An effective scaling frequency factor method (ESFF): review and local factors transferability problem

Karolina Gdula, Mariusz Barczak and Piotr Borowski

Faculty of Chemistry, Maria Curie-Skłodowska University,
Maria Curie-Sklodowska sq. 3, 20-031 Lublin, Poland

Scaling procedures are known to reproduce very accurate vibrational spectra provided that multiparameter scaling in conjunction with high-quality force fields is carried out. In contrast to purely theoretical approaches (variational and perturbational), they are applicable to large systems. In this work, a brief review of the scaling procedures is given. The emphasis is put on the recently proposed effective scaling frequency factor (ESFF) method [*Chem. Phys. Lett.,* **446**, 191, (2007), *J. Mol. Spectr.,* **264**, 66, (2010)] – the multiparameter frequency scaling method providing better scaled frequencies than the well-established scaled quantum mechanical (SQM) force field approach. In addition, the results of our calculations on the benzene-based related systems, i.e., benzene and most of its methyl derivatives, are presented. The calculations concern the middle- and low-frequency range of the vibrational spectra, where strong mixing of the local vibrations often takes place. The factors transferability problem is discussed.

1. INTRODUCTION

Harmonic vibrational frequencies of a molecule, which approximate the observed fundamentals, can be obtained from the quadratic force fields with the aid of the so-called Wilson-Decius-Cross (WDC) method [1, 2] (also called the GF matrix, when internal rather than Cartesian force fields are used). They are too high as compared with fundamentals in an

overwhelming majority of cases [3]. There are two main factors responsible for such a behavior. First, a real molecular potential energy surface (PES) is anharmonic, in that, e.g., the molecule is known to dissociate when the bond is stretched beyond its equilibrium length, and for this reason parabolic approximation of the PES is valid only in the direct vicinity of the equilibrium geometry. This effect lowers down the energy gap between the vibrational energy levels, which no longer corresponds to the classical harmonic frequency. Second, force fields computed using quantum chemistry (QC) methods may exhibit variable quality, depending on the curvature of the PES predicted by an approximate theoretical level. For example, a given theoretical method may lead to a curvature of PES that deviates from the exact one, in spite of predicting essentially correct equilibrium geometry. Such is the case with the MP2 method [4], which is capable of providing accurate geometries, but fails with respect to the accurate determination of vibrational spectra. In addition, there is a strong dependence of the force field on the equilibrium geometry, at which the force constants are to be calculated. For example, bond lengths obtained by means of the Hartree-Fock (HF) method for most molecules are typically too short and therefore the corresponding quadratic force constants are overestimated [5]. This leads to excessively high calculated frequencies (even if they are compared with the available experimental harmonic frequencies).

The formal theoretical procedure to account for the anharmonicity exists in including the effect of higher-order terms in the Taylor expansion of PES (which depend on cubic and higher force constants) within a variational or perturbational framework. A good review of the up-to-date literature can be found in the references [6, 7]. Unfortunately, that methodology cannot be efficiently applied to large systems. Nowadays, the density functional theory (DFT) methods (see, e.g., ref. [8]) are frequently used to predict reasonable equilibrium geometries and quadratic force fields of covalently bonded systems, at least with some density functionals. The computational cost of the DFT calculations is comparable to that of HF, which means that they are feasible on molecules containing even up to a few hundred atoms, in particular when parallel implementations of the QC software are used. However, calculations of higher energy derivatives for large systems for the subsequent use in purely quantum theoretical formalisms are still not practical. In such cases, the so-called scaling methods are the methods of choice. They are very efficient; they are capable of providing very accurate final (scaled) frequencies, at least in the case of multiparameter

scaling, at basically no additional computational cost. In addition, they include (at least to some extent) the effects of incomplete incorporation of correlation effects in the calculations, basis set incompleteness, and errors in the predicted geometry. All scaling methods use the so-called scaling factors, which are obtained by minimizing the least-squares merit function between the scaled theoretical and the observed fundamental frequencies for a set of well-defined fundamentals of the molecules from the so-called training (calibration) set. Thus, they should be regarded as empirical methods.

The present paper deals with the tests of the recently proposed effective scaling frequency factor (ESFF) method [9]. The main idea behind any scaling procedure is to use transferable scaling factors, in that they can be applied to any molecule, once computed on the basis of the limited set of the experimental data. It is demonstrated that the ESFF local scaling factors (cf. Section 2 for the definition) are highly transferable. This conclusion is drawn from the results of our calculations on the related systems, in that they all exhibit similar structural motifs. They are: benzene, toluene, as well as all di-, tri-, and tetramethyl derivatives of benzene. The paper is organized as follows. In Section 2 we will briefly review the most common scaling procedures. ESFF, the procedure applied in the present paper, will be also presented there. Computational details of our QC and ESFF calculations will be stated in Section 3. This section is followed by presenting and discussing our results (Section 4), and, finally, the conclusions are reported in Section 5.

2. REVIEW OF SCALING PROCEDURES

Two alternative scaling approaches are known at present: harmonic frequency scaling and harmonic force field scaling. The former one is very straightforward and easy to apply. The latter provides frequencies that are consistent with the final force fields; additional software to carry out such calculations is required, though.

Single-parameter frequency scaling procedures

In the early 1980s, single parameter scaling of harmonic frequencies, also known as uniform scaling, was proposed by Pople and co-workers [10]. Initially, harmonic frequencies determined from the HF force fields – the only ab initio approach commonly applicable in computational chemistry at that time – were subjected to scaling. This procedure was

then extended to other theoretical levels with typical basis sets [11,12], and nowadays a wide database of scaling factors for the harmonic frequencies, zero-point vibrational corrections and vibrational contributions to thermodynamic functions is available in the literature [13]. It should be emphasized that the accuracy of the scaled frequencies obtained in this way is not impressive, in that the root-mean-square (RMS) deviation between the scaled and observed frequencies is often quite large in the entire spectral range (30–50 cm^{-1}, sometimes even more, cf. ref. [13]). In addition, uniform scaling is not capable of providing the necessary frequency swaps – the scaled frequencies exhibit the same order as the harmonic ones. Thus, the assignment of two close-lying bands of comparable intensity on the vibrational spectrum (FT-IR, Raman, etc.) may become ambiguous in the case of totally different nature of the harmonic modes potentially responsible for their presence. However, this approach is frequently used due to its simplicity. Better in this regard turned out to be the so-called wavenumber linear scaling (WLS) procedure [14–16]. It takes advantage of the linear relationship between the scaling factor, defined as the ratio $\omega^{(\text{expt})}/\omega^{(\text{harm})}$, and the calculated frequency. For example, for B3LYP/6-311+G** harmonic frequencies this relationship reads

$$\frac{\omega^{(\text{expt})}}{\omega^{(\text{harm})}} = 1.0087(9) - 0.0000169(6)\omega^{(\text{harm})} \tag{1}$$

Recently, another frequency scaling approach, called effective scaling frequency factor (ESFF) approach, was proposed [9]. It belongs to the group of multiparameter scaling procedures, but in the limiting case of using only one local scaling factor (*vide infra*) it is demoted to the uniform scaling. The present paper deals with the results obtained within the ESFF scaling frame; thus, the method will be given more attention later in this section.

SQM force-field scaling procedure

Alternatively, scaling can be applied to the quadratic force constants. Multiparameter force field scaling was initiated in the 1970s by Botschwina and co-workers [17–19], as well as Bloom and co-workers [20–25]. It was further modified by Pulay et al. [26], who essentially proposed a more appropriate treatment of the coupling (off-diagonal) force constants. With this modification, the method became the famous scaled quantum mechanical (SQM) force field method. It is based on the

intuitive dividing of the internal coordinates basis (non-redundant natural [26,27], or redundant primitive [28]) into chemically similar groups sharing the same scaling factor s_i; the quadratic force constant matrix $F_{ij}^{(harm)}$ expressed in terms of these internal coordinates, is scaled according to $F_{ij}^{(scal)}=(s_i s_j)^{1/2}F_{ij}^{(harm)}$ prior to solving the vibrational secular problem. SQM was initially applied to HF force fields [26]; much more accurate, DFT force fields were then subjected to scaling [28]. With the appropriate classification of internal coordinates the obtained results turned out to be of high quality – the RMS deviation between fundamental and scaled theoretical frequencies obtained using DFT/B3LYP harmonic force fields was low, close to 12 cm^{-1}. A number of commercial SQM programs for routine applications is now available [29–31]. Recently, an extended database of the SQM scaling factors for the routine applications to molecules built of the second-row elements and chlorine was reported in the literature [32].

ESFF method

An effective scaling frequency factor method [9] uses effective scaling factors (ESFs) which – for a given vibrational mode – are constructed from a set of local scaling factors (LSFs) as the sums of LSFs weighed by the contributions of the local modes to the normal mode. LSFs are attributed to the types of internal coordinates, and the contributions are determined from the diagonal potential energy distribution (PED) matrix elements [33] after solving the vibrational problem. Note that the above-mentioned idea of classification of internal coordinates into chemically similar groups (types) is retained in ESFF. The assumptions of the method read as follows [9]:

1. all molecules are built of limited, intuitively chosen set of types of internal coordinates. They can be chosen as the ordinary primitive or natural internal coordinates forming a non-redundant set, or as all redundant primitive coordinates generated on the basis of the atomic connectivities. Let the number of types of internal coordinates be denoted n; we assume that in all molecules the Ith type has the same local scaling factor f_I, $I=1,2,…,n$;
2. the contributions $P_{I,k}$ of each type of local mode to a given, say, kth normal mode can be determined as

$$P_{I,k} = \sum_{i \in \text{type } I} p_{ii,k} \tag{2}$$

In the above expression $p_{ii,k}$ are the ordinary PED coefficients [33]

$$p_{ii,k} = \frac{\alpha_{ik}^2 F_{ii}}{\sum_j \alpha_{jk}^2 F_{jj}} \tag{3}$$

obtained from the force constants $\mathbf{F}$ expressed in internal coordinates, as well as the amplitudes (the $\boldsymbol{\alpha}$ matrix) that follow from the solution of an ordinary WDC equation of the form

$$\mathbf{F}\boldsymbol{\alpha} = \mathbf{G}^{-1}\boldsymbol{\alpha}\omega^2 \tag{4}$$

3. an effective scaling factor $f_k^{(\text{eff})}$ for the frequency associated with the kth normal mode is obtained as

$$f_k^{(\text{eff})} = \sum_I P_{I,k} f_I \tag{5}$$

4. finally, harmonic frequencies are scaled according to the following equation

$$\omega_k^{(\text{scal})} = f_k^{(\text{eff})} \omega_k^{(\text{harm})} \tag{6}$$

Local scaling factors can be determined – as in all scaling procedures – with the aid of the least-squares method. Namely, for a given, well defined set of set of the experimental fundamentals, $\omega_k^{(\text{expt})}, k = 1,2,...,M$, where M denotes the number of the selected frequencies (typically a few hundred) belonging to molecules of some training set typically consisting of a few dozen of (usually small) molecules, we minimize the least-square merit function of the form

$$F(\mathbf{f}) = F(f_1, f_2,...,f_n) = \sum_{k=1}^{M} \left\{ \omega_k^{(\text{scal})} - \omega_k^{(\text{expt})} \right\}^2 \tag{7}$$

with respect to the local scaling factors f_I. This results in a system of linear equations which can be easily solved using standard numerical procedures. At this point we must mention that this constitutes an advantages of ESFF over the well-established SQM approach [34], as the latter one requires an iterative search for the scaling factors. More details on further comparison of the both methods can be found in references cited in the text, in particular in [32] and [34].

It should be noted that – as the pioneers of the scaling methodology claim (see, e.g., ref. [26]) – none of the scaling procedures has a strict theoretical basis. They are all empirical procedures, and therefore their

validity is judged based on the agreement of the calculations and experiment.

The performance of the ESFF method was then tested [35] using Baker's training set of 30 small molecules [28], the set designed for calculations on systems built of the second-row elements and chlorine. It turned out to be superior to SQM, in that the RMS deviation between scaled and fundamental frequencies was somewhat lower. It has to be mentioned that the comparison of both approaches, though reliable, was not complete at that time. The SQM calculations were carried out using flexible redundant primitive internal coordinates, while those within the ESFF frame – using non-redundant natural internal coordinates. However, in both cases the originally proposed classification of the internal coordinates [28] was adopted. Then, the method was extended. First, a new, reduced set of scaling factors, particularly well suited to the frequency scaling in the fingerprint range of the vibrational spectrum, was proposed [36]. Second, the calculations on large systems, i.e., 1,2,4-triazole derivatives, were carried out [37]. The ESFF approach turned out to perform well, again somewhat better than SQM, in spite of using less flexible natural internal coordinate basis. Third, ESFF was reformulated, and in the present version full advantage of expressing local contributions to the normal modes in terms of redundant primitive valence coordinates is taken [38]. Fourth, the applicability of ESFF (as well as SQM) was extended to silicon and sulfur-II containing molecules and a number of typical basis sets in conjunction with the B3LYP density functional [34]. It should be mentioned that until 2012 both ESFF and SQM scaling, though widely used, were applied to the frequencies/force fields obtained at a very limited number of QC approaches (defined as METHOD/BASIS SET). Fifth, an extended database of ESFF local scaling factors (as well as SQM force field factors, as mentioned earlier) was created [32]. The scaling factors were obtained for a great variety of density functionals in conjunction with a great variety of typical basis sets. The ESFF method was shown to be superior to SQM in overwhelming majority of cases. As before the calculations were carried out using Baker's training set, which was shown to be consistent with that used in the works on the development of uniform scaling [13]. The recommendations related to the calculations oriented toward determination of the scaled frequencies were also given. Finally, the ESFF method was successfully applied do the determination of the IR spectra of the amorphous polymers [39].

As already mentioned, in contrast to uniform scaling, ESFF is a multi-parameter frequency scaling procedure. An obvious advantage of

using such a methodology is its higher flexibility – the scaled frequencies often exhibit different order as compared with harmonic ones, and the overall RMS value is also significantly lower. The former aspect is of particular importance when two (or more) close-lying bands are to be assigned, and their origin cannot be determined on the basis of other considerations, for example related to the molecular symmetry. In addition, ESFF offers a number of advantages as compared with the well-known and well-established SQM scaling procedure. They can be summarized as follows.

1. It is efficient; statistically it is superior even to the SQM method. The local scaling factors are well transferable between the molecules, somewhat better than in the case of SQM. In fact, it was shown a number of times in the reported literature [9, 32, 34–38].
2. It is simple and straightforward; some of the applications can be made "by hand", i.e., without the necessity of purchasing the additional software [37]. Moreover, the factors optimization can be performed by using typical statistical programs (e.g., Excel), though the relevant, commercial programs would surely facilitate the computational procedure.
3. Optimization of the scaling factors within the ESFF procedure is numerically stable, in contrast to SQM, which sometimes fails (which was already reported in the literature; cf. ref. [34]). Moreover, the iterative search for the least-squares minimum is not needed.
4. The ESFF method is well suited to the description of the macromolecular compounds [39]. When attempting to describe really large systems in the restricted frequency range at the levels of the mode-tracking [40, 41] and/or intensity-tracking [42, 43] procedures, utilization of the SQM scaling is not possible, on account of the determination of only the frequencies and normal modes (eigenvectors), rather than the molecular force fields. The ESFF scaling can be applied, though. Admittedly, the force constants enter the expression for the PED coefficients (cf. eq. 3); however, this problem is easily eliminated with the aid of the so-called ESFF2 procedure [38], according to which the contributions of the local vibrations to the normal modes can be determined only on the basis of the amplitudes of the atomic displacements from the equilibrium positions. This procedure turned out to be as efficient as the original one.

3. COMPUTATIONAL PROCEDURE

In the case of routine, analytical applications, the force field/frequency scaling procedures should be preferably based on transferable scaling factor(s) determined (optimized) using well-defined training set of molecules. Applications, in which factors are optimized for a limited set of new molecules may also be useful. Such is the case with an unusual behavior of some vibrations – the reoptimized values can provide better insight into the real molecular force field. This is also true when the transferability of scaling factors (as reported in this work) is to be investigated.

The calculations were carried out for benzene (**1**), toluene (**2**), *o*- (**3**), *m*- (**4**), *p*-xylene (**5**), 1,2,3- (**6**), 1,2,4- (**7**), 1,3,5-trimethylbenzene (**8**), 1,2,3,4- (**9**), 1,2,3,5- (**10**), and 1,2,4,5-tetramethylbenzene (**11**). The Cartesian force fields of these systems were calculated at DFT level of theory [8]. The B3LYP density functional [44], which was shown to provide the best vibrational spectra [32], was used. Utilization of the 6-311G** basis set [45] seems to be a good compromise between the results quality and computational effort for these compounds, though – to demonstrate the necessity of incorporating the polarization functions to the basis set – some results obtained with the 6-31G [46] basis set are also reported. Obviously, the geometry optimization was carried out for all molecules prior to the hessian calculations. All calculations were performed using the PQS QC package [47, 48].

The experimental vibrational frequencies were obtained from the FT-IR spectra included in the well-known Spectral Database of Organic Compounds (SDBS) [49]. We decided to concentrate on the frequencies below 1650 cm^{-1}, i.e., stretching vibrations involving hydrogen atoms were not considered. The reason is that in the high-frequency range large deviations between the scaled and the experimental frequencies are often observed, inter alia, on account of the frequency shifts due to the Fermi resonance, which cannot be handled by scaling.

The harmonic and ESFF-scaled frequencies were calculated using our home-made program. The program performs a few computational steps, which are executed in the loop over considered molecules, namely:

- generation of the so-called **B** matrix [1,2] and its inverse **B**$^{-1}$ (more precisely – the part corresponding to the selected internal coordinates) based on the Cartesian coordinates of the atoms and

the predefined internal coordinates (input data), followed by the transformation of the Cartesian force constant matrix $\mathbf{f}^{(x)}$ to the internal coordinate representation (the $\mathbf{F}$ matrix, cf. eq. 4),

– generation of the mass tensor $\mathbf{G}^{-1}$ based on the calculated $\mathbf{B}^{-1}$ matrix and atomic masses,

– solution the set of WDC equations (4) with the aid of the Jacobi diagonalization procedure [50] in conjunction with the canonical orthogonalization,

– determination of the PED coefficients, based on the diagonal elements of the $\mathbf{F}$ and the $\boldsymbol{\alpha}$ matrices (cf. eq. 3),

– optimization of the local scaling factors (if needed) and determination of the scaled frequencies.

The ESFF calculations were carried out using non-redundant internal coordinates. The main goal of the computational part of this paper is to demonstrate the transferability of LSFs; for this reason we did not follow any of the recommended classifications of internal coordinates [28, 36]. We used the types designed specifically for the present molecules instead. Thus, for benzene ring we used the following types: (i) C_{Ar}–C_{Ar} bond lengths (the so-called ν_{Ar} type), (ii) C_{Ar}–C_{Ar}–C_{Ar} valence angles (δ_{Ar} type), (iii) C_{Ar}–C_{Ar}–C_{Ar}–C_{Ar} torsion angles (τ_{Ar} type), (iv) C_{Ar}–H bond lengths (ν_{ArH} type), (v) C_{Ar}–C_{Ar}–H valence angles (ρ_{ArH} type), (vi) angles determining the bending vibrations of the hydrogen atoms, coming out of the benzene plane (γ_{ArH} type). Substitution of the ring with the methyl group introduces additional six types: (vii) C_{Ar}–C_{CH3} bond lengths (the so-called ν_{ArC} type), (viii) C_{Ar}–C_{Ar}–C_{CH3} valence angles (ρ_{ArC} type), (ix) angles determining the methyl carbon atoms bending vibrations out of the benzene ring plane (γ_{ArC} type), (x) C_{Met}–H bond lengths (ν_{CH} type), (xi) C_{Ar}–C_{Met}–H valence angles (δ_{Met} type), (xii) C_{Ar}–C_{Ar}–C_{Met}–H torsion angles (τ_{Met} type). As already pointed out, the middle- and low-frequency ranges in the vibrational spectrum, where frequently strong mixing of all but ν_{ArH} and ν_{CH} local modes takes place, were considered. For this reason the factors for the ν_{ArH} and ν_{CH} types were preset to unity. With this constraint a set of 10 scaling factors was obtained.

4. RESULTS AND DISCUSSION

The following sub-ranges below 1650 cm^{-1} can be distinguished on the FT-IR, or, in general, on the vibrational spectra of compounds **1–11** considered in this work:

- 1650–1450 cm^{-1}, where the stretching of aromatic ring ν_{Ar} coupled to variable extent to rocking of aromatic protons ρ_{ArH} takes place,
- 1450–1350 cm^{-1}, where in-plane bending vibrations of the methyl groups δ_{Met} take place,
- 1350–900 cm^{-1}, where delocalized and skeletal vibrations take place (this is, in fact, the main part of the so-called fingerprint range), and
- 900–650 cm^{-1}, where – apart from the delocalized "fingerprint vibrations" – strong bands corresponding to out-of-plane γ_{ArH} and torsion τ_{Ar} vibrations and/or a superposition of those take place (this is the "most structural" range in the entire FT-IR spectrum, in that it enables to identify the substitution of the aromatic ring without the necessity of comparing the fingerprint range with the extended database).

Table 1. Number of vibrational bands below 1650 cm^{-1} for each molecule considered in the present work. The sums for the molecules **1–5**, **1–8**, and **1–11**, respectively, are also reported.

No	Molecule	No. of bands	Sum
1	benzene	4	
2	toluene	12	
3	o-xylene	14	
4	m-xylene	14	
5	p-xylene	7	51
6	1,2,3-trimethylbenzene	13	
7	1,2,4-trimethylbenzene	13	
8	1,3,5-trimethylbenzene	10	87
9	1,2,3,4-tetramethylbenzene	14	
10	1,2,3,5-tetramethylbenzene	15	
11	1,2,4,5-tetramethylbenzene	10	126

As many as 126 well-resolved bands were selected from the FT-IR spectra of our compounds; bands corresponding to superposition of a few vibrational modes were rejected. Table 1 shows the number of bands per molecule included in the calculations. In the following part we will first concentrate on the above-mentioned "structural range" (900–650 cm^{-1}) to demonstrate the necessity of frequency scaling in routine applications, then we will deal with the factors transferability problem.

The out-of-plane vibrations

Table 2 reports the experimental and theoretical harmonic frequencies of the out-of-plane vibrations in the range of 900–650 cm^{-1}, as well as their computed and observed intensities. Note that the calculated values refer to B3LYP density functional and two basis sets: the small one, i.e., of VDZ quality (6-31G), and somewhat extended one of VTZ plus polarization quality (6-311G**, cf. Section 3).

As can be seen, the number of the intense bands in the above mentioned range is well reproduced even with a small basis set. The same applies to the relative intensities when two or more frequencies are present. However, the calculated harmonic frequencies exhibit significant error as compared with the fundamentals of the order of 30–45 cm–1. Substantial part of that error obviously follows from the anharmonic nature of the PES, but the results additionally suffer from the severe basis set deficiency. Indeed, increasing the basis set quality to 6-311G** (which for such systems treated with DFT should be sufficient) reduces the error by about 20 cm–1. It should be noted that the error, which mostly reflects the anharmonicity effects this time, is fairly constant and in most cases oscillates about the value of 15 cm–1. This constant error is to be expected due to the symmetric nature of the molecular potential associated with this kind of deformation. In the case of vibrations other than out-of-plane bending (e.g. bond stretching), the discrepancy is higher. Such disagreement clearly demonstrates the necessity of further improvements of the calculated frequencies. Empirical scaling is well suited for this kind of problems which we will demonstrate in the next subsection.

In addition, we observed the following trend with respect to the calculated vs. observed intensities. Namely (i) all bands classified as vs (very strong) correspond to vibrations having the calculated intensity A_∞ higher than 30 km mol^{-1}, (ii) s (strong) – $A_\infty \in$ (30–10 km mol^{-1}), (iii) m (medium) – $A_\infty \in$ (10–3 km mol^{-1}), (iv) w (weak) – $A_\infty \in$ (3–1 km mol^{-1}),

and (v) vw (very weak) – $A_\infty < 1$ km mol^{-1}. Those considerations refer to the 6-311G** basis set.

Table 2. The calculated harmonic, $\omega^{(harm)}$, and fundamental, $\omega^{(obs)}$, frequencies (in cm^{-1}), their differences, $\Delta\omega$, as well as calculated (A_∞, in km mol^{-1}) and observed (I, qualitative description) intensities for the out-of-plane vibrations of molecules **1–11**. The theoretical data refer to 6-31G / 6-311G** basis sets, respectively.

Molecule	$\omega^{(harm)}$	A_∞	$\omega^{(obs)}$	I	$\Delta\omega$
benzene	707 / 689	103 / 111	674	vs	+33 / +15
benzene monomethyl derivatives					
methylbenzene[a]	728 / 713	19.8 / 27.0	696	s	+32 / +17
	764 / 745	47.4 / 45.0	729	vs	+35 / +16
benzene dimethyl derivatives					
1,2-dimethylbenzene[a]	776 / 755	51.9 / 55.4	742	vs	+34 / +13
1,3-dimethylbenzene[a]	724 / 711	10.2 / 14.9	691	s	+33 / +20
	805 / 783	33.8 / 33.3	769	vs	+36 / +14
	918 / 893	1.91 / 2.33	876	w	+42 / +17
1,4-dimethylbenzene[a]	830 / 811	35.0 / 35.8	795	vs	+35 / +16
benzene trimethyl derivatives					
1,2,3-trimethylbenzene[a]	743 / 731	3.90 / 6.50	709	m	+34 / +22
	802 / 780	33.1 / 33.6	767	vs	+35 / +13
1,2,4-trimethylbenzene[a]	738 / 726	0.47 / 0.93	703	vw	+35 / +23
	843 / 821	23.2 / 23.7	805	vs	+38 / +16
	916 / 892	2.10 / 2.38	874	w	+42 / +18
1,3,5-trimethylbenzene[a]	720 / 709	7.23 / 12.5	688	s	+32 / +21
	880 / 855	20.7 / 20.6	836	s	+44 / +19
benzene tetramethyl derivatives					
1,2,3,4-tetrametlylbenzene[a]	839 / 816	19.3 / 20.3	802	s	+37 / +14
1,2,3,5-tetrametlylbenzene[a]	738 / 728	1.69 / 3.33	706	m	+32 / +22
	891 / 866	13.8 / 13.9	848	s	+43 / +18
1,2,4,5-tetrametlylbenzene[b]	900 / 879	12.6 / 12.7	868	s	+31 / +10

[a]Experimental spectrum – liquid film, [b]experimental spectrum – KBr pellet

ESFF scaling

The values of the scaling factors and therefore the quality of the final results depend on the fundamentals selected for the optimization procedure. However, in the limiting case of a large number of frequencies obtained from the vibrational spectra of a large number of various molecules they should asymptotically approach a constant value, and their statistical uncertainties should lower down. Four sets of scaling factors, denoted A, B, C, and D, respectively, were obtained. In the case of set A, the optimization of factors was based on 51 well resolved bands below 1650 cm^{-1} selected from the spectra of **1–5** (i.e., the tri- and tetramethyl derivatives were not considered). The removal of bands due to the delocalized skeletal vibrations, i.e., bands that appear in the range of 1350–900 cm^{-1}, gives set B. This means that B scaling factors were essentially determined on the basis of the group vibrations. Thus, the quality of the scaled frequencies reflects the transferability of scaling factors within the molecule (this term was introduced earlier [51] and should be understood as follows: the local scaling factors are transferable within the molecule if, after calculating them from rather localized modes, i.e., group vibrations, they are capable of reproducing frequencies of the delocalized modes, e.g., those forming the fingerprint range of a spectrum). In addition, the quality of the frequencies obtained for molecules **6–11** provides information on the transferability of the scaling factors between related molecules. When generating the set C we selected 87 bands from the spectra of **1–8** (this time only tetramethyl derivatives were not considered). Thus, it is equivalent to the set A, but from the statistical point of view it is more reliable, as there were more fundamentals the optimization of scaling factors was based on. Set D, like B, was obtained from the reduced set of frequencies obtained by excluding the skeletal vibrations. Note that the terms "set A", "set B" etc. may refer to both sets of factors, as well as sets of experimental frequencies (bands) their optimization was based on.

Table 3 reports the values of optimized scaling factors. The types of internal coordinates were described in Section 3. As can be seen they are all close to unity (as they should; this is the essence of all scaling procedures as – from the perturbational point of view – the smaller corrections, the more reliable the final quantity). In addition, most of them have similar values regardless of the set of frequencies used in the optimization.

Table 3. Optimized local scaling factors for 10 types of internal coordinates obtained from the A, B, C, and D sets of selected frequencies, respectively (see text).

Type	Scaling A	Scaling B	Scaling C	Scaling D
ν_{Ar}	0.9772	0.9833	0.9753	0.9827
δ_{Ar}	0.9796	0.9691	0.9794	0.9741
τ_{Ar}	0.9725	0.9674	0.9630	0.9728
ρ_{ArH}	0.9779	0.9792	0.9786	0.9776
γ_{ArH}	0.9835	0.9825	0.9849	0.9842
ν_{ArC}	0.9915	1.0416	0.9987	1.0159
ρ_{ArC}	0.9594	0.9568	0.9781	0.9524
γ_{ArC}	0.9416	1.0099	0.9866	0.9581
δ_{Met}	0.9732	0.9664	0.9723	0.9691
τ_{Met}	0.9794	0.9723	0.9801	0.9726

Only three factors exhibit fluctuations larger than 0.01; they are ν_{ArC}, ρ_{ArC}, and γ_{ArC}. This is not a surprise – there are basically no pure vibrations in the overall set of the selected bands involving those kinds of the local motions. They participate mostly in highly delocalized modes and for this reason their values are strongly dependent of the set of bands used in their determination. Nowadays, the common practice is to include them into other, better defined types. However, the main goal of these calculations is to discuss the factors transferability problem rather than to obtain factors for the routine calculations.

Table 4 includes the RMS values between the scaled and fundamental frequencies obtained with the aid of various sets of scaling factors (as described above). It should be recalled that sets A and B were generated on the basis of fundamentals of molecules **1–5**, while C and D – those of molecules **1–8**. The following remarks can be made.

Table 4. RMS values between the scaled and fundamental frequencies obtained for various sets of scaling factors considered in the present paper for molecules **1–5**, **1–8**, and **1–11**, respectively (see text), as well as the number of frequencies those values ere obtained from.

Molecules	**1–5**	**1–8**	**1–11**
Scaling A	4.1	4.6	5.3
Scaling B	5.9	6.2	6.7
Scaling C		4.5	5.0
Scaling D		5.1	5.6
Number of frequencies	51	87	126

1. Extension of the frequency set to be scaled by adding the frequencies of molecules **6–11** for A and B type of scaling and **9–11** for C and D (i.e., those not included in the factors optimization procedure) slightly deteriorates the obtained results. This is quite natural as the factors transferability between molecules, though very good, is not expected to be perfect. However, an increase in the overall RMS value is not high, and ranges between $0.5–1.2 \text{ cm}^{-1}$. This shows that the scaling factors are well transferable between the related systems considered in this work. This is in accord with the conclusions from the up to date literature.

2. Removal of the frequencies from the fingerprint range in the optimization procedure (set A vs. B and set C vs. D) also slightly deteriorates the RMS values obtained for all frequencies. However, an increase in the RMS value not larger than 1.8 cm^{-1} is observed. Analysis of individual frequencies shows that this is mostly due to a slight increase in the deviations from the experimental values for the removed frequencies (for those included – to a lesser extent). Only in two cases the RMS values exceed 6 cm^{-1} (which is just 1.5 times the resolution of typical analytical measurement). This proves good transferability of the local scaling factors within a molecule. Note that the capability of the scaling procedures to properly reproduce the frequencies of the delocalized modes is an important issue.

Nearly all individual scaled frequencies for the molecules considered in the present work show very good agreement with experiment. This is shown in Table 5, which is split into 4 parts (denoted 5a–5d) concerning

benzene and toluene (Table 5a), all xylenes (Table 5b), three- (Table 5c), and tetramethyl (Table 5d) derivatives of benzene, respectively. The reported values refer to set C of scaling factors – the most reliable set determined in the present work. In addition, the RMS values for each molecule are also provided. As can be observed, they do not exceed 7 cm^{-1}, which is an excellent result. Remember that in the case of uniform scaling RMS is frequently larger than 30 cm^{-1} [13] (sometimes larger than 50 cm^{-1} for poor-quality force fields) in spite of optimizing the scaling factor.

Table 5a. Benzene and toluene. The calculated harmonic, $\omega^{(harm)}$, fundamental, $\omega^{(obs)}$, and scaled, $\omega^{(scal)}$ (using the set C of factors, see text) frequencies (in cm^{-1}), as well as the differences between the scaled and fundamental frequencies ($\Delta\omega$) are given. The RMS values are also reported.

Benzene				Toluene			
$\omega^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$	$\omega^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$
1513	1479	1479	0	1648	1605	1612	7
1060	1038	1037	−1	1529	1496	1497	1
1013	996[a]	988	−8	1415	1379	1380	1
689	674	678	4	1229	1210	1207	−3
				1204	1179	1177	−2
				1182	1156	1155	−1
				1112	1082	1087	5
				1062	1030	1037	7
				910	895	895	0
				799	786	783	−3
				745	729	729	0
				713	696	693	−3
		RMS	**4.5**			**RMS**	**3.6**

[a]The so-called ring breathing vibration; frequency obtained from Raman spectrum.

Table 5b. Dimethyl derivatives of benzene (see Table 5a for details).

o-xylene				m-xylene				p-xylene			
$\omega^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$	$\omega^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$	(harm)	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$
1649	1606	1613	7	1649	1614	1614	0	1549	1516	1517	1
1524	1495	1491	-4	1629	1592	1594	2	1414	1378	1380	2
1505	1467	1471	4	1526	1492	1493	1	1237	1220	1214	-6
1497	1456	1465	9	1507	1482	1474	-8	1142	1120	1116	-4
1422	1384	1386	2	1491	1462	1459	-3	1063	1043	1038	-5
1412	1374	1377	3	1412	1376	1377	1	1038	1023	1017	-6
1317	1291	1289	-2	1272	1250	1248	-2	811	795	796	1
1243	1223	1224	1	1196	1170	1170	0				
1144	1120	1120	0	1122	1095	1097	2				
1074	1053	1050	-3	1062	1040	1035	-5				
1044	1022	1019	-3	920	905	903	-2				
1003	986	978	-8	893	876	877	1				
942	932	926	-6	783	769	769	0				
755	742	743	1	711	691	691	0				
		RMS 4.6				**RMS 2.9**				**RMS 4.1**	

5. CONCLUSIONS AND FINAL REMARKS

The present work briefly reviews the advances in the development of empirical scaling – a methodology supporting interpretation of the vibrational spectra. The main attention is given to the recently proposed ESFF harmonic frequency scaling method – the method developed at the authors' University. The method was described, and all the applications reported to date in the literature were summarized. Then, ESFF was applied to reproduce the FT-IR spectra of benzene-based molecules exhibiting similar structural motifs. In fact, we used benzene itself and its various methyl derivatives. The calculations clearly indicate the necessity of correcting frequencies calculated within the harmonic approximation. Harmonic frequencies of the out-of-plane vibrations – vibrations, for which the potential is symmetric (like parabola), i.e., qualitatively resembling the harmonic one – deviate significantly from the observed fundamentals in spite of using reliable force fields (B3LYP/6-311G**). The calculated

scaling factors were proven to be well transferable, both between the related molecules and within the molecule. The obtained frequencies turned out to be in good agreement with experiment. The final, scaled frequencies show small deviations as compared with the observed fundamentals. It directly translates into the low value of the root-mean-square deviation, which in most cases is lower than 1.5 times the resolution of the typical, analytical measurement.

Table 5c. Trimethyl derivatives of benzene (see Table 5a for details).

1,2,3-trimethylbenzene				1,2,4-trimethylbenzene				1,3,5-trimethylbenzene			
$\omega^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$	$\omega^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$	$^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$
1627	1588	1592	4	1657	1619	1621	2	1643	1609	1607	-2
1493	1443	1462	19	1613	1576	1578	2	1514	1472	1480	8
1426	1386	1390	4	1536	1505	1504	-1	1491	1443	1459	16
1412	1378	1378	0	1422	1383	1387	4	1413	1375	1377	2
1270	1249	1248	-1	1232	1211	1208	-3	1187	1165	1164	-1
1190	1163	1165	2	1181	1156	1157	1	1060	1037	1035	-2
1118	1094	1092	-2	1145	1125	1123	-2	945	929	927	-2
1092	1075	1070	-5	1061	1038	1035	-3	895	881	880	-1
1033	1010	1008	-2	1044	1021	1018	-3	855	836	837	1
1010	988	985	-3	1015	999	992	-7	709	688	689	1
892	879	874	-5	892	874	876	2				
780	767	766	-1	821	805	807	2				
731	709	712	3	726	703	709	6				
	RMS		**6.0**		**RMS**		**3.4**		**RMS**		**5.8**

The main shortcoming of the ESFF procedure, like any other frequency scaling ones, is the lack of consistency between the calculated frequencies and the final molecular force field. In the case of SQM, the force constants and the final frequencies are clearly consistent. This is advantageous in the case when further analysis on the force constants is required. However, the lack of the above-mentioned consistency does not seem to constitute a severe problem.

Table 5d. Tetramethyl derivatives of benzene (see Table 5a for details).

1,2,3,4-tetramethylbenzene				1,2,3,5-tetramethylbenzene				1,2,4,5-tetramethylbenzene			
$\omega^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$	$\omega^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$	$^{(harm)}$	$\omega^{(obs)}$	$\omega^{(scal)}$	$\Delta\omega$
1639	1604	1603	-1	1651	1615	1615	0	1537	1503	1505	2
1524	1489	1492	3	1617	1579	1582	3	1498	1462	1466	4
1518	1477	1484	7	1524	1509	1491	-18	1422	1388	1387	-1
1500	1464	1468	4	1512	1486	1479	-7	1405	1375	1371	-4
1489	1444	1457	13	1496	1461	1464	3	1222	1205	1200	-5
1425	1381	1389	8	1492	1454	1460	6	1214	1196	1190	-6
1272	1262	1252	-10	1481	1446	1450	4	1046	1026	1021	-5
1278	1254	1253	-1	1425	1384	1389	5	1018	1001	993	-8
1192	1166	1167	1	1416	1377	1381	4	1001	984	978	-6
1101	1080	1077	-3	1229	1213	1206	-7	879	868	863	-5
1079	1063	1058	-5	1161	1143	1139	-4				
1064	1039	1038	-1	1092	1077	1070	-7				
1030	1005	1005	0	1060	1024	1034	10				
816	802	803	1	866	848	849	1				
				728	706	710	4				
	RMS		**5.6**		**RMS**		**6.9**		**RMS**		**5.0**

Obviously, the $\omega \rightarrow \mathbf{F},\alpha$ transformation is not defined, as opposed to the inverse one, i.e., $\mathbf{F} \rightarrow \omega,\alpha$; the force fields corresponding to the scaled frequencies are, therefore, not uniquely defined. Nevertheless, the problem may be circumvented in the following way. Namely, after solving the WDC set of equations (4) and finding the ESFF-scaled frequencies, $\omega^{(scal)}$, the corresponding, scaled force fields can be found according to the equation

$$\mathbf{F}^{(scal)} = \mathbf{G}^{-1}\alpha(\omega^{(scal)})^2\alpha^{-1} \qquad (8)$$

using the obtained eigenvectors α; then the required analysis can be carried out.

An initial test of the described procedure was recently carried out using the formaldehyde molecule. Namely, the scaled force constants obtained according to eq. (8) were compared with those determined using the SQM procedure. Then, the relative deviations defined as $|F^{(scal)SQM} - F^{(scal)ESFF}|/F^{(scal)SQM}$ were found. The largest deviation for the diagonal elements of the scaled force constants (observed for $F_{C=O}$) was only 0.6%. Somewhat larger deviations, not exceeding 1% in most cases though, were observed for the off-diagonal elements. Such good agreement with the SQM force fields allows us for the time being to think about the comparable capabilities of the ESFF method in this respect. However, the $\mathbf{F}^{(scal)}$ matrix determined from equation (8), by definition corresponds to the same normal modes (the $\boldsymbol{\alpha}$ matrix) as harmonic frequencies. Scaling of the force fields according to the SQM procedure leads to new, potentially better eigenvectors, which can be used for calculating the improved intensities. In this case our experience clearly reveals that SQM is not really so profitable in this respect. The intensities determined from the SQM-scaled force fields frequently only marginally change as compared with those computed for the harmonic force fields. The differences do not exceed a few percent in an overwhelming majority of cases (1% in more than half of the cases for the molecules in Baker's training set), and our observations do not confirm a noticeable advantage of the corrected eigenvectors as compared with the harmonic ones.

REFERENCES

[1] E.B. Wilson, Jr., J.C. Decius, P.C. Cross, Molecular Vibrations. *The Theory of Infrared and Raman Vibrational Spectra*, Dover Publications, Inc., New York, (1955).

[2] S. Califano, *Vibrational States*, John Wiley & Sons, Ltd., London, (1976).

[3] W.J. Hehre, L. Radom, P.v.R. Schleyer, J.A. Pople, *Ab initio Molecular Orbital Theory*, Wiley, New York, (1986); and references therein.

[4] T. Helgaker, T.A. Ruden, P. Jørgensen, J. Olsen, W. Klopper, *J. Phys. Org. Chem.* **17**, 913, (2004); and references therein.

[5] G. Fogarasi, P. Pulay, *Ab initio Calculation of Force Fields and Vibrational Spectra*, [in:] *Vibrational Spectra and Structure*, (J.R. Durig, Ed.), Elsevier, Amsterdam, Vol. **14**, pp. 125–219, (1985).

[6] J.M. Bowman, T. Carrington, H.-D. Meyer, *Mol. Phys.* **106**, 2145, (2008).

[7] S.V. Krasnoshchekov, E.V. Isayeva, N.F. Stepanov, *J. Phys. Chem. A* **116**, 3691, (2012).

[8] R.G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, (1989).

[9] P. Borowski, M. Fernández-Gómez, M.P. Fernández-Liencres, T. Peña Ruiz, *Chem. Phys. Lett.,* **446**, 191, (2007).

[10] J.A. Pople, H.B. Schlegel, R. Krishnan, D.J. DeFrees, J.S. Binkley, M.J. Frisch, R.A. Whiteside, R.F. Hout, W.J. Hehre, *Int. J. Quantum Chem.,* **20**, 269, (1981).

[11] A.P. Scott, L. Radom, *J. Phys. Chem.* **100**, 16502, (1996).

[12] P. Sinha, S.E. Boesch, C. Gu, R.A. Wheeler, A.K. Wilson, *J. Phys. Chem. A,* **108**, 9213, (2004).

[13] J.P. Merrick, D. Moran, L. Radom, *J. Phys. Chem. A,* **111**, 11683, (2007).

[14] H. Yoshida, A. Ehara, H. Matsuura, *Chem. Phys. Lett.,* **325**, 477, (2000).

[15] H. Yoshida, K. Takeda, J. Okamura, A. Ehara, H. Matsuura, *J. Phys. Chem. A,* **106**, 3580, (2002).

[16] K.V. Berezin, V.V. Nechaev, T.V. Krivokhizhina, *Optics and Spectroscopy,* **94**, 357, (2003) (translated from *Optika i Spektro-skopiya* **94**, 398, (2003)).

[17] P. Botschwina, *Chem. Phys. Lett.,* **29** 98, (1974) .

[18] W. Bleicher, P. Botschwina, *Mol. Phys.,* **30**, 1029, (1975) .

[19] P. Botschwina, W. Meyer, A.M. Semkow, *Chem. Phys.,* **15**, 25, (1976).

[20] C.E. Blom, P. J. Slingerland, C. Altona, *Mol. Phys.,* **31**, 1359, (1976).

[21] C.E. Blom, C. Altona, *Mol. Phys.,* **31**, 1377, (1976).

[22] C.E. Blom, L.P. Otto, C. Altona, *Mol. Phys.,* **32**, 1137, (1976).

[23] C.E. Blom, C. Altona, *Mol. Phys.,* **33**, 875, (1977).

[24] C.E. Blom, C. Altona, *Mol. Phys.,* **34**, 177, (1977).

[25] C.E. Blom, C. Altona, A. Oskam, *Mol. Phys.,* **34**, 557, (1977).

[26] P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, A. Vargha, *J. Am. Chem. Soc.,* **105**, 7037, (1983).

[27] G. Rauhut, P. Pulay, *J. Phys. Chem.,* **99**, 3093, (1995).

[28] J. Baker, A.A. Jarzecki, P. Pulay, *J. Phys. Chem. A,* **102** 1412, (1998).

[29] L. Hedberg, I.M. Mills, *J. Mol. Spectrosc.,* **203**, 82, (2000).

[30] T. Sundius, *Vib. Spect.,* **29** 89, (2002).
[31] SQM version 1.0, *Scaled Quantum Mechanical Force Field*, 2013 Green Acres Road, Fayetteville, Arkansas 72703.
[32] P. Borowski, *J. Phys. Chem. A,* **116**, 3866, (2012) .
[33] Y. Morino, K. Kuchitsu, *J. Chem. Phys.,* **20**, 1809, (1952).
[34] P. Borowski, T. Peña Ruiz, M. Barczak, K. Pilorz, S. Pasieczna-Patkowska, *Spect. Acta Part A,* **86**, 571, (2012).
[35] P. Borowski, A. Drzewiecka, M. Fernández-Gómez, M.P. Fernández-Liencres, T. Peña Ruiz, *Chem. Phys. Lett.,* **465**, 290, (2008).
[36] P. Borowski, A. Drzewiecka, M. Fernández-Gómez, M.P. Fernández-Liencres, T. Peña Ruiz, *Vib. Spect.,* **52**, 16, (2010).
[37] P. Borowski, K. Pilorz, M. Pitucha, *Spect. Acta Part A,* **75**, 1470, (2010).
[38] P. Borowski, *J. Mol. Spectr.,* **264**, 66, (2010).
[39] P. Borowski, S. Pasieczna-Patkowska, M. Barczak, K. Pilorz, *J. Phys. Chem. A*, **116**, 7424, (2012).
[40] M. Reiher, J. Neugebauer, *J. Chem. Phys.* **118**, 1634, (2003).
[41] C. Herrmann, J. Neugebauer, M. Reiher, *New J. Chem.* **31**, 818, (2007).
[42] S. Luber, J. Neugebauer, M. Reiher, *J. Chem. Phys.,* **130,** 064105, (2009).
[43] K. Kiewisch, S. Luber, J. Neugebauer, M. Reiher, *Chimia,* **63**, 270, (2009).
[44] A.D. Becke, *J. Chem. Phys.,* **98**, 5648, (1993).
[45] R. Krishnan, J.S. Binkley, R. Seeger, J.A. Pople, *J. Chem. Phys.,* **72**, 650, (1980).
[46] P.C. Hariharan, J.A. Pople, *Theor. Chim. Acta,* **28**, 213, (1973).
[47] J. Baker, K. Woliński, M. Malagoli, D. Kinghorn, P. Woliński, G. Magyarfalvi, S. Saebo, T. Janowski, P. Pulay, *J. Comput. Chem.,* **3**0, 317, (2009).
[48] PQS version 3.3, Parallel Quantum Solutions, 2013 Green Acres Road, Fayetteville, Arkansas 72703.
[49] http://sdbs.riodb.aist.go.jp/sdbs/cgi-bin/cre_index.cgi?lang=eng
[50] W.H. Press, S.A. Teulolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recepies in Fortran*, Cambridge University Press, Cambridge, (1992).
[51] P. Borowski, M. Fernández-Gómez, M.P. Fernández-Liencres, T. Peña Ruiz, M. Que-sada Rincón, *J. Mol. Struct.,* **924–926**, 493, (2009) .