


# AUTHOR QUERY FORM

	<p><b>Journal:</b> J. Chem. Phys.</p> <p><b>Article Number:</b> 016807JCP</p>	<p>Please provide your responses and any corrections by annotating this PDF and uploading it to AIP's eProof website as detailed in the Welcome email.</p>
---	--	--

Dear Author,

Below are the queries associated with your article. Please answer all of these queries before sending the proof back to AIP.

**Article checklist:** In order to ensure greater accuracy, please check the following and make all necessary corrections before returning your proof.

1. Is the title of your article accurate and spelled correctly?
2. Please check affiliations including spelling, completeness, and correct linking to authors.
3. Did you remember to include acknowledgment of funding, if required, and is it accurate?

Location in article	Query/Remark: click on the Q link to navigate to the appropriate spot in the proof. There, insert your comments as a PDF annotation.
<a href="#">Q1</a>	Please check that the author names are in the proper order and spelled correctly. Also, please ensure that each author's given and surnames have been correctly identified (given names are highlighted in red and surnames appear in blue).
<a href="#">Q2</a>	Please confirm that the asterisk after "Carlos Larriba-Andaluz" indicates the corresponding author of this paper.
<a href="#">Q3</a>	Please reword the sentence beginning with "The questions that..." so that your meaning will be clear to the reader.
<a href="#">Q4</a>	Please define MM2 at first occurrence.
<a href="#">Q5</a>	Please define QTAIM at first occurrence.
<a href="#">Q6</a>	Please define NBO at first occurrence.
<a href="#">Q7</a>	For the benefit of the reader, please insert a brief explanation of the meaning of the boldface values that appear in Table II. The explanation may be added to the caption or set as a footnote to the table.
<a href="#">Q8</a>	As per standard journal style, we have changed the footnote labels "1" and "2" to the letters "b" and "c" in Table II.
<a href="#">Q9</a>	Please define VdW at first occurrence.
<a href="#">Q10</a>	The source file for Fig. 5 did not match the figure that appeared in your manuscript. We have used the source file. Please check.
<a href="#">Q11</a>	We were unable to locate a digital object identifier (doi) for Ref. 46. Please verify and correct author names and journal details (journal title, volume number, page number, and year) as needed and provide the doi. If a doi is not available, no other information is needed from you. For additional information on doi's, please select this link: <a href="http://www.doi.org/">http://www.doi.org/</a> .

Thank you for your assistance.

# Optimization of long range potential interaction parameters in ion mobility spectrometry

Tianyang Wu,<sup>1</sup> Joseph Derrick,<sup>1</sup> Minal Nahin,<sup>1</sup> Xi Chen,<sup>1,2</sup> and Carlos Larriba-Andaluz<sup>1,3,a)</sup>

<sup>1</sup>Mechanical Engineering, Indiana University-Purdue University Indianapolis, 723 W Michigan St., Indianapolis, Indiana 46202, USA

<sup>2</sup>Mechanical Engineering, Purdue University, 585 Purdue Mall, West Lafayette, Indiana 47907, USA

<sup>3</sup>Integrated Nanosystems Development Institute (INDI), 420 University Blvd., Indianapolis, Indiana 46202, USA

(Received 16 November 2017; accepted 30 January 2018; published online XX XX XXXX)

The problem of optimizing Lennard-Jones (L-J) potential parameters to perform collision cross section (CCS) calculations in ion mobility spectrometry has been undertaken. The experimental CCS of 16 small organic molecules containing carbon, hydrogen, oxygen, nitrogen, and fluoride in N<sub>2</sub> was compared to numerical calculations using Density Functional Theory (DFT). CCS calculations were performed using the momentum transfer algorithm *IMoS* and a 4-6-12 potential without incorporating the ion-quadrupole potential. A *ceteris paribus* optimization method was used to optimize the intercept  $\sigma$  and potential well-depth  $\epsilon$  for the given atoms. This method yields important information that otherwise would remain concealed. Results show that the optimized L-J parameters are not necessarily unique with intercept and well-depth following an exponential relation at an existing line of minimums. Similarly, the method shows that some molecules containing atoms of interest may be ill-conditioned candidates to perform optimizations of the L-J parameters. The final calculated CCSs for the chosen parameters differ 1% on average from their experimental counterparts. This result conveys the notion that DFT calculations can indeed be used as potential candidates for CCS calculations and that effects, such as the ion-quadrupole potential or diffuse scattering, can be embedded into the L-J parameters without loss of accuracy but with a large increase in computational efficiency. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5016170>

## INTRODUCTION

Electrical mobility is a property that relates the average drift velocity of an ion to the electric field in a neutral buffer gas. In that respect, the mobility is a function of ion and gas properties (e.g., pressure, temperature, and mass), the ion geometrical structure, and any interaction between the gas and ion. The average effective area of interaction between the gas and the ion is normally referred to as the collision cross section (CCS),  $\Omega$ , and its calculation is key for the proper assessment of electrical mobility. The zero-field mobility equation, useful only when the electric field to gas density,  $E/N$ , is small, is given by the Mason-Schamp approximation,<sup>1</sup>

$$K = \sqrt{\frac{1}{m_{ion}} + \frac{1}{m_{gas}}} \frac{3ze}{N} \sqrt{\frac{\pi}{2k_bT}} \frac{1}{\Omega_{avg}^{1,1}}, \quad (1)$$

where  $K$  is the mobility of the ion,  $m_{gas}$  and  $m_{ion}$  are the molecular masses of the gas and the ion,  $ze$  is the ion charge,  $T$  is the effective temperature, and  $\Omega_{avg}^{1,1}$  is the orientationally averaged collision integral when all directions are assumed to be equally probable. A simplified CCS expression can be obtained analytically from a momentum transfer approach by averaging over all collision geometries, velocities, and impact parameters

as<sup>2,3</sup>

$$\Omega_{avg}^{1,1} = \frac{1}{8\pi^2} \int_0^{2\pi} d\theta \int_0^\pi \sin\phi d\phi \int_0^{2\pi} d\gamma \frac{\pi}{8} \left(\frac{m_{red}}{k_bT}\right)^3 \times \int_0^\infty g^5 e^{-\frac{m_{red}g^2}{2k_bT}} dg \int_0^\infty 2b(1 - \cos\chi(\theta, \phi, \gamma, g, b))db. \quad (2)$$

Here,  $\theta$ ,  $\phi$ , and  $\gamma$  are the 3 orientation angles,  $g$  is the relative velocity,  $b$  is the impact parameter,  $m_{red}$  is the reduced mass, and  $\chi$  is the deflection angle which depends on the ion-gas interaction potentials, the relative velocity of the ion-gas, the impact parameter, and the orientation of the ion as<sup>4</sup>

$$\chi(\phi, \theta, \gamma, g, b) = \pi - 2b \int_{r_m}^\infty \frac{dr}{r^2 \sqrt{1 - \frac{b^2}{r^2} - \frac{\Phi(r)}{\frac{m_{red}g^2}{2}}}}, \quad (3)$$

where  $r$  is the ion-gas molecule distance,  $r_m$  is the distance of closest approach, and  $\Phi(r)$  is the ion-gas interaction potential. Even when the interaction potential  $\Phi(r)$  is well known, the integral in Eq. (3) cannot be calculated analytically except for the simplest cases and therefore requires the deflection angle  $\chi$  to be calculated numerically.

The deflection angle  $\chi$  can be easily calculated numerically when a suitable potential governs the equations of motion. A 4-6-12 potential is currently the most commonly used interaction potential in trajectory method algorithms and

<sup>a)</sup>Author to whom correspondence should be addressed: [clarriba@iupui.edu](mailto:clarriba@iupui.edu)

it is described by

$$\Phi(x, y, z) = 4\epsilon \sum_{i=1}^n \left[ \left( \frac{\sigma}{r_i} \right)^{12} - \left( \frac{\sigma}{r_i} \right)^6 \right] - \frac{\alpha}{2} \left( \frac{ze}{n} \right)^2 \times \left[ \left( \sum_{i=1}^n \frac{x_i}{r_i^3} \right)^2 + \left( \sum_{i=1}^n \frac{y_i}{r_i^3} \right)^2 + \left( \sum_{i=1}^n \frac{z_i}{r_i^3} \right)^2 \right]. \quad (4)$$

In this case,  $r_i = (x_i, y_i, z_i)$  is the distance between the gas molecule and each of the  $n$  atoms,  $\alpha$  is the polarizability of the buffer gas that governs the ion-induced dipole term, and  $\epsilon$  and  $\sigma$  are the Lennard-Jones (L-J) parameters corresponding to the well-depth and intercept (zero potential crossing), respectively.

While the effect of polarizability is well established, there is still considerable ambiguity regarding the Lennard-Jones parameters employed in Eq. (4). The regular  $\sigma - \epsilon$  pairs generally employed in other calculations, e.g., molecular dynamics, cannot be inserted directly in expression (4) without incurring in large errors. In fact, very few of such parameters are optimized or even known for CCS calculations. The rationale behind this marked difference can be explained from the fact that Eqs. (1)–(4) assume several simplifications including that the atoms are fixed, that higher order potential interactions are present, and that the energy/momentum is conserved upon collisions. While neglecting all these effects might work reasonably well for He with smaller or very rigid molecules, this is not expected to be the case when using larger heavier gases such as  $N_2$  or  $CO_2$ .<sup>5–9</sup> This has been well observed when comparing  $N_2$  CCS ratios of experimental values with Exact Hard Sphere Scattering (EHSS) or Projection Approximation (PA) calculations. Under such scenarios, the ratios vary between 1 and 1.3 for EHSS scattering calculations and seem to be a constant value tending to  $\sim 1.36$  over that of PA calculations.<sup>10–14</sup>

To account for this particular difference, modifications have been proposed to the different numerical methods when dealing with  $N_2$  gas. Most notably two of them stand out: (1) the addition of ion-quadrupole potential interaction together with L-J parameter optimization<sup>15,16</sup> and (2) the addition of artificial diffuse hard sphere scattering to account for the vibration of the atoms and the energy and momentum exchange.<sup>17,18</sup>

The first modification is the optimization of L-J potentials with the addition of an ion-quadrupole potential. The quadrupole moment is obtained by placing one negative charge of  $0.4825 e$  on each nitrogen and one positive charge of  $0.965 e$  in the center of the molecule. In such a way, the quadrupole potential can be expressed as<sup>15,16</sup>

$$\Phi_{IQ}(x, y, z) = \sum_{j=1}^3 \sum_{i=1}^n \frac{q_i q_j e^2}{r_{ij}^3}. \quad (5)$$

Here the index  $j$  denotes 3 different  $N_2$  charges (2 being the center charge) and index  $i$  indicates the charges on the ion/atoms. Simultaneously, the orientation of the  $N_2$  molecule is taken into account by assuming an appropriately weighted impact parameter. While this addition yields very accurate results, it makes the calculation rather inefficient and still requires the

L-J parameters to be artificially modified to account for other effects.

A second modification is to resort to hard sphere diffuse scattering (as opposed to specular) instead of using L-J parameters. The resulting CCS when using artificial diffuse scattering agrees extremely well with experimental results for large ions.<sup>19,20</sup> In fact, a diffuse reemission is the calculation of choice that Millikan, Epstein, Rosner,<sup>11,12,21,22</sup> and many authors use when dealing with larger particles. These diffuse scattering calculations (with the addition of the ion-induced dipole potential) may be considered as an upper bound for CCSs.

Interestingly enough, both modifications are in agreement for CCSs larger than  $2 \text{ nm}^2$  ( $200 \text{ \AA}^2$ ).<sup>23</sup> The reason behind the coincidence cannot be related to the effect of the ion-quadrupole potential (which can be integrated into the effective size of the  $N_2$  molecule) but to the artificially modified L-J parameters. From a physical point of view, increasing the well-depth or decreasing the intercept of L-J parameters of nearby atoms will indirectly enhance their contribution to the trajectory of the gas molecule. This indirect effect can be thought of as a series of multidirectional contributions which overall make the reemission effectively diffuse (in contrast to effectively specular if only one atom was involved). Given that using appropriate density functional theory (DFT) calculations with trajectory methods, results of calculated CCS normally agree within 4% of experimental counterparts,<sup>24–33</sup> it is clear that many of the additional contributions to the CCS such as vibration of atoms, rotational effects, exchange of energy due to non-zero relative velocity effects, ion-quadrupole, and higher order interactions can be safely added to artificially modified L-J potentials without loss of accuracy but with a large gain of computational efficiency. This effect should be particularly true for small semi-rigid organic-type molecules.

~~The questions that arise are,~~ how does one optimize the CCS L-J parameters correctly for a given atom in a molecule?, and is any molecule with a given element an appropriate candidate to carry out the optimization? Previous trials at optimizing L-J potentials have been quite successful,<sup>3,31,34</sup> but the optimization methods employed cannot be extrapolated or generalized as they rely on trial and error approaches. In those trials, the ratio of  $\epsilon/\sigma$  is kept constant for every element, starting from some initial candidate values and multiplying both parameters by the same factor. The reason for this constancy might have its support on some type of coupling between  $\epsilon$  and  $\sigma$  although an explanation is not given.

The purpose of this work is to design a straightforward optimization method that will correctly assign  $\epsilon$ ,  $\sigma$  pairs if corresponding values of CCSs are known experimentally and structural models of the appropriate ions exist. Here, 16 molecules for which their CCSs have been well determined experimentally in  $N_2$  and which have well-defined DFT structures have been used for the optimization of carbon, hydrogen, oxygen, nitrogen, and fluoride. In particular, it is shown that the ion-quadrupole potential effect can be adsorbed into the L-J parameters, increasing computational efficiency without loss of accuracy. As a consequence of the optimization method of choice, the effect of L-J pairs can be studied for one element at

a time, yielding a lot of insights into the process. While there is no need to have any knowledge of initial values for any Lennard-Jones pair, a known interval range for the parameters is important to determine the optimized values more efficiently. Simultaneously, the importance of choosing an appropriate molecule for a given element is also studied. Interestingly, a deterministic  $\sigma - \epsilon$  relation appears as a consequence of trying to determine L-J pairs through CCS calculations. New well-depths and zero intercept distances for C, H, O, N, and F elements with respect to  $N_2$  are established with minimal errors.

Finally, the optimization brought forth in this manuscript and the choice to remove the ion-quadrupole potential should not be confused with the need to continuously improve the way CCSs are theoretically calculated, and that additional CCS contributions should continuously be studied to improve existing methods. Some recent publications have been aimed towards this purpose. For example, recent work has focused on testing other empirical formulation for the repulsion exponent of the L-J potentials<sup>35</sup> pointing out its importance for more accurate CCS calculations. Similarly, several authors<sup>35–37</sup> have pointed out that the use of different parameters for elements such as carbon should be employed under different scenarios much like it is done in molecular dynamics. Other effects that have been studied are the effect of dipole moments<sup>38</sup> and planar molecules<sup>31</sup> that might lead to preferred orientations.<sup>39</sup> The importance of this work relies on the fact that one could optimize the L-J potential to yield even more accurate data if these, or other, contributions were added and if reliable data were available.

## METHODS AND THEORY

### Optimization

In order to generalize the study so that it can be used under any conditions and any elements, a quadratic function is used to minimize the error between experimental and theoretical collision cross sections. Given  $n$  molecules with a total of  $k$  elements, the multivariate function mapping  $\mathbb{R}^{2k} \rightarrow \mathbb{R}$  to be minimized is given by

$$\mathcal{F}(\epsilon_1, \sigma_1, \epsilon_2, \sigma_2, \dots, \epsilon_{k-1}, \sigma_{k-1}, \epsilon_k, \sigma_k) = \sum_{i=1}^n \left( 1 - \frac{\Omega_{iCalc}(\epsilon_1, \sigma_1, \epsilon_2, \sigma_2, \dots, \epsilon_{k-1}, \sigma_{k-1}, \epsilon_k, \sigma_k)}{\Omega_{iExp}} \right)^2, \quad (6)$$

where  $\Omega_{iCalc}$  and  $\Omega_{iExp}$  are the calculated and experimental CCSs for the  $i$ -th molecule, which can depend on the Lennard-Jones parameters  $\sigma$  and  $\epsilon$  of any given number of elements as long as  $n \geq 2k$  to avoid indetermination. To be mathematically unambiguous, this last requirement has to be met partially as well so that there has to be at least two molecules that contain any given element. Note that if the experimental and numerical CCSs are equal for all molecules, Eq. (6) will be zero. The problem of optimization is therefore that of a multivariate minimization process subject to

$$\nabla \mathcal{F} = \left( \frac{\partial \mathcal{F}}{\partial \epsilon_1}, \frac{\partial \mathcal{F}}{\partial \sigma_1}, \frac{\partial \mathcal{F}}{\partial \epsilon_2}, \frac{\partial \mathcal{F}}{\partial \sigma_2}, \dots, \frac{\partial \mathcal{F}}{\partial \epsilon_k}, \frac{\partial \mathcal{F}}{\partial \sigma_k} \right) = 0. \quad (7)$$

While this is a straight forward mathematical calculation using the gradient descent (or conjugate gradient) method, the physical

interpretation of the results can be somewhat complicated since the quadratic function  $\mathcal{F}$  is only known at discrete points on a  $2k$ -th dimensional space and might contain non-physical solutions. The [supplementary material](#) has an example of the application of this optimization method for 5 molecules and 2 elements while underlining its problems.

A simple option instead is to minimize only a L-J pair for one element at a time while keeping the rest of the parameters constant through a *ceteris paribus* all-mapping technique. While this method is more computationally involved, it will give a lot more insight into the physical explanation of the results obtained. In such case, the process is as follows:

- (1) Select an element (e.g., carbon) for which to optimize  $\mathcal{F}$ .
- (2) Optimize  $\mathcal{F}(\sigma_C, \epsilon_C)$  for the L-J pair,  $\sigma_C, \epsilon_C$ , of the chosen element while keeping the rest of the parameters constant (e.g.,  $\sigma_H, \epsilon_H, \sigma_N, \epsilon_N, \dots$ ). Note that this optimization does not necessarily yield  $\mathcal{F} \sim 0$  as the rest of the elements have not been modified. To acquire the most information possible, the whole surface  $\mathcal{F}$  is mapped around the region of interest.
- (3) Optimize  $\mathcal{F}$  using steps (1) and (2) for each of the rest of the elements (one by one) to complete the rest of the iteration.
- (4) Iterate starting from the first element using the newly corrected L-J pairs obtained in (1)–(3) in the previous iteration.
- (5) Continue iterating until the Lennard-Jones pairs satisfy a convergence criterion.

The concept of surface mapping is generalized as a block diagram in Fig. 1 which depicts the algorithm layout for one element. In short, an interval and a step size are chosen for the potential well-depth and intercept of the element of interest. For each well-depth/intercept pair, a CCS for every molecule is calculated using *IMoS*.<sup>40</sup> Leaving the rest of the L-J parameters constant, the value of  $\mathcal{F}(\sigma_A, \epsilon_A)$  is calculated, where the subindex  $A$  stands for the element of choice. This leads normally to several thousand CCS calculations in order to map the whole surface. As an example, Fig. 2 shows  $\mathcal{F}$  as a

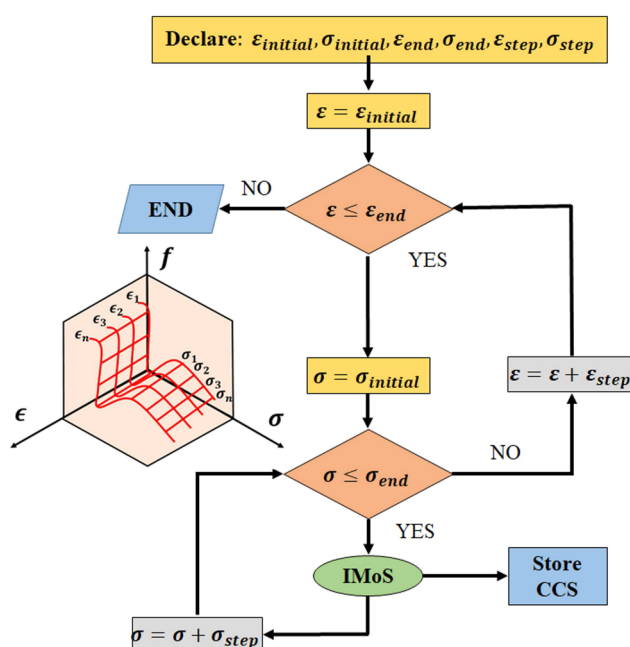


FIG. 1. Roadmap of the mapping process of the optimization function  $\mathcal{F}$  is depicted for any atom.



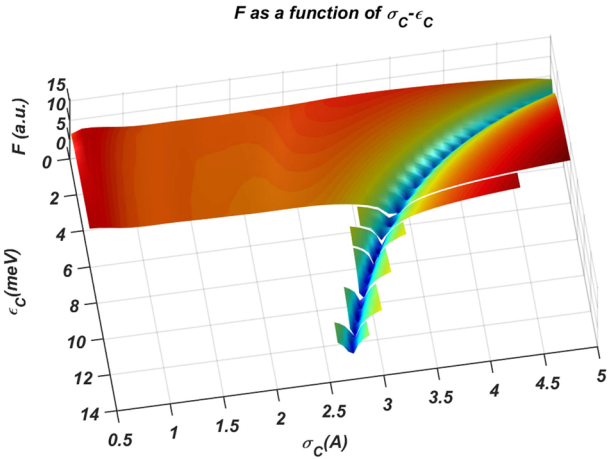


FIG. 2. A wide surface mapping of  $\mathcal{F}$  for the carbon atom extended in the region of interest.

function of the *ceteris paribus* pair  $\sigma_C - \epsilon_C$  for the carbon element. If the initial interval does not cover the region of interest, it can be extended as shown in the figure. The fully mapped surface can now be used to find the expected minimum for the selected element parameter set.

**Structures, experimental CCS, and theoretical CCS—DFT calculation effectiveness**

16 molecules were employed for the optimization of the L-J parameters for C, H, O, N, and F. The experimental CCSs for these 16 molecules, which appear in Table I, were obtained directly from a careful analysis from the work of Campuzano *et al.* at 301 K.<sup>31</sup> The reason for this choice is that Campuzano *et al.* used the same results to optimize N<sub>2</sub> L-J parameters when an ion-quadrupole potential [Eq. (5)] was employed

using the trajectory method algorithm. Their resulting optimized parameters are a good comparison point to use for our own optimization, which now excludes the ion-quadrupole potential.

In order to numerically calculate the CCS of a molecule of choice by the procedures stated, there is a need to establish the geometrical location of each of the atoms. This is by no means a simple endeavor as the atoms are in constant motion and the molecule is subject to possible deformations from gas molecule collisions. It has been shown that hybrid functionals in density functional theory (DFT) can be used as a good estimate for the optimal position of the atoms.<sup>24–33</sup> Note, however, that resulting DFT structures are not a function of temperature or pressure and therefore cannot be considered the actual geometry at that temperature, not even in the average sense. Nonetheless, DFT calculations provide a well-defined geometry which can be obtained by unambiguous calculations and can serve as models for L-J parameter optimization and CCS calculations, with emphasis on rigid and semi-rigid *small* molecules. With this in mind, CCS L-J parameters obtained using DFT calculations should then be used with caution when other calculation methods are employed for geometrical characterization of the molecule (e.g., molecular dynamics) as the parameters will contain the effect of temperature and pressure, albeit generally small, embedded in them.

Initial structures for the 16 molecules appearing in Table I were generated using *Chem3D Pro v.12*(Perkin Elmer) and *MM2* calculations. Calculations of frequency and geometry optimization were undertaken using different Becke-Lee-Yang-Parr and Becke-Perdew type functionals including cam-b3lyp, b3pw91, bpv86, and b3lyp.<sup>41–43</sup> The Pople split-valence basis set used was 6-31++G(d,p) with diffuse and polarization functions. Most calculations yielded very similar results

TABLE I. CCS for 16 molecules studied. Expt. CCS—Experimental CCS, Ref. 31. CCS ion-quad—CCS calculated in the N2 using ion-quad potential, Ref. 31. CCS LJ—CCS calculated with previously reported L-J parameters with no ion-quadrupole. CCS1—CCS calculated using carbon first. CCS2—CCS calculated using hydrogen first. Error—% of absolute error when comparing calculated CCS with experimental CCS.

Molecule	Expt. CCS (Å <sup>2</sup> )	CCS ion- quad (Å <sup>2</sup> )	Error (%)	CCS LJ (Å <sup>2</sup> )	Error (%)	CCS1 (Å <sup>2</sup> )	Error (%)	CCS2 (Å <sup>2</sup> )	Error (%)
Triphenylene	143.30	149.20	4.12	140.44	2.00	145.87	1.79	144.61	0.92
Nethylaniline	124.50	124.50	0.00	112.95	9.28	120.42	3.28	119.42	4.08
Dexamethasone	190.70	190.40	0.16	176.77	7.30	190.28	0.22	188.75	1.02
Acetaminophen	131.10	136.20	3.89	120.45	8.12	129.84	0.96	128.99	1.61
Betamethasone	189.60	189.40	0.11	176.25	7.04	189.32	0.15	187.51	1.10
Anthracene	129.60	136.20	5.09	127.38	1.71	132.28	2.07	130.86	0.97
Choline	115.40	115.90	0.43	104.17	9.73	114.53	0.75	113.27	1.85
Phenanthrene	129.10	135.70	5.11	126.74	1.83	131.69	2.00	130.32	0.95
Acetylcholine	127.80	128.00	0.16	117.80	7.82	128.54	0.58	127.13	0.52
c60	213.10	212.90	0.09	214.04	0.44	213.61	0.24	214.35	0.59
c70	231.40	229.80	0.69	231.70	0.13	230.81	0.25	231.43	0.01
Naphthalene	115.80	121.10	4.58	111.28	3.90	116.28	0.42	114.97	0.72
Paracetamol2	131.10	136.20	3.89	121.25	7.51	130.63	0.36	130.09	0.77
Pyrene	135.00	139.90	3.63	131.78	2.39	136.88	1.39	135.77	0.57
TtEA	122.20	122.20	0.00	113.43	7.18	122.75	0.45	120.75	1.18
TMA	107.40	107.20	0.19	95.31	11.26	104.06	3.11	102.79	4.30
Error avg(%)			2.01		5.48		<b>1.13</b>		1.32

(with negligible effects in the CCS) and resulting structures have been added to the [supplementary material](#). Both Mulliken and Hirshfeld partial charges were calculated. While significant differences were obtained in the placement of the partial charges, no large differences were observed in the CCS calculations. QTAIM and NBO methods were not employed. For generalization purposes, the choice of partial charges in the calculations shown is from the Mulliken approach.

CCSs are then calculated numerically using the above DFT structures and the ion mobility calculator *IMoS*. The trajectory method with a 4-6-12 potential (L-J and ion-induced dipole) and no ion-quadrupole interaction was the calculation of choice with a polarizability of  $1.7 \cdot 10^{-24} \text{ cm}^3$ . The program was run with 8, 16, or 32 cores indistinctively.  $9 \cdot 10^5$  gas molecules were used for every calculation with each molecule taking just a few seconds to complete under most scenarios. The need to calculate several thousand CCSs to map the surfaces of some atoms took several hours. The number of gas molecules was increased for the gradient descent method (provided in the [supplementary material](#)) to  $3 \cdot 10^7$  gas molecules to increase accuracy when the program error was too large (brought down from  $\sim 1\%$  to  $0.1\%$ ) to appropriately define the gradient. The output of the *IMoS* code for a sample calculation has been added to the [supplementary material](#).

## RESULTS AND DISCUSSION

The *ceteris paribus* optimization process was carried out starting from two different initial elements, carbon and hydrogen. The reason behind this double choice is to verify that our results are not biased by the initial condition or element chosen. For both studies, there is no need to establish initial L-J parameter pairs for any of the atoms. However, for comparison purposes, L-J parameters were taken from previous studies<sup>31</sup> and are shown in the first portion of Table II. For those specific parameters, using the TM method with and without the ion-quadrupole potential, Table I also shows the numerical CCS and its error when compared to the experimental CCS. The average error is also provided at the bottom of the table.

### Optimization starting with carbon first

#### 1st iteration

Figure 3(a) shows the mapped  $\mathcal{F}$  surface as a function of  $\sigma_{C,1}$  and  $\epsilon_{C,1}$  while keeping the rest of the L-J pairs

TABLE II. L-J parameters. MD-LJ—average molecular dynamics L-J parameters pairs using combinatorial rules and  $N_2$   $\sigma = 3.64 \text{ \AA}$  and  $\epsilon = 102 \text{ K}$ . Refs. 44 and 45.

L-J Elements	Previous <sup>a</sup>		Opt. <sup>b</sup>		Opt. <sup>c</sup>		MD-LJ	
	$\epsilon$ (meV)	$\sigma$ (Å)	$\epsilon$ (meV)	$\sigma$ (Å)	$\epsilon$ (meV)	$\sigma$ (Å)	$\epsilon$ (meV)	$\sigma$ (Å)
C-N <sub>2</sub>	4.2319	3.5814	<b>4.65</b>	<b>3.5</b>	4.1	3.6	6.32	3.535
H-N <sub>2</sub>	0.8204	1.2409	<b>3</b>	<b>2.3</b>	3.3	2.2	4.09	2.77
O-N <sub>2</sub>	2.4195	3.255	<b>4.3</b>	<b>3.5</b>	4.3	3.5	4.78	3.38
N-N <sub>2</sub>	3.5902	4.392	<b>5.8</b>	<b>4.2</b>	5.8	4.2	5.13	3.64
F-N <sub>2</sub>	<b>2.0162</b>	<b>3.1285</b>	<b>1.2</b>	<b>3.4</b>	<b>1.2</b>	<b>3.4</b>	<b>1.51</b>	<b>3.375</b>

<sup>a</sup>Taken from Ref. 31.

<sup>b</sup>Parameters refer to optimization using carbon first.

<sup>c</sup>Refers to Opt. using hydrogen first.

constant. Here, index 1 corresponds to the first iteration. The  $\sigma_{C,1}$  and  $\epsilon_{C,1}$  values, given in Angstroms and meV, respectively, were varied in increments of 0.1, and the resulting CCS were used together with the experimental CCS to obtain  $\mathcal{F}(\sigma_C, \epsilon_C)$  using Eq. (6). The value of  $\mathcal{F}$  appearing in Fig. 3(a) (and in some surface mapping plots) has been modified and cut at a given height to accent the valley of the minimum for representation purposes. The initial domain chosen was  $\epsilon_C \in [2, 9]$  and  $\sigma_C \in [2.5, 4.5]$  which was expected to contain the minimum. Figure 3(a) shows a smooth decrease throughout most of the values until a minimum is reached at some  $\sigma_C - \epsilon_C$  relation. Unexpectedly, this minimum is not unique and the aforementioned  $\sigma_C - \epsilon_C$  relation happens for a series of well-depths and corresponding intercepts. The optimized value of  $\mathcal{F}$  is therefore not exclusive but contained in a region which can be referred to as the “line of global minimums.” To verify that there was no defined global minimum, the plot was extended (up to  $\epsilon_C = 13 \text{ meV}$ ) in the region of interest. However, no significant minimum was found for any given pair. The optimal point is therefore not singular, and there is a relation between the well-depth and the intercept that corresponds to the line of global minimums. This relation can be graphed [see Fig. 3(a)] and corresponds to an exponential relation between  $\sigma_C$  and  $\epsilon_C$  [see Fig. 3(a),  $R^2 \sim 0.99$ ]. This exponential relation gives a lot more importance to the intercept value than to the well-depth in the region of interest, as small fluctuations in the well-depth affect much less than those in the intercept. While there is a simple mathematical explanation for the relation, i.e., a similar effect can be achieved by reducing the intercept

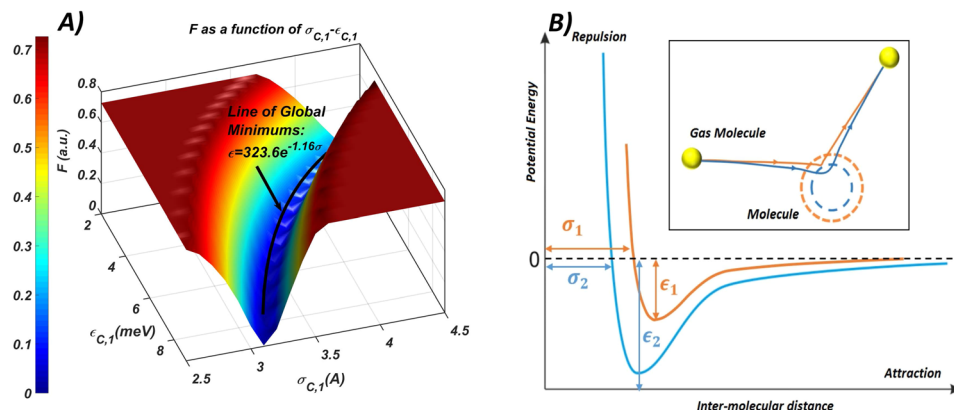


FIG. 3. (a) Mapped optimization surface for carbon atoms in the first iteration. A “line of global minimums” yielding semi-optimized CCSs can be observed. The line of global minimums corresponds to an exponential relation ( $R^2 \sim 0.99$ ). (b) Sketch of two very distinct L-J pairs yielding the same deflection angle and thus the same CCS.

and increasing the well-depth or vice versa, there is no particular physical reasoning, to our knowledge, other than that there is no enough information in the given data to fully separate the parameter interaction. Figure 3(b) shows a depiction of two very different instances of  $\sigma - \epsilon$  pairs that will yield very similar values of CCS and therefore similar values of  $\mathcal{F}$ . Nonetheless, there are very interesting consequences to this analysis, having the possibility of correctly defining the optimized CCS the most important one. As is demonstrated in the [supplementary material](#), two sets of L-J parameters for carbon which are very different from each other can give very similar CCSs. Perhaps if there were sufficient data of CCSs as a function of temperature, a unique answer to the pair could be found.

To continue with the rest of the iteration, a choice must be made for the well-depth intercept pair. The intercept value was chosen to be that of the sum of the [VdW](#) radii of N<sub>2</sub> (1.82 Å) and C (1.7 Å) corresponding to  $\sigma_{C,1} = 3.52$  Å. For this chosen value, and using the exponential relation, the corresponding well-depth that minimizes  $\mathcal{F}$  is given by  $\epsilon_{C,1} = 5.49$  meV. These first values for carbon are in agreement with previous values used in CCS calculations in N<sub>2</sub>, albeit high for the well-depth.

For the rest of the first iteration, atom parameters were optimized in the order H, O, N, and F. The results are all shown in Fig. 4. Hydrogen and oxygen, in Figs. 4(a) and 4(b), both follow a very similar pattern to carbon with their respective exponential relation. For hydrogen, the plot is markedly non-symmetric. The reason for this unbalance is understood from the effect of hydrogen as the external-most atom in a molecule. For any constant value of the potential well-depth, and starting from the line of global minimums, increasing the intercept over the line of global minimums has a strong effect on the  $\mathcal{F}$  function and therefore on the CCS, which

is expected as the hydrogen atom is enlarged. On the other hand, decreasing the intercept at constant well-depth starting once again from the line of global minimums leads to a much less pronounced change on  $\mathcal{F}$ . In such a case, the reduction of the effect of the repulsive potential in the hydrogen atom is confronted by the increased effect of the heavier atoms close to the hydrogen atom which softens the overall effect. This result is quite important, as it may lead to the possibility of completely removing the hydrogen element from CCS calculations, if its effect is embedded into the heavier atoms, which may lead to much faster computational times. A choice was made for the values of the hydrogen atom L-J pair with  $\epsilon_{H,1} = 2.33$  meV and  $\sigma_{H,1} = 2.22$  Å. The option to use a smaller intercept than that of the sum of the VdW radii was made in this case, choosing H to be closer to its covalent radius (40 pm) to observe if it would affect the optimization procedure. With this new set of parameters for carbon and hydrogen, oxygen was optimized. Oxygen has a wider valley than any of the previous atoms. This might be reasoned from its smaller influence on the set of molecules as only 6 out of 16 molecules contain oxygen with a very small percentage (~3% of the total number of atoms). Another interesting feature for oxygen is that there seems to be a well-localized global minimum in the line of minimums. However, one must note here that there is really no appreciable difference in the CCS for any of the L-J pairs belonging to the line of minimums. Since having a very large potential well-depth seems physically unreasonable, a choice was made for  $\epsilon_{O,1} = 4$  meV and  $\sigma_{O,1} = 3.45$  Å.

Following the same procedure, Figs. 4(c) and 4(d) correspond to nitrogen and fluoride elements. Nitrogen mimics oxygen with a wide valley that contains a line of minimums leading to a global minimum. Nitrogen only appears in 7 out of 16 molecules with 1% overall abundance. In Fig. 4(c), there

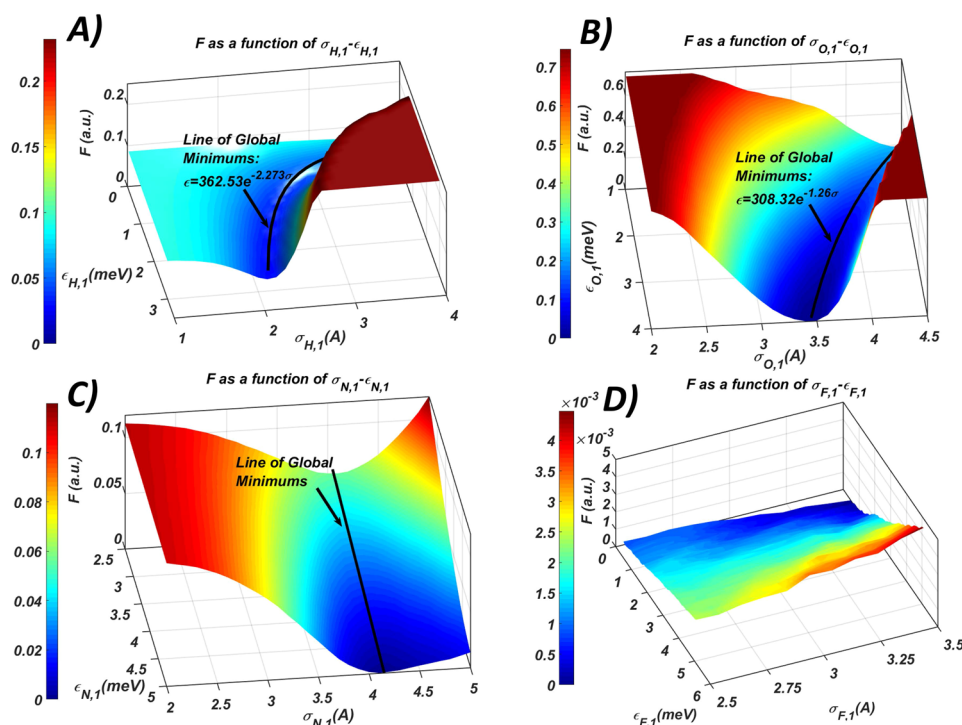


FIG. 4. Mapped surfaces of the optimization function  $\mathcal{F}$  as a function of the L-J pairs for the first iteration for (a) hydrogen, (b) oxygen, (c) nitrogen, and (d) fluoride.

does not seem almost any dependence of the intercept  $\sigma$  on the value of  $\mathcal{F}$ , where as now the valley of minimums drops as the well-depth is increased. An interpretation can be made from these results; given that the position of the nitrogen atom in the set of molecules studied is normally centered on the molecule, e.g., tetra-alkyl chains, the influence of this atom on the CCS is exerted indirectly through the well-depth, as direct collisions rarely happen.  $\epsilon_{N,1} = 5$  meV and  $\sigma_{N,1} = 4.2$  Å were chosen for nitrogen as the parameters for the first iteration. Finally, fluoride was optimized. Fluoride only appears in 2 molecules (and only 2 atoms) of two very similar epimers: betamethasone and dexamethasone. Fluoride has also a very small influence on such molecules and therefore, any value of the L-J pairs will not modify the optimization function  $\mathcal{F}$  substantially. In particular, for this first iteration, there is no marked minimum anywhere for the fluoride atom, in contrast to the rest of the atoms. To continue with the next iteration, a pair of values were chosen close to the interval minimum,  $\epsilon_{F,1} = 0.75$  meV and  $\sigma_{F,1} = 3.5$  Å.

## 2nd iteration

Figure 5 shows the results for carbon (top) and hydrogen (bottom) for the second iteration when compared to the first iteration. The calculated range of values has been narrowed to the region of interest to avoid excessive computational time. From the comparison, it is quite clear that the line of global minimums has slightly shifted and the valleys have become narrower as well, and the values of  $\mathcal{F}$  are closer to zero, indicative of convergence. Overall, the new L-J well-depth

and intercept for the new iteration were set to be  $\epsilon_{C,2} = 5$  meV and  $\sigma_{C,2} = 3.52$  Å for carbon and  $\epsilon_{H,2} = 3$  meV and  $\sigma_{H,2} = 2.2$  Å for hydrogen.

Figures 6(a)–6(c) show the results for the second iteration for oxygen, nitrogen, and fluoride. The trends for oxygen and nitrogen are very similar with values shifting slightly from those of iteration 1. An emphasis must be made on the fact that the range of values of  $\mathcal{F}$  varies very little in the scrutinized region, so the error in the CCS from any of the values in the range selected is expected to be not more than 2%. The choice of parameters was  $\epsilon_{O,2} = 4.3$  meV and  $\sigma_{O,2} = 3.5$  Å for oxygen and  $\epsilon_{N,2} = 5$  meV and  $\sigma_{N,2} = 4.1$  Å for nitrogen. The fluoride atom in Fig. 6(c) now shows a global minimum. The change in the optimization of  $\mathcal{F}$  is minimal, and the limit in accuracy of *IMoS* using  $9 \cdot 10^5$  gas molecules is becoming more apparent from the wavy pattern. A value can be chosen for the L-J parameters of fluoride as  $\epsilon_{F,2} = 1.2$  meV and  $\sigma_{O,2} = 3.4$  Å.

## 3rd iteration

Only carbon and hydrogen made any significant changes on the third iteration which was considered the final iteration. Figure 7 shows the optimization function  $\mathcal{F}$  as a function of the L-J parameters for (A) carbon and (B) hydrogen where a very small shift is again observed in the well-depth. The final values are  $\epsilon_{C,3} = 4.65$  meV and  $\sigma_{C,3} = 3.5$  Å for carbon and  $\epsilon_{H,3} = 3$  meV and  $\sigma_{H,3} = 2.3$  Å for hydrogen. The final L-J parameters resulting from the optimization are portrayed in

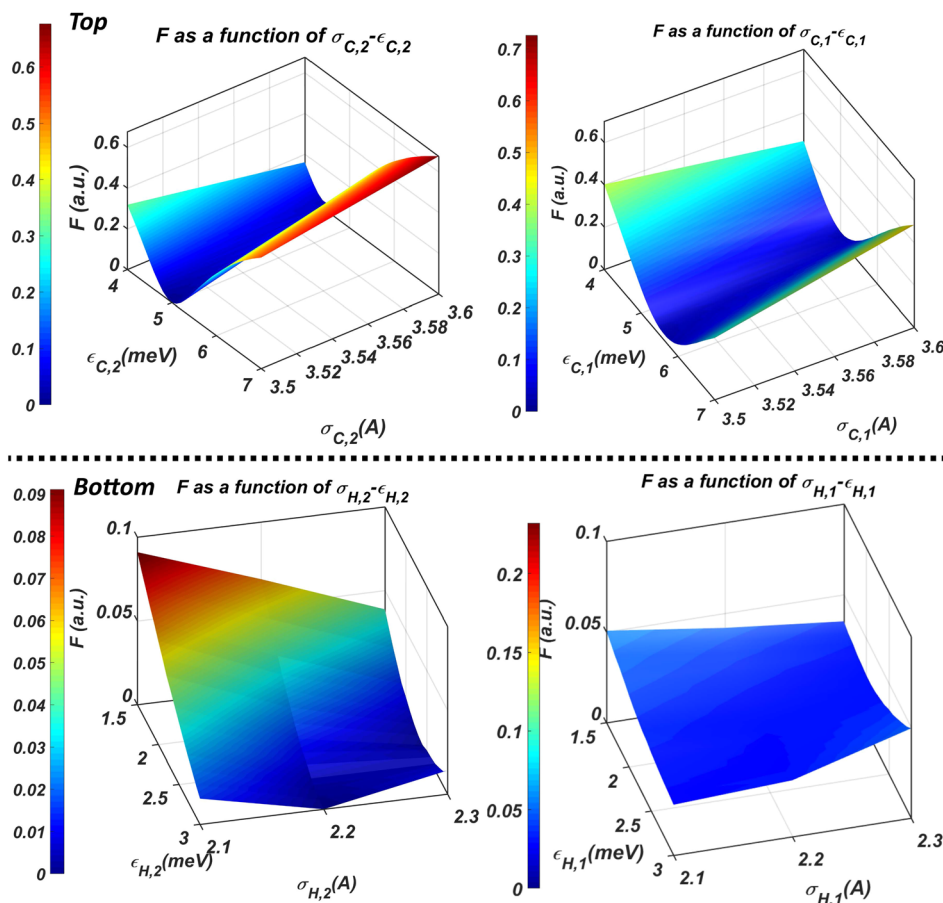


FIG. 5. Comparison between two successive iterations of the (top) carbon and (bottom) hydrogen atoms for the region of interest. On new iterations, the valley shifts and narrows.  $\mathcal{F}$  also yields a lower value (closer to optimization).



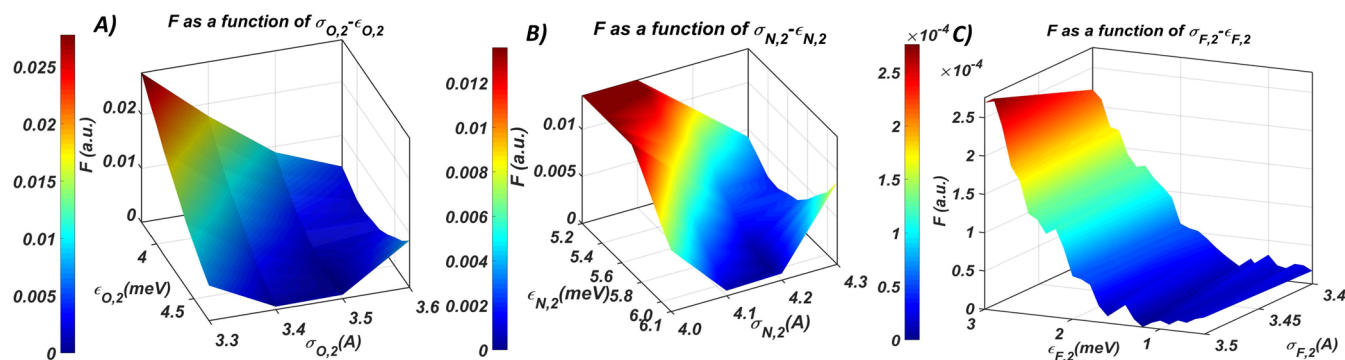


FIG. 6. Optimization function  $\mathcal{F}$  as a function of L-J parameters for the second iteration for (a) oxygen, (b) nitrogen, and (c) fluoride.

Table II in comparison to the initial parameters. When comparing it to previously reported L-J parameters, the intercept of all atoms except for the hydrogen atom has not been strongly modified. However, the hydrogen intercept used as a comparison was undoubtedly small for an  $\text{N}_2$ -H pair. On the other hand, the potential well-depth has been strongly increased for all atoms. While variations of the potential well-depth have a small effect on the CCS, this strong variation is indicative of its importance in the optimization of the calculation, and how perhaps some of the effects of the ion-quadrupole potential have been integrated into the potential wells. Table II also shows the expected Lennard-Jones parameters using Lorentz-Berthelot<sup>46,47</sup> rules of combination and regular molecular dynamics L-J pairs. The CCS optimized parameters are somewhat in agreement with those generally used in other fields.

### Optimization starting with hydrogen first

Since the optimized values of the L-J pairs are not unambiguously defined, there exists the possibility that optimizing the atoms L-J pairs in a different order could provide more weight to some atoms in detriment of others while not seeing substantial differences in the CCSs. For this reason, a new optimization routine was initiated where hydrogen was chosen as the starting element, followed by C, O, N, and F. The results are shown in Fig. 8 for the elements with the largest weight, carbon and hydrogen, after 3 iterations. The Lennard-Jones parameters are slightly different from those previously obtained yielding  $\epsilon_{\text{C},3} = 4.1$  meV and  $\sigma_{\text{C},3} = 3.6$  Å for carbon and

$\epsilon_{\text{H},3} = 3.3$  meV and  $\sigma_{\text{H},3} = 2.2$  Å for hydrogen with no significant change for the other elements. The results can be observed in Table II.

### Results and discussion of CCS calculations using optimized L-J parameters

Table I shows the CCS calculations using the optimized L-J parameters for both cases studied. With the new optimized parameters, the errors associated with the calculations are minimal with most CCS within 1% of their experimental counterparts. The L-J parameters could have been further optimized (e.g., to 4-digit accuracy), yielding slightly improved results (see the [supplementary material](#) for errors within 0.7% for 5 molecules), but higher precision for this set of molecules is not guaranteed to yield better results for other molecules. It is clear however from the results that there is no need to explicitly calculate the ion-quadrupole potential when trying to obtain approximate CCSs in  $\text{N}_2$  if the L-J potentials are properly optimized. The addition of the ion-quadrupole potential provides no gain in accuracy while greatly increasing the computational time. This should not be mistaken with the fact that the ion-quadrupole potential does exist along with other effects which should not be ignored. However, some of these effects can be embedded into L-J potentials without significant errors in the CCS calculations under most circumstances. Other effects such as the CCS differences arising from planar polyaromatic hydrocarbon structures encountered by Campuzano *et al.*<sup>31</sup> have disappeared when the L-J potentials are optimized in this way. Since there is not enough information of how the

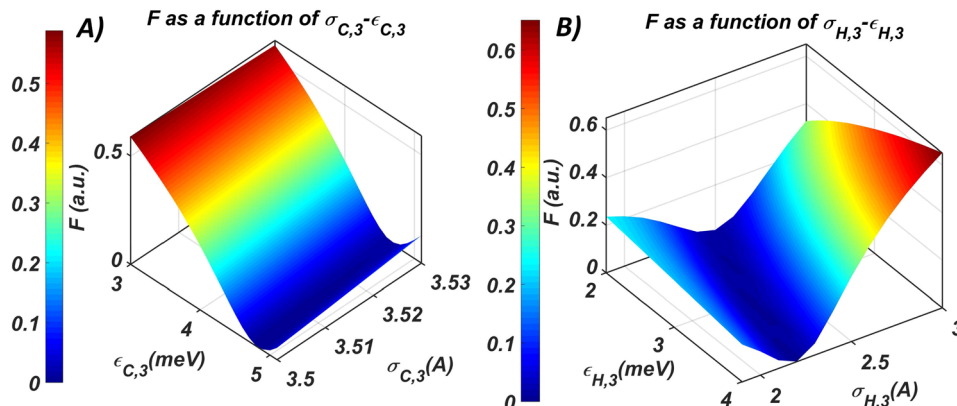


FIG. 7. Final iteration using the results for (a) carbon and (b) hydrogen.

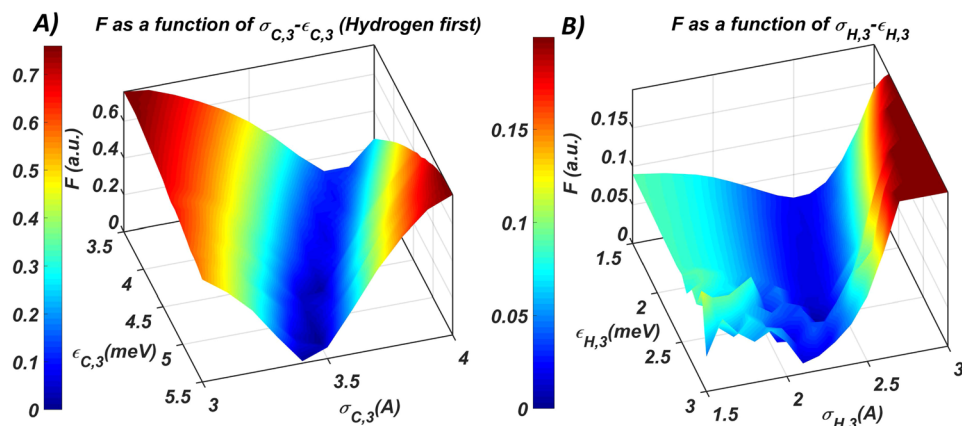


FIG. 8. Final iteration for (a) carbon and (b) hydrogen when hydrogen L-J pairs are optimized first.

parameters were optimized in their work, it is difficult to fully assess the reason for such discrepancy. There is reason to believe that very planar molecules could have some preferred alignment, but this effect could not be characterized with the existing calculation methods which assume all orientations are equally probable.

## CONCLUSIONS

A general method for the optimization of the Lennard-Jones potential parameters, intercept and potential well-depth, has been proposed. The following conjectures were used for the calculations:

- DFT calculations were chosen as the method of choice for rigid and semi-rigid small molecules. While the geometry in the gas phase at given conditions of temperature and pressure is expected to be different from the DFT geometry and frequency optimization values, DFT calculations provide a well-defined and reproducible way to generate reliable geometries. DFT calculations have been shown to yield good approximations for CCS.
- 16 molecules with 5 different elements C, H, O, N, and F were chosen for the optimization. These molecules have been previously used to optimize existing Lennard-Jones potentials using an ion-quadrupole potential in N<sub>2</sub>,<sup>31</sup> so they are excellent candidates for the study.
- A quadratic function  $\mathcal{F}$  has been chosen to optimize the Lennard-Jones potentials when comparing numerical and experimental CCSs. This function will yield zero at its minimization value.
- The numerical CCSs were obtained using *IMoS* in N<sub>2</sub> with a 4-6-12 potential while ignoring the ion-quadrupole effect. It is hypothesized that the effect of the ion-quadrupole potential together with the vibrations, translation and rotation of atoms can all be embedded in well-optimized L-J potentials.
- Because of the difficulty of interpreting a general optimization method with multiple unknowns, a *ceteris paribus* assumption was incorporated to understand the causation of each of the different parameters and each of the elements. This requires the method to be iterative.

- Surface mapping of the optimization function using several thousands of CCS calculations was obtained. It is revealed that a line of global minimums exists for most of the elements studied, and an exponential relation between the intercept and the potential well-depth can be determined. The reason is that the effect of increasing the potential well-depth can be equilibrated by a proper reduction of the intercept. A choice must be made for the L-J pair. To the extent that was possible, the Lorentz-Berthelot combination rule was used to establish a value for the intercept.
- To check whether different iteration routines could lead to different final results, the order of the elements carbon and hydrogen was altered.

The following conclusions have been obtained:

- New optimized Lennard-Jones parameters have been obtained for C, H, O, N, and F in N<sub>2</sub>. The resulting numerical CCSs are mostly within 1% of their experimental counterpart.
- The ion-quadrupole potential has not been used in the numerical calculations. It is proven here that the effect of this potential together with other effects such as atom vibrations and diffuse collisions can all be embedded into the L-J parameters without significantly affecting the CCS. Avoiding such calculations leads to a large gain in computational efficiency. Although not shown here, there is a possibility that the hydrogen effect can also be embedded into the influence of heavier atoms although this simplification should be studied in detail.
- While a set of L-J parameters were selected as the optimization candidate, other sets with largely different values could have been selected while providing accurate CCSs. See the [supplementary material](#) for an example.
- DFT geometric structures are excellent candidates to be used in CCS calculations of rigid and semi-rigid *small* molecules. However, it is recommended that the L-J parameters calculated for these molecules are not used in conjunction with molecular dynamics calculations or with very flexible molecules unless previously tested. These L-J parameters should not also be used in conjunction with ion-quadrupole potential methods in N<sub>2</sub>.

- To correctly optimize L-J parameters, it is important that elements in a specific molecule of interest have a strong weight in the CCS calculation (i.e., they are on or very near the surface). Finding appropriate molecules for a specific atom will lead to better defined parameters.

## SUPPLEMENTARY MATERIAL

See [supplementary material](#) for (1) an L-J gradient descent optimization example of carbon and hydrogen using 5 molecules, (2) an example of the input and output of *IMoS*, and (3) coordinates of DFT calculations used for the optimization of the parameters.

## ACKNOWLEDGMENTS

The authors of the paper would like to thank Vaibhav Shrivastav for providing the schematic in Fig. 1. The authors would also like to thank Iain Campuzano and Matt Bush for the procurement of the experimental data used in the optimization. This research was supported in part by Lilly Endowment, Inc., through its support for the Indiana University Pervasive Technology Institute, and in part by the Indiana METACyt Initiative. The Indiana METACyt Initiative at IU was also supported in part by Lilly Endowment, Inc.

- 1 E. A. Mason and E. W. McDaniel, *Transport Properties of Ions in Gases* (John Wiley & Sons, New York, 1988).
- 2 A. A. Shvartsburg and M. F. Jarrold, "An exact hard-spheres scattering model for the mobilities of polyatomic ions," *Chem. Phys. Lett.* **261**, 86–91 (1996).
- 3 M. F. Mesle, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz, and M. F. Jarrold, "Structural information from ion mobility measurements: Effects of the long-range potential," *J. Phys. Chem. A* **101**, 968 (1997); **100**, 16082 (1996).
- 4 Z. G. Li and H. Wang, "Drag force, diffusion coefficient, and electric mobility of small particles. I. Theory applicable to the free-molecule regime," *Phys. Rev. E* **68**, 061206 (2003).
- 5 J. C. May, C. R. Goodwin, N. M. Lareau, K. L. Leaptrout, C. B. Morris, R. T. Kurulugama, A. Mordehai, C. Klein, W. Barry, E. Darland, G. Overney, K. Imatani, G. C. Stafford, J. C. Fjeldsted, and J. A. McLean, "Conformational ordering of biomolecules in the gas phase: Nitrogen collision cross sections measured on a prototype high resolution drift tube ion mobility-mass spectrometer," *Anal. Chem.* **86**, 2107–2116 (2014).
- 6 K. L. Davidson and M. F. Bush, "Effects of drift gas selection on the ambient-temperature, ion mobility mass spectrometry analysis of amino acids," *Anal. Chem.* **89**, 2017–2023 (2017).
- 7 C. Bleiholder, N. R. Johnson, S. Contreras, T. Wyttenbach, and M. T. Bowers, "Molecular structures and ion mobility cross sections: Analysis of the effects of He and N<sub>2</sub> buffer gas," *Anal. Chem.* **87**, 7196–7203 (2015).
- 8 R. T. Kurulugama, E. Darland, F. Kuhlmann, G. Stafford, and J. Fjeldsted, "Evaluation of drift gas selection in complex sample analyses using a high performance drift tube ion mobility-QTOF mass spectrometer," *Analyst* **14**, 6834–6844 (2015).
- 9 C. J. Hogan, B. T. Ruotolo, C. V. Robinson, and J. Fernandez de la Mora, "Tandem differential mobility analysis-mass spectrometry reveals partial gas-phase collapse of the GroEL complex," *J. Phys. Chem. B* **115**, 3614–3621 (2011).
- 10 J. F. de la Mora, "Free-molecule mobility of polyhedra and other convex hard-bodies," *J. Aerosol Sci.* **33**, 477–489 (2002).
- 11 R. A. Millikan, "The general law of fall of a small spherical body through a gas, and its bearing upon the nature of molecular reflection from surfaces," *Phys. Rev.* **22**, 1–23 (1923).
- 12 P. S. Epstein, "On the resistance experienced by spheres in their motion through gases," *Phys. Rev.* **23**, 710 (1924).
- 13 C. Larriba, C. J. Hogan, M. Attoui, R. Borrajo, J. F. Garcia, and J. F. de la Mora, "The mobility-volume relationship below 3.0 nm examined by

- tandem mobility-mass measurement," *Aerosol Sci. Technol.* **45**, 453–467 (2011).
- 14 J. Fernandez-Garcia and J. F. de la Mora, "Electrical mobilities of multiply charged ionic-liquid nanodrops in air and carbon dioxide over a wide temperature range: Influence of ion-induced dipole interactions," *Phys. Chem. Chem. Phys.* **16**, 20500–20513 (2014).
- 15 H. Kim, H. I. Kim, P. V. Johnson, L. W. Beegle, J. L. Beauchamp, W. A. Goddard, and I. Kanik, "Experimental and theoretical investigation into the correlation between mass and ion mobility for choline and other ammonium cations in N<sub>2</sub>," *Anal. Chem.* **80**, 1928–1936 (2008).
- 16 H. I. Kim, H. Kim, E. S. Pang, E. K. Ryu, L. W. Beegle, J. A. Loo, W. A. Goddard, and I. Kanik, "Structural characterization of unsaturated phosphatidylcholines using traveling wave ion mobility spectrometry," *Anal. Chem.* **81**, 8289–8297 (2009).
- 17 C. Larriba and C. J. Hogan, "Free molecular collision cross section calculation methods for nanoparticles and complex ions with energy accommodation," *J. Comput. Phys.* **251**, 344–363 (2013).
- 18 C. Larriba and C. J. Hogan, "Ion mobilities in diatomic gases: Measurement versus prediction with non-specular scattering models," *J. Phys. Chem. A* **117**, 3887–3901 (2013).
- 19 J. Fernandez-Garcia and J. F. de la Mora, "Measuring the effect of ion-induced drift-gas polarization on the electrical mobilities of multiply-charged ionic liquid nanodrops in air," *J. Am. Soc. Mass. Spectrom.* **24**, 1872–1889 (2013).
- 20 C. Larriba and J. F. de la Mora, "The gas phase structure of coulombically stretched polyethylene glycol ions," *J. Phys. Chem. B* **116**, 593–598 (2012).
- 21 P. Garcaybarra and D. E. Rosner, "Thermophoretic properties of nonspherical particles and large molecules," *AIChE J.* **35**, 139–147 (1989).
- 22 D. W. Mackowski, "Monte Carlo simulation of hydrodynamic drag and thermophoresis of fractal aggregates of spheres in the free-molecule flow regime," *J. Aerosol Sci.* **37**, 242–259 (2006).
- 23 C. Larriba-Andaluz, J. Fernandez-Garcia, M. A. Ewing, C. J. Hogan, and D. E. Clemmer, "Gas molecule scattering and ion mobility measurements for organic macro-ions in He versus N<sub>2</sub> environments," *Phys. Chem. Chem. Phys.* **17**, 15019–15029 (2015).
- 24 C. Laphorn, F. S. Pullen, B. Z. Chowdhry, P. Wright, G. L. Perkins, and Y. Heredia, "How useful is molecular modelling in combination with ion mobility mass spectrometry for 'small molecule' ion mobility collision cross-sections?," *Analyst* **14**, 6814–6823 (2015).
- 25 N. L. Zakharova, C. L. Crawford, B. C. Hauck, J. K. Quinton, W. F. Seims, H. H. Hill, and A. E. Clark, "An assessment of computational methods for obtaining structural information of moderately flexible biomolecules from ion mobility spectrometry," *J. Am. Soc. Mass Spectrom.* **23**, 792–805 (2012).
- 26 J. Boschmans, S. Jacobs, J. P. Williams, M. Palmer, K. Richardson, K. Giles, C. Laphorn, W. A. Herrebout, F. Lemiére, and F. Sobott, "Combining density functional theory (DFT) and collision cross-section (CCS) calculations to analyze the gas-phase behaviour of small molecules and their protonation site isomers," *Analyst* **141**, 4044–4054 (2016).
- 27 C. Laphorn, T. J. Dines, B. Z. Chowdhry, G. L. Perkins, and F. S. Pullen, "Can ion mobility mass spectrometry and density functional theory help elucidate protonation sites in 'small' molecules?," *Rapid Commun. Mass Spectrom.* **27**, 2399–2410 (2013).
- 28 F. Furche, R. Ahlrichs, P. Weis, C. Jacob, S. Gilb, T. Bierweiler, and M. M. Kappes, "The structures of small gold cluster anions as determined by a combination of ion mobility measurements and density functional calculations," *J. Chem. Phys.* **117**, 6982–6990 (2002).
- 29 S. Gilb, P. Weis, F. Furche, R. Ahlrichs, and M. M. Kappes, "Structures of small gold cluster cations (Au<sup>+</sup><sub>n</sub>, n < 14): Ion mobility measurements versus density functional calculations," *J. Chem. Phys.* **116**, 4094–4101 (2002).
- 30 H. Ouyang, C. Larriba-Andaluz, D. R. Oberreit, and C. J. Hogan, "The collision cross sections of iodide salt cluster ions in air via differential mobility analysis-mass spectrometry," *J. Am. Soc. Mass Spectrom.* **24**, 1833–1847 (2013).
- 31 I. Campuzano, M. F. Bush, C. V. Robinson, C. Beaumont, K. Richardson, H. Kim, and H. I. Kim, "Structural characterization of drug-like compounds by ion mobility mass spectrometry: Comparison of theoretical and experimentally derived nitrogen collision cross sections," *Anal. Chem.* **84**, 1026–1033 (2012).
- 32 I. Czerwinska, J. Far, C. Kune, C. Larriba-Andaluz, L. Delaude, and E. De Pauw, "Structural analysis of ruthenium-arene complexes using ion mobility mass spectrometry, collision-induced dissociation, and DFT," *Dalton Trans.* **45**, 6361–6370 (2016).

- 849 <sup>33</sup>P. Benigni, J. D. DeBord, C. J. Thompson, P. Gardinali, and F. Fernandez-  
850 Lima, "Increasing polyaromatic hydrocarbon (PAH) molecular cover-  
851 age during fossil oil analysis by combining gas chromatography and  
852 atmospheric-pressure laser ionization fourier transform ion cyclotron res-  
853 onance mass spectrometry (FT-ICR MS)," *Energy Fuel* **30**, 196–203  
(2016).
- 854 <sup>34</sup>C. K. Siu, Y. Z. Guo, I. S. Saminathan, A. C. Hopkinson, and K. W.  
855 M. Siu, "Optimization of parameters used in algorithms of ion-mobility cal-  
856 culation for conformational analyses," *J. Phys. Chem. B* **114**, 1204–1212  
(2010).
- 857 <sup>35</sup>J. W. Lee, K. L. Davidson, M. F. Bush, and H. I. Kim, "Collision cross  
858 sections and ion structures: Development of a general calculation method via  
859 high-quality ion mobility measurements and theoretical modeling," *Analyst*  
860 **142**, 4289–4298 (2017).
- 861 <sup>36</sup>T. W. Knapman, J. T. Berryman, I. Campuzano, S. A. Harris, and A.  
862 E. Ashcroft, "Considerations in experimental and theoretical collision cross-  
863 section measurements of small molecules using travelling wave ion mobil-  
864 ity spectrometry-mass spectrometry," *Int. J. Mass Spectrom.* **298**, 17–23  
(2010).
- 865 <sup>37</sup>A. A. Shvartsburg, G. C. Schatz, and M. F. Jarrold, "Mobilities of carbon  
866 cluster ions: Critical importance of the molecular attractive potential," *J.*  
867 *Chem. Phys.* **108**, 2416–2423 (1998).
- 868 <sup>38</sup>P. M. Lalli, Y. E. Corilo, M. Fasciotti, M. F. Riccio, G. F. Sa, R. J. Daroda,  
869 G. H. Souza, M. McCullagh, M. D. Bartberger, and M. N. Eberlin, "Base-  
870 line resolution of isomers by traveling wave ion mobility mass spectrometry:  
Investigating the effects of polarizable drift gases and ionic charge distribu-  
tion," *J. Mass Spectrom.* **48**, 989–997 (2013).
- 871 <sup>39</sup>C. Larriba-Andaluz, M. Nahin, and V. Shrivastav, "A contribution to the  
872 amaranthine quarrel between true and average electrical mobility in the free  
873 molecular regime," *Aerosol Sci. Technol.* **51**, 887–895 (2017).
- 874 <sup>40</sup>V. Shrivastav, M. Nahin, C. J. Hogan, and C. Larriba-Andaluz, "Bench-  
875 mark comparison for a multi-processing ion mobility calculator in the free  
876 molecular regime," *J. Am. Soc. Mass Spectrom.* **28**, 1540–1551 (2017).
- 877 <sup>41</sup>P. M. W. Gill, B. G. Johnson, J. A. Pople, and M. J. Frisch, "The performance  
878 of the Becke-Lee-Yang-Parr (B-Lyp) density functional theory with various  
879 basis-sets," *Chem. Phys. Lett.* **197**, 499–505 (1992).
- 880 <sup>42</sup>A. D. Becke, "A new mixing of Hartree-Fock and local density-functional  
881 theories," *J. Chem. Phys.* **98**, 1372–1377 (1993).
- 882 <sup>43</sup>J. P. Perdew, M. Emzerhof, and K. Burke, "Rationale for mixing exact  
883 exchange with density functional approximations," *J. Chem. Phys.* **105**,  
884 9982–9985 (1996).
- 885 <sup>44</sup>F. Cuadros, I. Cachadina, and W. Ahumada, "Determination of Lennard-  
886 Jones interaction parameters using a new procedure," *Mol. Eng.* **6**, 319–325  
(1996).
- 887 <sup>45</sup>R. W. Zwanzig, "High-temperature equation of state by a perturbation  
888 method. I. Nonpolar gases," *J. Chem. Phys.* **22**, 1420–1426 (1954).
- 889 <sup>46</sup>D. Berthelot, "le m<sup>me</sup> des gaz," *Comput. Rendus* **126**, 1703–1706  
890 (1898).
- 891 <sup>47</sup>H. Lorentz, "Ueber die anwendung des satzes vom virial in der kinetischen  
892 theorie der gase," *Ann. Phys.* **248**, 127–136 (1881).
- 893  
894  
895