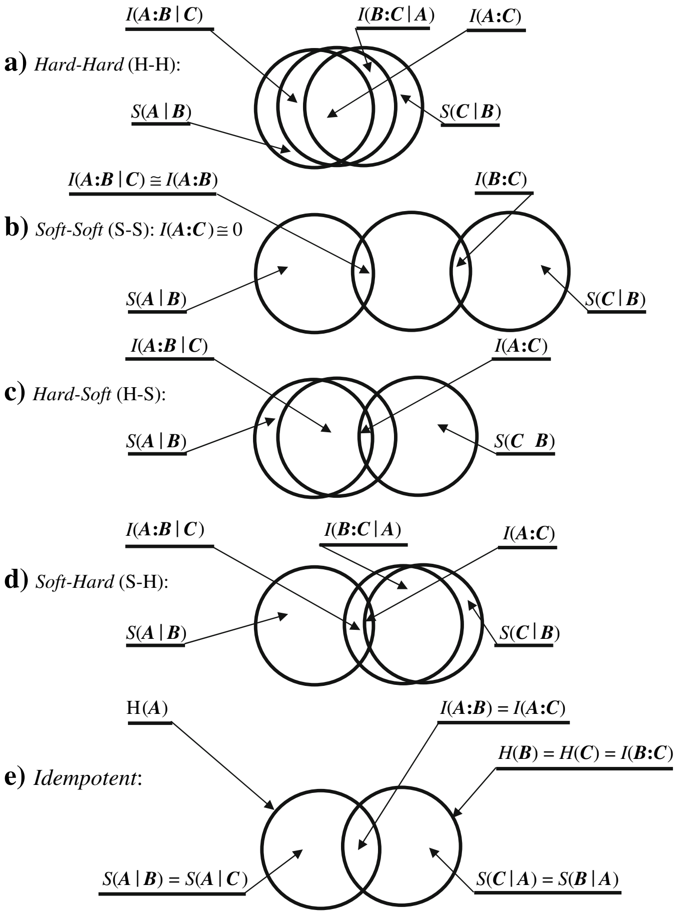


Fig. 5 Qualitative diagrams of possible arrangements of the entropy circles in a sequential cascade of two elementary molecular channels: IT-"similar" (Panels **a**, **b**) and IT-"dissimilar" (Panels **c**, **d**) communication systems. The diagram of Panel e represents the sequential cascade involving a repetition of the idempotent channel, $\mathbf{d} \equiv \mathbf{P}(\boldsymbol{B}|\boldsymbol{A}) = \mathbf{P}(\boldsymbol{C}|\boldsymbol{B})$, when $\mathbf{P}(\boldsymbol{C}|\boldsymbol{A}) = \mathbf{d}^2 = \mathbf{d}$

$$S(\boldsymbol{A}|\boldsymbol{B})/I(\boldsymbol{A}{:}\boldsymbol{B}) \equiv S^{\mathrm{I}}/I^{\mathrm{I}} \approx S(\boldsymbol{B}|\boldsymbol{C})/I(\boldsymbol{B}{:}\boldsymbol{C}) \equiv S^{\mathrm{II}}/I^{\mathrm{II}}. \tag{27}$$

Accordingly, the two systems would be regarded as "dissimilar", when these proportions are notably different, e.g., $S^{\mathrm{I}}/I^{\mathrm{I}} \ll S^{\mathrm{II}}/I^{\mathrm{II}}$. In qualitative diagrams of Fig. 5 we have represented possible similarity (Panels a, b) and dissimilarity (Panels c, d) cases in the sequential cascades, with the noise/information proportions of the compared networks reflected by the corresponding S and I areas in the plots.

One should distinguish the following two general categories of molecular communication systems, as characterized by the ratio of Eq. 27. The strongly *ionic* [IT-"hard" (H), to a large extent deterministic] channels, in which $S^{\alpha} < I^{\alpha}$, preserve a substantial fraction of their initial information content and exhibit a high degree of localization the system valence electrons. The other class of systems represents a strongly covalent

[IT-"soft" (S), strongly scattering] channels, with a dominating electron-delocalization ("noise") component, $S^{\alpha} > I^{\alpha}$, and hence only a small amount of the initial information content surviving in the channel output probabilities. Examples of such general communication networks are schematically depicted in Fig. 5, where alternative combinations of such elementary channels in the sequence are qualitatively examined. In the last panel of the figure the special case of a repetition of the same (*idempotent*) information system in the cascade is shown. This channel corresponds to the idempotent conditional-probability matrix $\mathbf{d} \equiv \mathbf{P}(\boldsymbol{B}|\boldsymbol{A}) = \mathbf{P}(\boldsymbol{C}|\boldsymbol{B}) : \mathbf{d}^2 = \mathbf{d}$, so that there is no additional amount of noise created (information lost) in the second stage of such a repetitive sequential system, compared to that already generated in the first stage.

The H–H (a) and S–S (b) panels in the figure deal with two cases of the information-similarity, while its H–S (c) and S–H (d) panels focus on two cases of the information-dissimilarity between the two compared communication channels. It follows from the figure (see also Table 1) that only the sequence of two ionic (H) channels, or a repetition of a single (ionic) idempotent information system (Panel e), can produce a substantial amount of the cascade mutual information. Indeed, the sequential cascade of two similar, strongly ionic ("hard") species, with $S(\boldsymbol{A}|\boldsymbol{B}) < I(\boldsymbol{A}{:}\boldsymbol{B})$ and $S(\boldsymbol{C}|\boldsymbol{B}) < I(\boldsymbol{B}{:}\boldsymbol{C})$, conserves in $\boldsymbol{C}$ a high proportion of the initial information content $H(\boldsymbol{A})$, as reflected by the sizable amount of the mutual information (ionicity) area $I(\boldsymbol{A}:\boldsymbol{C}) \equiv I^{\mathrm{III}}$ and a medium-size of the complementary (noise, covalency) areas $S(\boldsymbol{A}|\boldsymbol{C}) = H(\boldsymbol{A}) - I^{\mathrm{III}}$ and $S(\boldsymbol{C}|\boldsymbol{A}) = H(\boldsymbol{C}) - I^{\mathrm{III}}$. The other S–S case of the channel similarity (Panel b), with $S(\boldsymbol{A}|\boldsymbol{B}) > I(\boldsymbol{A}{:}\boldsymbol{B})$ and $S(\boldsymbol{C}|\boldsymbol{B}) > I(\boldsymbol{B}{:}\boldsymbol{C})$, gives rise to a practically vanishing amount of information flowing through the cascade [no overlap between $H(\boldsymbol{A})$ and $H(\boldsymbol{C})$ circles], and a large conditional entropies $S(\boldsymbol{A}|\boldsymbol{B})$ and $S(\boldsymbol{C}|\boldsymbol{B})$. This conclusion also applies to the repetition of a soft (non-idempotent) molecular channel. All remaining mixed (dissimilarity) combinations of H and S elementary channels also generate in the sequential arrangement almost a pure-covalent effective information dissipation of the cascade. The least amount of the cascade initial information is expected to be preserved at the output of the sequence of two strongly covalent (IT-soft) channels of Panel b, when $\boldsymbol{C}$ remains practically independent of $\boldsymbol{A}$. Panels c and d of Fig. 5 summarize the two information-dissimilarity cases. A reference to Fig. 5c indicates, that information system I is strongly ionic (electron-localized), as reflected by a relatively high value of the $I(\boldsymbol{A}{:}\boldsymbol{B}) > S(\boldsymbol{A}|\boldsymbol{B})$ area, while the second channel exhibits the dominance of the communication noise, i.e., the information dissipation due to electron delocalization, over the information-preservation due to the electron localization: noise, $I(\boldsymbol{C}:\boldsymbol{B}) < S(\boldsymbol{C}|\boldsymbol{B})$. As qualitatively argued in Fig. 5, the information similarity of the two ionic networks (Panel a) generates a relatively high value of the amount of information I^{III} flowing through the cascade as a whole, and hence a relatively low value of the complementary parts $I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C})$ of $I^{\mathrm{I}} = I^{\mathrm{III}} + I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C})$ and $I(\boldsymbol{B}{:}\boldsymbol{C}|\boldsymbol{A})$ of $I^{\mathrm{II}} = I^{\mathrm{III}} + I(\boldsymbol{B}{:}\boldsymbol{C}|\boldsymbol{A})$, compared to the information dissimilarity cases of Panels c and d. The H–S pair of Panel c is seen to give rise to generally low value of I^{III} and hence to high values of the complementary quantities $I_1^{cond.} = I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C})$ or $S(\boldsymbol{C}|\boldsymbol{A}) = H(\boldsymbol{A}) - I^{\mathrm{III}}$. Accordingly, the S–H pair of Panel d is expected to generate generally low values of both I^{III} and $I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C})$, and hence high values of the complementary quantities $I_2^{cond.} = I(\boldsymbol{B}{:}\boldsymbol{C}|\boldsymbol{A})$ and $S(\boldsymbol{C}|\boldsymbol{A})$.

Table 1 Summary of qualitative trends in magnitudes of the entropy/information indices of sequential communication systems consisting of two compared elementary channels of varying degree of the information ionicity/covalency (see Panels a–d of Fig. 5)

Sequential cascade (Fig. 5)	Bond indices					
	$I(\boldsymbol{A}{:}\boldsymbol{C})$	$S(\boldsymbol{C}\vert\boldsymbol{A})$	$I(\boldsymbol{A}{:}\boldsymbol{B}\vert\boldsymbol{C})$	$I(\boldsymbol{B}{:}\boldsymbol{C}\vert\boldsymbol{A})$	$S(\boldsymbol{A}\vert\boldsymbol{B})$	$S(\boldsymbol{C}\vert\boldsymbol{B})$
H–H	Medium	Medium	Small	Small	Small	Small
S–S	$\cong 0$	$\cong S(\boldsymbol{C})$	$\cong I(\boldsymbol{A}{:}\boldsymbol{B})$	$\cong I(\boldsymbol{B}{:}\boldsymbol{C})$	Large	Large
H–S	Small	Large	Large	Small	Small	Large
S–H	Small	Large	Small	Large	Large	Small

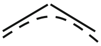
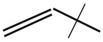
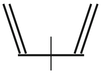
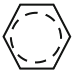
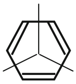
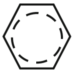
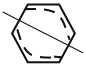
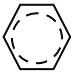
In Table 1 we have summarized all these qualitative predictions, which should help establishing the type of information/entropy matching of the elementary communication systems of the compared molecular systems in the sequential probability scattering network. It follows from these qualitative predictions that each case of the subsystem similarity (H–H or S–S) or dissimilarity (H–S or S–H) has a distinct combination of sizes of these complementary entropy/information bond-indices. Within a given class identified by such a qualitative IT- “fingerprint” the predicted values of the cascade amount of information $I(\boldsymbol{A}{:}\boldsymbol{C})$ reflect a relative similarity in the compared series of compounds. In other words, the higher the value of I^{III} [lower the value of $I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C})$] the more similar are the compared information systems. The other conditional mutual-information index, $I(\boldsymbol{B}{:}\boldsymbol{C}|\boldsymbol{A}) = I^{\mathrm{II}} - I^{\mathrm{III}} \equiv I_2^{cond.}$ can alternatively be used as the similarity indicator.

In the next section we shall apply these IT criteria within the Hückel theory, to compare the π-electron systems of three simple alternant hydrocarbons [5], allyl, butadiene, and benzene, which exhibit a varying degree of the bond delocalization, against their selected (hypothetical) VB-structures, which involve an increased electron localization, i.e., a more emphasized bond alternation, compared to the molecular information channel.

4 Illustrative numerical applications

The representative *molecular* communication systems of π-electrons in ethylene, allyl, butadiene, and benzene in the Hückel approximation have been reported elsewhere (see, e.g., [5]). These elementary channels also provide the building blocks for generating the model information channels of the hypothetical VB-structures of larger systems, the similarity of which to the corresponding molecular channels will be examined in terms of the conditional-entropy (S) and mutual-information (I) indices of the associated sequential cascades of Eq. 24. A variety of the predicted IT bond indices is summarized in Table 2. Both positions of the molecular channels in the cascade have been tested. As explicitly indicated in the table, the information systems of allyl and butadiene have been placed as the first stages in their respective cascades, while the benzene communication network follows the model channels for the current VB structure $\boldsymbol{X} = (X_1|X_2|\ldots)$, where $\boldsymbol{X} = \{X_\alpha\}$ groups the mutually decoupled (closed, non-bonded) molecular fragments of the molecule. The corresponding VB

Table 2 Conditional-entropy (S) and mutual-information (I) indices (in bits) of the sequential information cascades (III, *casc.*) consisting of the communication channels of the molecular (*mol.*) π-electron systems in selected alternant hydrocarbons (Hückel theory) and model channels of their alternative VB-structures (*VB*). The extra (thin) lines in VB-structures separate the mutually closed (non-bonded) fragments, thus delineating the extent of their allowed electron delocalization. The conditional mutual-information indices $I_1^{cond.} = I(\boldsymbol{A}{:}\boldsymbol{B}|\boldsymbol{C}) = I(\boldsymbol{A}{:}\boldsymbol{B}) - I(\boldsymbol{A}{:}\boldsymbol{C}) = I^{\mathrm{I}} - I^{casc.}$ and $I_2^{cond.} = I(\boldsymbol{B}{:}\boldsymbol{C}|\boldsymbol{A}) = I(\boldsymbol{B}{:}\boldsymbol{C}) - I(\boldsymbol{A}{:}\boldsymbol{C}) = I^{\mathrm{II}} - I^{casc.}$ are reported in the two last columns of the table

Compared systems:			Bond indices							
Molecule	Cascade, VB-structure		S^{VB}	I^{VB}	$S^{mol.}$	$I^{mol.}$	$S^{casc.}$	$I^{casc.}$	$I_1^{cond.}$	$I_2^{cond.}$
Allyl (I):	→allyl→VB→									
		(ethylene | carbon)	0.67	0.92	1.52	0.06	1.55	0.04	0.023	0.880
Butadiene (I):	→butadiene→VB→									
		(ethylene | ethylene)	1.00	1.00	1.94	0.06	1.95	0.05	0.004	0.948
		(allyl | carbon)	1.14	0.86	1.94	0.06	1.98	0.02	0.036	0.980
Benzene (II):	→VB→benzene→									
		cyclohexatriene = (ethylene | ethylene | ethylene)	1.00	1.59	2.55	0.03	2.56	0.02	1.563	0.012
		(allyl | allyl)	1.52	1.06	2.55	0.03	2.57	0.01	1.040	0.021
		(butadiene | ethylene)	1.63	0.96	2.55	0.03	2.57	0.01	1.606	0.023

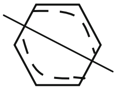
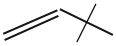
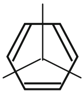
channels, defined by their conditional probability matrix $\mathbf{d}^{VB} = \{\mathbf{d}_\alpha \delta_{\alpha,\beta}\}$, have been constructed using the communication networks of individual subsystems $\boldsymbol{X}$, defined by the respective subsystem conditional probabilities $\{\mathbf{d}_\alpha\}$. For the alternant hydrocarbons all probability vectors combine equal molecular/promolecular probabilities for each carbon atom, $\boldsymbol{A} = \boldsymbol{A}^0 = \boldsymbol{B} = \boldsymbol{C} = \{1/n\}$, where n stands for the number of atoms in the π-electron system. This determines the overall bond index for each elementary channel (Eq. 5): $N^{\mathrm{I}} = S^{\mathrm{I}} + I^{\mathrm{I}} = N^{\mathrm{II}} = S^{\mathrm{II}} + I^{\mathrm{II}} = H(\boldsymbol{A}^0) = \log_2 n$.

We begin the discussion of these results by examining the IT bond indices for individual molecular and VB channels. It follows from the table that, as intuitively expected, the information channels of VB structures, in which the electron delocalization is constrained relative to the associated molecular system, are characterized by a distinctly higher IT-ionicity (lower IT-covalency), compared to the corresponding molecular descriptors. The molecular data are seen to predict almost purely covalent π-bonds between the system carbon atoms, while the increasing localization of bonds gives rise to their increasing IT-ionicity. For example, in the benzene VB-structures the highest ionicity index corresponds to the cyclohexatriene structure, involving 3 localized bonds, i.e., the hypothetical ethylene subsystems, while the (allyl|allyl) and (butadiene|ethylene) structures, which give rise to a similar bond composition, exhibit a much lower ionicity contribution due to a partial delocalization of π-bonds already accounted for.

Let us now focus on entries in Table 2, which report the cascade descriptors. It follows from the table that all pairs of the information cascades including the molecular and VB stages correspond to the mixed pairs of Figs. 5c, d, when there is a marked difference between the ratios of Eq. 27 for the two systems compared. This explains rather small values of the predicted cascade mutual information index $I^{casc.} = I(\boldsymbol{A}:\boldsymbol{C})$, which can be thus considered as valid similarity indicator within a given class of cascades distinguished in Fig. 5: S–H [(allyl, butadiene)-VB] or H–S (VB-benzene).

It should be realized at this point that there are two major factors determining the similarity of these information systems: the symmetry of the trial VB structure and the extend of the bond delocalization it allows. These factors often act in opposite directions. For example, the (allyl|carbon) VB structure of butadiene involves a higher level of the π-electron delocalization compared to the (ethylene|ethylene) case, which increases its similarity to the butadiene π-system. However, such more delocalized structure breaks the symmetry of the π-electron system, which decreases its resemblance to the molecule. The same competition between delocalization and symmetry aspect of the electronic structure is involved in the case of the (butadiene|ethylene) structure of benzene. The $I^{casc.}$ results for the sequential systems involving benzene show that the symmetry constitutes the decisive factor: the cyclohexatriene structure, which preserves the symmetry the most, gives the highest value of $I^{casc.}$, in comparison to the remaining structures involving a higher degree of the bond delocalization but a diminished symmetry. It is also reflected by a lower value of the complementary

Table 3 Illustrative examples of the molecular quadratic cascades, of two identical stages, representing the H–H (Fig. 2a), S–S (Fig. 2b) and idempotent (Fig. 2d) stages of information scattering and their complementary entropy/information descriptors (in bits)

Type	Stage	Bond indices:			
		$I^{\mathrm{I}} = I(\boldsymbol{A}:\boldsymbol{B})$, $I^{\mathrm{II}} = I(\boldsymbol{B}:\boldsymbol{C})$	$S^{\mathrm{I}} = S(\boldsymbol{B}\vert\boldsymbol{A})$, $S^{\mathrm{II}} = S(\boldsymbol{C}\vert\boldsymbol{B})$	$I^{\mathrm{III}} = I(\boldsymbol{A}:\boldsymbol{C})$	$S^{\mathrm{III}} = S(\boldsymbol{C}\vert\boldsymbol{A})$
H–H		1.061	1.524	1.003	1.582
S–S	allyl	0.061	1.524	0.003	1.582
Idempotent		0.918	0.667	0.918	0.667
		1.585	1.000	1.585	1.000

index $I_2^{cond.}$. The other conditional mutual-information index $I_1^{cond.}$ is not relevant for such comparisons, due to dramatically different values of $I^{\mathrm{I}} = I^{VB}$ for the trial VB structures. The same conclusion follows from an inspection of the butadiene results: the symmetry-preserving (ethylene|ethylene) structure gives the highest $I^{casc.}$ (lowest $I_1^{cond.}$) index, despite containing the least degree of delocalization. From what we know about the bond localization in butadiene, this structure is indeed the closest to the relatively localized π-bond pattern in this molecule [5].

Illustrative examples of the model H–H, S–S and idempotent cascades are reported in Table 3. These *quadratic* cascades involve identical stages I and II, with the idempotent channels being generated by the VB structures involving ethylene fragments of allyl and benzene. These examples are seen to confirm the qualitative conjectures of Fig. 2 and Table 1. For example, only the H–H and idempotent channels are seen to preserve a sizable part of the cascade initial amount of information, while the S–S combination gives a practically vanishing bond-ionicity, i.e., an almost pure bond-covalency. Clearly, the same result of a very low magnitude of the cascade information flow can be expected in the quadratic information systems involving the remaining strongly covalent (S) π-channels of butadiene and benzene.

References

1. R.A. Fisher, Proc. Cambridge Phil. Soc. **22**, 700 (1925); see also: B.R. Frieden, *Physics from the Fisher Information—A Unification* (Cambridge University Press, Cambridge, 2000)
2. C.E. Shannon, Bell System Tech. J. **27**, 379, 623 (1948); see also: C.E. Shannon, W. Weaver, *The Mathematical Theory of Communication* (University of Illinois, Urbana, 1949)
3. S. Kullback, R.A. Leibler, Ann. Math. Stat. **22**, 79 (1951); see also: S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959)
4. N. Abramson, *Information Theory and Coding* (McGraw-Hill, New York, 1963)
5. R.F. Nalewajski, *Information Theory of Molecular Systems* (Elsevier, Amsterdam, 2006) and refs. therein
6. F.L. Hirshfeld, Theoret. Chim. Acta (Berl.) **44**, 129 (1977)
7. R.F. Nalewajski, Int. J. Mol. Sci. **3**, 237 (2002); Adv. Quant. Chem. **43**, 119 (2003); Chem. Phys. Lett. **372**, 28 (2003); **375**, 196 (2003)
8. R.F. Nalewajski, R.G. Parr, Proc. Natl. Acad. Sci. USA **97**, 8879 (2000); J. Phys. Chem. A **105**, 7391 (2001)
9. R.F. Nalewajski, R. Loska, Theoret. Chem. Acc. **105**, 374 (2001)
10. R.F. Nalewajski, Phys. Chem. Chem. Phys. **4**, 1710 (2002); Chem. Phys. Lett. **372**, 28 (2003); J. Phys. Chem. A **107**, 3792 (2003)
11. R.G. Parr, P.W. Ayers, R.F. Nalewajski, J. Phys. Chem. A **109**, 3957 (2005)
12. R.F. Nalewajski, E. Broniatowska, Theor. Chem. Acc. **117**, 7 (2006)
13. R.F. Nalewajski, E. Świtka, A. Michalak, Int. J. Quantum Chem. **87**, 198 (2002); R.F. Nalewajski, E. Świtka, Phys. Chem. Chem. Phys. **4**, 4952 (2002); R.F. Nalewajski, Use of fisher information in quantum chemistry, Int. J. Quantum Chem. (K. Jankowski issue). (in press)
14. R.F. Nalewajski, E. Broniatowska, J. Phys. Chem. A **107**, 6270 (2003); Int. J. Quantum Chem. **101**, 349 (2005)
15. R.F. Nalewajski, A.M. Köster, S. Escalante, J. Phys. Chem. A **109**, 10038 (2005)
16. R.F. Nalewajski, E. Broniatowska, Chem. Phys. Lett. **376**, 33 (2003)
17. R.F. Nalewajski, J. Phys. Chem. A **107**, 3792 (2003); Ann. Phys. (Leipzig) **13**, 201 (2004); Mol. Phys. **104**, 255 (2006)
18. R.F. Nalewajski, J. Phys. Chem. A **104**, 11940 (2000)
19. R.F. Nalewajski, Mol. Phys. **102**, 531, 547 (2004)
20. R.F. Nalewajski, Mol. Phys. **103**, 451 (2005)
21. R.F. Nalewajski, Mol. Phys. **104**, 365, 493, 1977, 2533 (2006)

22. R.F. Nalewajski, Struct. Chem. **15**, 391 (2004)
23. R.F. Nalewajski, J. Math. Chem. **38**, 43 (2005)
24. R.F. Nalewajski, Theoret. Chem. Acc. **114**, 4 (2005)
25. R.F. Nalewajski, K. Jug, in *Reviews of Modern Quantum Chemistry: A Celebration of the Contributions of Robert G. Parr*, vol. I, ed. by K.D. Sen (World Scientific, Singapore, 2002), p. 148
26. R.F. Nalewajski, Chem. Phys. Lett. **386**, 265 (2004)
27. R.F. Nalewajski, Mol. Phys. **104**, 2533, 3339 (2006)
28. R.F. Nalewajski, J. Phys. Chem. A **111**, 4855 (2007)
29. R.F. Nalewajski, Entropic bond indices from molecular information channels in orbital resolution: ground-state systems. J. Math. Chem. (in press)
30. R.F. Nalewajski, Chemical bonds through probability scattering: information channels for intermediate-orbital stages. J. Math. Chem. (in press)
31. R.F. Nalewajski, Entropic bond-descriptors of molecular information systems in local resolution. J. Math. Chem. (in press)
32. K.A. Wiberg, Tetrahedron **24**, 1083 (1968)
33. M.S. Gopinathan, K. Jug, Theor. Chim. Acta (Berl.) **63**, 497, 511 (1983); see also: K. Jug, M.S. Gopinathan, in *Theoretical Models of Chemical Bonding*, vol. II, ed. by Z.B. Maksić (Springer, Heidelberg, 1990), p. 77
34. I. Mayer, Chem. Phys. Lett. **97**, 270 (1983)
35. R.F. Nalewajski, A.M. Köster, K. Jug, Theoret. Chim. Acta (Berl.) **85**, 463 (1993)
36. R.F. Nalewajski, J. Mrozek, Int. J. Quantum Chem. **51**, 187 (1994)
37. R.F. Nalewajski, S.J. Formosinho, A.J.C. Varandas, J. Mrozek, Int. J. Quantum Chem. **52**, 1153 (1994)
38. R.F. Nalewajski, J. Mrozek, G. Mazur, Can. J. Chem. **100**, 1121 (1996)
39. R.F. Nalewajski, J. Mrozek, A. Michalak, Int. J. Quantum Chem. **61**, 589 (1997)
40. J. Mrozek, R.F. Nalewajski, A. Michalak, Polish J. Chem. **72**, 1779 (1998)
41. W. Heitler, F. London, Z. Physik **44**, 455 (1927); for an English translation see: H. Hettema, *Quantum Chemistry Classic Scientific Paper* (World Scientific, Singapore, 2000); F. London, Z. Phys. **455**, 46 (1928)
42. S. Shaik, in *New Theoretical Concepts for Understanding Organic Reactions*, NATO ASI Series, Vol. C267, ed. by J. Bertran, I.G. Czismadia (Kluwer Academic Publ., Dordrecht, 1989), p. 165; S. Shaik, P.C. Hiberty, Adv. Quant. Chem. **26**, 100 (1995); K. Jug, A.M. Köster, J. Am. Chem. Soc. **112**, 6772 (1990); S. Shaik, A. Shurki, D. Danovich, P.C. Hiberty, Chem, Rev, **101**, 1501 (2001)