

# Global Minimum Potential Energy Conformations of Small Molecules

COSTAS D. MARANAS AND CHRISTODOULOS A. FLOUDAS\*

*Department of Chemical Engineering, Princeton University, Princeton, N.J. 08544-5263*

**Abstract.** A global optimization algorithm is proposed for finding the global minimum potential energy conformations of small molecules. The minimization of the total potential energy is formulated on an independent set of internal coordinates involving only torsion (dihedral) angles. Analytical expressions for the Euclidean distances between non-bonded atoms, which are required for evaluating the individual pairwise potential terms, are obtained as functions of bond lengths, covalent bond angles, and torsion angles. A novel procedure for deriving convex lower bounding functions for the total potential energy function is also introduced. These underestimating functions satisfy a number of important theoretical properties. A global optimization algorithm is then proposed based on an efficient partitioning strategy which is guaranteed to attain  $\epsilon$ -convergence to the global minimum potential energy configuration of a molecule through the solution of a series of nonlinear convex optimization problems. Moreover, lower and upper bounds on the total finite number of required iterations are also provided. Finally, this global optimization approach is illustrated with a number of example problems.

**Keywords:** Molecular Conformations, Global Minimum Total Potential Energy, Consistent Convex Lower Bounding, Exhaustive Subdivision Process, Bound Improving Selection Operation

## 1. Introduction

The study of molecular conformations is as fascinating a subject as it is utterly complex. Although the basic building blocks of molecules, the atoms, remain virtually unchanged in different compounds the versatility of the ways that they can be combined and reconfigure the obtained atom chains results in many different spatial configurations for a given molecule. One of these configurations, the most stable one, is of particular importance because it dictates most of the properties of the molecule. This provides enough motivation to predict the most stable conformation of a molecule based solely on the energetics of the interactions between the atoms composing the molecule. Molecular mechanics is a widely used method designed to provide a priori accurate representations of structures and energies for molecules. It originated in 1946 when Hill [13] proposed that van der Waals interactions along with stretching and bending deformations can be used for expressing the potential energy of a molecule. Dostrovsky, Hughes and Ingold [9, 7] utilized this same basic principle in an effort to understand the rates at which various halides underwent the  $S_N2$  reaction. However, it was Westheimer and Mayer [34, 24] who first used molecular mechanics ideas successfully in explaining the rate of racemization of

---

\*AUTHOR TO WHOM ALL CORRESPONDENCE SHOULD BE ADDRESSED.

some optically active compounds. With the advent of computers during the 1950's molecular mechanics became one of the standard methods of structural chemistry.

The most accurate representation of the potential energy of a molecule is the *ab initio* quantum mechanical approach. Quantum mechanical calculations, however, for all but the simplest molecules such as  $H_2$  or  $HCl$  are not feasible because of the large associated computational effort. As a result, tractable potential energy expressions had to be derived which adequately captured the energy contributions resulting from various types of atom interactions. In doing so, trade-offs between maximizing the accuracy of the potential energy expression and at the same time minimizing the computational effort to evaluate these functions had to be addressed.

In quantum mechanics the Born–Oppenheimer approximation is routinely used. It states that the Schrödinger equation for the molecule can be separated into a part describing the motions of the electrons and a part that describes the motion of the nuclei; and thus these two sets of motions can be studied independently. This means that the potential energy of a molecule in the *ground electronic state* is a function of only the nuclear positions. This defines a potential energy multi-dimensional hypersurface in the coordinate set of the nuclear positions. In molecular mechanics this hypersurface is simply called potential energy surface and encompasses the effect on the potential energy of all possible conformations that a molecule can resume. Any point of this surface corresponds to a different *conformation* of the molecule and local minimum points on this surface are referred to as *conformers*. The systematic identification of the global minimum point on this surface is the focus of this work. Experimental evidence [28] shows that in the great majority of cases this total potential energy global minimum point corresponds to the most stable conformation of the molecule. However, there exists examples where this is not true due to the interplay of rotational–vibrational motions [25].

Molecular mechanics calculations employ an empirically derived set of potential energy contributions for approximating the Born–Oppenheimer surface [3]. This set of potential energy contributions, called the force field, contains adjustable parameters that are selected in a such a way as to provide the best possible agreement with experimental data. The main assumption introduced in molecular mechanics is that every parameter is associated with a *specific interaction* rather than a specific molecule. These parameters can be bond lengths; covalent bond angles; bond stretching, bending, or rotating constants; non-bonded atom interaction constants, etc. Thus, whenever a specific interaction is present, the same value for the parameter can be used even if this interaction occurs in different molecules [14]. Although it is not possible to prove the validity of this assumption, experimental results provide sufficient evidence that it is a reasonable assumption in most cases. Molecular mechanics force fields include two-body, three-body or even n-body interaction terms.

$$V = \sum_{i>j} V_{ij} + \sum_{i>j>k} V_{ijk} + \sum_{i>j>k>l} V_{ijkl} + \dots$$

It can be seen that the total potential energy  $V$  is not simply the sum of all pair interactions; it may involve energy terms which depend on the position and prop-

erties of three or even more atoms due to the fact that the charges on the particles are in general polarized. However, for nonpolar atoms these terms can be neglected with little effect on the total potential energy [14]. This means that for nonpolar molecules the total potential energy expression is given approximately by,

$$V = \sum_{i>j} V_{ij}(B) + \sum_{i>j} V_{ij}(NB)$$

where B indicates bonded atom interactions and NB nonbonded atom interactions. In most conformational calculations the bonded atom interactions are assumed to be constant and independent of the actual conformation and thus they can be omitted since they do not contribute to the configuration of the molecule. Also, in general only *pairwise additive terms* are significant in conformational calculations although higher order terms can become important under certain conditions.

The general behavior of two-body potentials is very well studied. The force field is repulsive due to coulombic nuclear-nuclear and electron-electron interactions and attractive due to electron-nuclear interactions. In practice the pairwise potential interactions are usually represented by Lennard-Jones, Buckingham's, or Kihara's classical empirical potential functions:

### 1. Lennard-Jones 6-12 Potential Function

$$v_{ij} = \frac{B_{ij}}{r^{12}} - \frac{A_{ij}}{r^6}$$

The attractive term  $-A_{ij}/r^6$  is rigorously derived for a pair of identical spherically symmetrical and chemically saturated molecules, and  $B_{ij}/r^{12}$  is an approximation for the repulsive component for which no rigorously derived expression exists.

### 2. Buckingham Potential Function

$$v_{ij} = A_{ij} \exp(B_{ij} r_{ij}) + \frac{C_{ij}}{r_{ij}^6}$$

The Buckingham potential function differs from the Lennard-Jones function only in the form of the repulsive term. A two-parameter exponential function is used which in principle should be more specific in describing the repulsive interaction because of the additional parameter.

### 3. Kihara Potential Function

$$v_{ij} = 4\epsilon_o \left[ \left( \frac{1-\gamma}{r_{ij}/\sigma - \gamma} \right)^{12} - \left( \frac{1-\gamma}{r_{ij}/\sigma - \gamma} \right)^6 \right]$$

With the help of the parameters  $\gamma, \sigma$  Kihara's potential introduces an effective core and a shape dependence.

Before solving the resulting energy minimization problem it is necessary to decide whether or not the energy minimization will be performed on an independent set of internal coordinates (all bond lengths, covalent bond angles, and torsion angles) or on the Cartesian coordinates. Because it is easier to calculate the internal coordinates which are needed for evaluating the potential function from Cartesian coordinates rather than from an independent set of internal coordinates, most minimization methods use Cartesian coordinates. A number of methods have been proposed for finding the most stable conformation of a molecule through the identification of the global minimum point of the potential energy surface. All methods attempt to locate this point by tracing paths on the potential energy surface conjecturing that some of them will converge to the global minimum point. There are two main groups of methods; simulation type methods and gradient methods. Simulation type methods, including Monte-Carlo minimization and simulated annealing share the problem of selecting good strategies for "temperature" reduction dependence, optimal step size selection, and efficient random generator algorithms for generating the random walk. Gradient type methods are the most widespread methods for potential energy minimization of molecules. They can be divided into first derivative and second derivative minimization techniques. First derivative techniques (steepest descent) [35] follow a path defined by the steepest descent direction at every point. They perform satisfactorily only if all first derivatives are of the same order of magnitude, otherwise scaling problems cause oscillations. It has been also reported [2] that torsion angles are often not well minimized with these methods. Second derivative techniques, although more complicated, are far superior to first order methods. By utilizing second order derivative information calculated either analytically [1, 33, 29] or numerically [16], improved convergence rates are achieved.

It is important to note that most energy minimization procedures do not locate energy minima, but rather stationary points which occur when all first order derivatives are equal to zero. Therefore such procedures may converge to a saddlepoint or even a energy maximum on the potential energy surface rather than to an energy minimum. For example, in the case of cyclohexane the boat conformation ( $C_{2v}$ ) or the conformation that has all six carbon atoms on a plane ( $D_{6h}$ ) may erroneously appear as energy minima. However, the main limitation shared by all the aforementioned methods is that unless there is a single potential well, the obtained minimum energy conformation depends heavily on the supplied initial conformation. This is not surprising because all currently available methods are local optimization methods guaranteed to find a local minimum at best. This is why in practice many trial geometries serve as initial points for the employed optimization method. These geometries are usually chosen from Dreiding models, or other similar considerations and thus there is no guarantee that important conformations are not overlooked. Therefore, unless a systematic method capable of always converging to the global minimum potential energy independent of the initial conformation is employed, the obtained minimum energy conformation will be limited by which initial conformations seemed appropriate to the researcher [3]. The need for a method that can

guarantee convergence to the global minimum potential energy conformation is the motivation for this initial effort to introduce such a method for small molecules interacting with relatively simple force fields.

## 2. Problem Discussion

The problem to be addressed in this work can be stated simply as follows:

*“Given the connectivity of the atoms in a molecule and the force field according to which they interact, find the molecular conformation(s) in the three-dimensional Euclidean space involving the global minimum total potential energy”*

The simplifications employed herein are as follows:

1. The molecular mechanical approximation of the Born–Oppenheimer surface is adopted.
2. Only pairwise, additive, two–body interaction terms are considered.
3. Covalent bond lengths and angles are assumed to remain at their equilibrium values.
4. The expression representing the pairwise non–bonded atom potential interactions is assumed to be a function of only the Euclidean distance between the interacting atoms.

Approximation (1) is well established in the study of molecular conformations providing tractable expressions for the total potential energy of the molecule. Simplification (2) is valid for nonpolar molecules whose most stable conformations are sought in this work. Assumption (3) is adopted for the sake of convenience and it is fairly accurate in most cases because covalent bond lengths and angles do not deform significantly from their equilibrium values without substantial increase in the potential energy of the molecule. Finally, assumption (4) is made routinely in the field of molecular mechanics without significant loss of accuracy.

Under the aforementioned simplifications the expression for the total potential energy  $V$  of a molecule involves only the sum of a number of pairwise potential interaction terms. Each pairwise potential term is a function of only the Euclidean distances  $r_{ij}$  between the interacting atoms which are directly related to the Cartesian coordinates  $x_i, y_i, z_i$  of the atoms forming the molecule. Therefore,  $V$  can be fully represented in the coordinate space  $x_i, y_i, z_i$  of the atomic coordinates. However, a number of equality constraints must be added in the formulation to reflect the fact that all covalent bond lengths and covalent bond angles are assumed to be fixed at their equilibrium values. After defining  $\mathcal{B}$  to be the set of bonded atoms

and  $\mathcal{NB}$  the set of non-bonded atoms,  $V$  can be formulated in the Cartesian atomic coordinate space as follows:

$$V = \sum_{\substack{i < j \\ ij \in \mathcal{NB}}} v(r_{ij})$$

where

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}, \quad \forall ij \in \mathcal{NB}$$

subject to

$$r_{ij}^o = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}, \quad \forall ij \in \mathcal{B}$$

$$\theta_{ijk}^o = -\frac{(x_i - x_j)(x_j - x_k) + (y_i - y_j)(y_j - y_k) + (z_i - z_j)(z_j - z_k)}{r_{ij}^o r_{jk}^o},$$

$$\forall ijk \text{ such that } ij, jk \in \mathcal{B}$$

$$\text{and} \quad x_1 = y_1 = z_1 = y_2 = z_2 = z_3 = 0$$

where  $V$  is the total potential energy of the molecule;  $v(r_{ij})$  is the pairwise potential expression;  $r_{ij}$  is the Euclidean distance between nonbonded atom  $i$  with atom  $j$ ;  $r_{ij}^o, \theta_{ijk}^o$  are the fixed covalent bond lengths and angles respectively; and  $x_i, y_i, z_i$  are the Cartesian coordinates. Note that in the summation  $i < j$  so that we avoid doublecounting pair interactions and the interaction of an atom with itself. Furthermore, by imposing  $x_1 = y_1 = z_1 = y_2 = z_2 = z_3 = 0$  we eliminate the translational and rotational degrees of freedom of the molecule. This means that the total number of degrees of freedom are:

$$F = (3N - 6) - N_{bonds} - N_{angles}$$

where  $N$  is the total number of atoms,  $N_{bonds}$  the total number of covalent bonds in the molecule, and  $N_{angles}$  the total number of independent covalent bond angles. Note that only five out of six covalent bond angles are independent for tetravalent atoms.

As it has been discussed in the introduction, the use of Cartesian coordinates in local optimization techniques greatly facilitates the calculation of internal coordinates. However, by employing Cartesian coordinates we introduce a number of complex, highly nonlinear equality constraints, one for each  $r_{ij}^o, \theta_{ijk}^o$ , which do not allow the application of a global optimization algorithm. On the other hand, because the covalent bond lengths and angles are assumed to be constant, the set of independent internal coordinates reduces to a set of independent torsion angles. One simple possibility is to select one torsion angle per rotational degree of freedom. In polycyclic molecules, however, it may become more involved to define a set of independent internal coordinates because the bond and torsion angles are correlated. In principle, the independent set of torsion angles is rather small which is a considerable advantage for global optimization algorithms where the number of variables is typically the bottleneck of the computational effort. Accordingly, in this work independent internal coordinates rather than Cartesian coordinates are

employed. The total potential energy can then be written in the set of independent internal coordinates as follows:

$$V = \sum_{\substack{i < j \\ ij \in \mathcal{NB}}} v(r_{ij}(t_k))$$

where  $t_k \in$  set of independent internal coordinates,  $k = 1, \dots, K$

Note that the above formulation does not involve any equality constraints; however, the functionality between  $r_{ij}$  and the set of independent internal variables (independent torsion angles) has not been established yet. In the next section, an efficient procedure for obtaining explicit expressions for  $r_{ij}$  as functions of independent internal parameters is introduced.

### 3. Problem Formulation in Internal Coordinates

Although it is quite straightforward to derive expressions for the  $r_{ij}$ 's in the Cartesian coordinate set, unfortunately this is not the case when  $r_{ij}$  must be expressed as a function of an independent set of internal coordinates. In doing so it is first necessary to establish a connecting path, formed by a sequence of covalent bonds, between every atom  $i$  and atom  $j$ . For all pairs of atoms  $(i,j)$  in a molecule there is always an acyclic chain of atoms, connected by covalent bonds, which links atom  $i$  with atom  $j$ , because the connectivity of atoms in a molecule always defines a *undirected connected graph* [5]. This introduces a natural way of *partitioning* the set of nonbonded  $r_{ij}$ 's according to the minimum number of atoms participating in an acyclic chain connecting atom  $i$  with atom  $j$ . The aforementioned partitioning of  $r_{ij}$ 's can be realized with the definition of the following sets:

$$\begin{aligned} \mathcal{B} &= \{(i, j) : i, j \text{ bonded atoms}\} \\ \mathcal{NB} &= \{(i, j) : i, j \text{ nonbonded atoms}\} \\ \mathcal{B}^2 &= \{(i, j) : i < j, \text{ and } (i, j) \in \mathcal{B}\} \\ \mathcal{B}^3 &= \{(i, j, k) : i < k, \text{ and } (i, j) \in \mathcal{B}, (j, k) \in \mathcal{B}\} \\ \mathcal{B}^4 &= \{(i, j, k, l) : i < l, \text{ and } (i, j) \in \mathcal{B}, (j, k) \in \mathcal{B}, (k, l) \in \mathcal{B}, \\ &\quad \text{where } (i, j, k, l) \text{ acyclic chain}\} \end{aligned}$$

etc.

Note that  $\mathcal{B}$  is the set of pairs of atoms which are connected with covalent bonds and  $\mathcal{NB}$  is the set of pairs of atoms which are not. Also,  $\mathcal{B}^2, \mathcal{B}^3, \mathcal{B}^4$ , correspond to *directed* sets of chains of covalently bonded atoms involving two, three, and four atoms respectively. By generalizing the above analysis, the directed set of chains of  $m$  covalently bonded atoms is defined as,

$$\mathcal{B}^m = \{(i_1, i_2, \dots, i_m) : i_1 < i_m, \text{ and } (i_1, i_2) \in \mathcal{B}, (i_2, i_3) \in \mathcal{B}, \dots, (i_{m-1}, i_m) \in \mathcal{B}, \text{ where } (i_1, i_2, \dots, i_m) \text{ acyclic chain}\}$$

If only a single atom (2) is between atoms (1) and (3) such that  $(1, 2, 3) \in \mathcal{B}^3$ , then it is well known from trigonometry that,

$$r_{13}^2 = r_{12}^2 + r_{23}^2 - 2r_{12}r_{23} \cos(\theta_{123})$$

where  $\theta_{123}$  is the angle formed by covalent bonds (1-2) and (2-3). However, when more than one atom is required to connect atoms  $i$  and  $j$  the derivation of  $r_{ij}$  is much more involved.

Let us now consider three atoms, numbered (1), (2), and (3) such that atom (1) is bonded with atom (2), and atom (2) is bonded with atom (3). The bond lengths  $r_{12}$  and  $r_{23}$  as well as the covalent bond angle  $\theta_{123}$  are given. First, we select a base coordinate system defined by the positions of the first three atoms, (1), (2), and (3) such that the origin is chosen at atom (1), the negative  $x$ -axis passes through atom (2), and atom (3) lies in the first or second quadrant of the  $xy$  plane (See Figure 1). It is easy to show that the Cartesian coordinates of these three atoms are:

$$\begin{pmatrix} x_1 = 0, & y_1 = 0, & z_1 = 0 \\ x_2 = -r_{12}, & y_2 = 0, & z_2 = 0 \\ x_3 = r_{23} \cos(\theta_{123}) - r_{12}, & y_3 = r_{23} \sin(\theta_{123}), & z_3 = 0 \end{pmatrix}$$

The position of an additional atom (4), bonded to atom (3), can be fully specified if the length of the bond  $r_{34}$ , the bond angle  $\theta_{234}$ , and the dihedral (torsion) angle  $\phi_{1234}$  are known.  $\phi_{1234}$  is the angle between the plane defined by atoms (1), (2), (3) and the plane defined by atoms (2), (3), (4). The coordinates of atom (4) after rotational transformation and a translation are given by [27],

$$\begin{bmatrix} x_4 \\ y_4 \\ z_4 \\ 1 \end{bmatrix} = \mathbf{B}_2 \mathbf{B}_3 \mathbf{B}_4 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

where  $\mathbf{B}_2$ ,  $\mathbf{B}_3$  and  $\mathbf{B}_4$  are  $4 \times 4$  transformation matrices.

$$\mathbf{B}_2 = \begin{bmatrix} -1 & 0 & 0 & -r_{12} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{B}_3 = \begin{bmatrix} -\cos(\theta_{123}) & -\sin(\theta_{123}) & 0 & -r_{23} \cos(\theta_{123}) \\ \sin(\theta_{123}) & -\cos(\theta_{123}) & 0 & r_{23} \sin(\theta_{123}) \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and  $\mathbf{B}_4$  is equal to,

$$\begin{bmatrix} -\cos(\theta_{234}) & -\sin(\theta_{234}) & 0 & -r_{34} \cos(\theta_{234}) \\ \sin(\theta_{234}) \cos(\phi_{1234}) & -\cos(\theta_{234}) \cos(\phi_{1234}) & -\sin(\phi_{1234}) & r_{34} \sin(\theta_{234}) \cos(\phi_{1234}) \\ \sin(\theta_{234}) \sin(\phi_{1234}) & -\cos(\theta_{234}) \sin(\phi_{1234}) & \cos(\phi_{1234}) & r_{34} \sin(\theta_{234}) \sin(\phi_{1234}) \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

By repetitively placing extra atoms in the initial four atom chain, an  $m$ -chain is formed composed by atoms connected sequentially with covalent bonds. By generalizing the presented analysis the coordinates of the  $m^{th}$  atom in the coordinate system that the first three atoms define are,

$$\begin{bmatrix} x_m \\ y_m \\ z_m \\ 1 \end{bmatrix} = \mathbf{B}_2 \mathbf{B}_3 \mathbf{B}_4 \dots \mathbf{B}_m \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

where  $\mathbf{B}_m$  is defined as,

$$\mathbf{B}_m = \begin{bmatrix} -\cos(\theta_m) & -\sin(\theta_m) & 0 & -r_m \cos(\theta_m) \\ \sin(\theta_m) \cos(\phi_m) & -\cos(\theta_m) \cos(\phi_m) & -\sin(\phi_m) & r_m \sin(\theta_m) \cos(\phi_m) \\ \sin(\theta_m) \sin(\phi_m) & -\cos(\theta_m) \sin(\phi_m) & \cos(\phi_m) & r_m \sin(\theta_m) \sin(\phi_m) \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

with the following simplified notation,

$$\begin{aligned} r_{m-1,m} &\leftarrow r_m \\ \theta_{m-2,m-1,m} &\leftarrow \theta_m \\ \phi_{m-3,m-2,m-1,m} &\leftarrow \phi_m \end{aligned}$$

Based on the above analysis, explicit expressions for the Cartesian coordinates  $x_m, y_m, z_m$  of the  $m^{th}$  atom in a  $m$ -atom chain can be obtained as functions of the bond lengths, covalent bond angles, and torsion angles. After recognizing that  $x_1 = y_1 = z_1 = 0$ , the Euclidean distance  $r_{1m}$  can then be easily calculated,

$$r_{1m}^2 = x_m^2 + y_m^2 + z_m^2$$

After expanding the squared terms, the expression for  $r_{1m}^2$  becomes the sum of a very large number of terms involving the product of bond lengths with sines and/or cosines of covalent bond lengths and/or dihedral (torsion) angles. Due to the repetitive matrix operations, the resulting formulas lose any special structure which is necessary for developing a global optimization algorithm. This special structure, however, can be uncovered with the repetitive application of the trigonometric formula,

$$\sin(x)^2 + \cos(x)^2 = 1$$

through symbolic computations. After performing the aforementioned analysis, general expressions for the Euclidean distances between the end atoms in  $m$ -atom ( $m \leq 6$ ) chains are obtained. These expressions are given in Appendix A.

Based on the above analysis, the expression for the total potential can be partitioned as follows:

$$V = \sum_{j=4}^m \sum_{(i_1, \dots, i_j) \in \mathcal{B}^j} v(r_{i_1, i_j})$$

where  $m$  is the number of atoms in the longest chain connecting two interacting atoms;  $r_{i_1, i_j}$  is the expression derived earlier for the Euclidean distance between the two end atoms  $i_1$  and  $i_j$ ; and  $v(r_{i_1, i_j})$  is the pairwise potential interaction function. Note that  $j$  starts from ( $j = 4$ ) because there must be at least two atoms between, the interacting atoms before considering their contribution to the total potential energy of the molecule [3]. This is because for  $j = 2, 3$  the Euclidean distances between the end atoms  $i_1$  and  $i_j$  are functions of only the fixed covalent bond angles and covalent bond lengths. Presumably, the corresponding pairwise interaction terms are constant independent of the conformation of the molecule and therefore they can be neglected.

The main difficulty in finding the global minimum of  $V$  arises from the fact that it is a *complicated nonconvex* function. As it has been discussed earlier, any local optimization technique is likely to find only a local minimum at best and complete enumeration of all local minima is impossible. In the next section, a procedure is proposed for obtaining a convex lower bounding function of  $V$ .

#### 4. Convex Lower Bounding Function of $V$

In the previous section, explicit relations for the nonbonded atom distances have been derived as functions of the dihedral (torsion) angles  $\phi_{ijkl}$ . However, not all of these dihedral angles are independent. In fact, there are only as many independent dihedral angles as the number of rotational degrees of freedom. For example, for ethane-like molecules all dihedral angles can be expressed as functions of a *single* dihedral angle. For propane-like molecules two dihedral angles are required for fully specifying the conformation of the molecule. This means that all nonbonded distances  $r_{ij}$  and consequently the expression for the total potential energy  $V$  can be expressed as a function of only these independent dihedral angles, defined one for each rotational degree of freedom.

$$V = V(t_k), \quad k = 1, \dots, K$$

where  $t_k$ 's correspond to the  $K$  independent dihedral angles.

The next step in this analysis is to define the function  $\mathcal{L}$  by augmenting  $V$  as it has been proposed in [19, 20].

$$\mathcal{L} = V + \alpha \sum_{k=1}^K (t_k^L - t_k)(t_k^U - t_k)$$

where 
$$\alpha \geq \max \left\{ \max_{t_k^L \leq t_k \leq t_k^U} 0, \left( -\frac{1}{2} \lambda_k^V \right) \right\}$$

Note that  $t_k^L, t_k^U$  correspond to the lower and upper bounds on  $t_k$  which are typically set to  $t_k^L = 0$  and  $t_k^U = 2\pi$ . Also,  $\alpha$  is a nonnegative parameter which must be greater or equal to the negative one half of the minimum eigenvalue of  $V$  over  $t_k^L \leq t_k \leq t_k^U$ . Guidelines for estimating  $\alpha$  are provided in Appendix B. The effect of adding the term,

$$\alpha \sum_{k=1}^K (t_k^L - t_k)(t_k^U - t_k)$$

to  $V$  is to make  $\mathcal{L}$  convex by overpowering the nonconvexity characteristics of  $V$  with the addition of the term  $2\alpha$  to all of its eigenvalues,

$$\lambda_k^{\mathcal{L}} = \lambda_k^V + 2\alpha.$$

Here  $\lambda_k^{\mathcal{L}}, \lambda_k^V$  are the  $k^{\text{th}}$  eigenvalues of  $\mathcal{L}, V$  respectively. In the following, a number of important properties, which  $\mathcal{L}$  share, will be stated and proven. These properties will enable us to construct a global optimization algorithm for finding the global minimum of  $V$  in the space defined by the independent dihedral angles  $t_k$ .

**Property 1:** For every  $t_k \in [t_k^L, t_k^U]$ , we have  $\mathcal{L}(t_k) \leq V(t_k)$ .

**Proof:**

$$\begin{aligned} \forall t_k \in [t_k^L, t_k^U], \quad & \alpha \geq 0, \quad (t_k^L - t_k) \leq 0, \quad (t_k^U - t_k) \geq 0, \\ \implies & \alpha(t_k^L - t_k)(t_k^U - t_k) \leq 0 \\ \implies & \alpha \sum_{k=1}^K (t_k^L - t_k)(t_k^U - t_k) \leq 0 \\ \implies & \mathcal{L}(t_k) \leq V(t_k). \quad \square \end{aligned}$$

**Remark:**  $\mathcal{L}$  is always a valid underestimator of  $V$  inside the box constraints  $[t_k^L, t_k^U]$ .

**Property 2:** For every corner point in  $[t_k^L, t_k^U]$ , we have  $\mathcal{L} = V$ .

**Proof:**

Let  $t_k^c$  be a corner point of  $[t_k^L, t_k^U]$ ,  $k = 1, \dots, K$ .

$$\begin{aligned} \forall k = 1, \dots, K, \quad & (t_k^L - t_k^c) = 0 \quad \text{or} \quad (t_k^U - t_k^c) = 0, \\ \implies & \alpha(t_k^L - t_k^c)(t_k^U - t_k^c) = 0, \\ \implies & \sum_{k=1}^K \alpha(t_k^L - t_k^c)(t_k^U - t_k^c) = 0, \\ \implies & \mathcal{L}(t_k^c) = V(t_k^c) \quad \square \end{aligned}$$

**Remark:**  $\mathcal{L}$  matches  $V$  at all corner points.

**Property 3:**  $\mathcal{L}$  is convex in  $[t_k^L, t_k^U]$ .

**Proof:**

$\mathcal{L}$  is convex in  $[t_k^L, t_k^U]$  if and only if for every  $t_k \in [t_k^L, t_k^U]$ , all eigenvalues  $\lambda_k^{\mathcal{L}}(t_k)$  of  $\mathcal{L}$  are nonnegative. This is equivalent with requiring the minimum eigenvalue of  $\mathcal{L}$  over  $t_k$  to be nonnegative.

$$\min_{t_k^L \leq t_k \leq t_k^U} \lambda_k^{\mathcal{L}}(t_k) \geq 0.$$

After substituting  $\lambda_k^{\mathcal{L}}(t_k) = \lambda_k^V(t_k) + 2\alpha$

and  $\alpha \geq \max \left\{ 0, \max_{t_k^L \leq t_k \leq t_k^U} \left( -\frac{1}{2} \lambda_k^V(t_k) \right) \right\}$  in the previous inequality we have,

$$\min_{t_k^L \leq t_k \leq t_k^U} \left\{ \lambda_k^V(t_k) + \max \left[ 0, \max_{t_k^L \leq t_k \leq t_k^U} \left( -\lambda_k^V(t_k) \right) \right] \right\} \geq 0.$$

We can further simplify by combining the two max operators and transforming the resulting single max operator into a single min operator.

$$\min_{t_k^L \leq t_k \leq t_k^U} \lambda_k^V(t_k) - \min_{t_k^L \leq t_k \leq t_k^U} [0, \lambda_k^V(t_k)] \geq 0.$$

Finally, we combine the two min operators by considering the two cases for the sign of  $\lambda_k^V(t_k)$ . The resulting inequality is true for every  $t_k$  in  $[t_k^L, t_k^U]$ :

$$\max \left[ 0, \min_{t_k^L \leq t_k \leq t_k^U} \lambda_k^V(t_k) \right] \geq 0.$$

Therefore,  $\mathcal{L}(t_k)$  is convex for every  $t_k$  in  $[t_k^L, t_k^U]$ .  $\square$

**Remark:**  $\mathcal{L}$  is convex in  $[t_k^L, t_k^U]$  by construction.

**Property 4:** 
$$\max_{t_k^L \leq t_k \leq t_k^U} (V - \mathcal{L}) = \frac{1}{4} \alpha \sum_{k=1}^K (t_k^U - t_k^L)^2$$

**Proof:**

$$\begin{aligned} & \max_{t_k^L \leq t_k \leq t_k^U} (V - \mathcal{L}) \\ &= \max_{t_k^L \leq t_k \leq t_k^U} -\alpha \sum_{k=1}^K (t_k^L - t_k)(t_k^U - t_k) \\ &= \min_{t_k^L \leq t_k \leq t_k^U} \alpha \sum_{k=1}^K (t_k^L - t_k)(t_k^U - t_k) \end{aligned}$$

The last function is convex and separable in  $t_k$ . This means that if the single stationary point is inside the box constraints  $[t_k^L, t_k^U]$  then it corresponds to the single global minimum [18]. After applying the stationarity conditions we find that the single stationary point  $t_k^{stat}$  is indeed inside the box constraints,

$$t_k^{stat} = \frac{(t_k^L + t_k^U)}{2}, \quad k = 1, \dots, K,$$

and therefore it corresponds to the single global minimum. After substituting the expression for  $t_k^{stat}$  in the original function we have,

$$\max_{t_k^L \leq t_k \leq t_k^U} (V - \mathcal{L}) = \frac{1}{4} \alpha \sum_{k=1}^K (t_k^U - t_k^L)^2 \quad \square$$

**Remark:** The maximum separation between  $\mathcal{L}$  and  $V$  is proportional to  $\alpha$  and to the square of the diagonal of the current box constraints.

**Property 5:**

$$\forall t_k \in [t_k^{L1}, t_k^{U1}], \quad \forall [t_k^{L2}, t_k^{U2}] \text{ such that } [t_k^{L1}, t_k^{U1}] \subseteq [t_k^{L2}, t_k^{U2}],$$

$$\mathcal{L}(t_k, \alpha^1, t_k^{L1}, t_k^{U1}) \geq \mathcal{L}(t_k, \alpha^2, t_k^{L2}, t_k^{U2}),$$

after imposing  $\alpha^2 \geq \alpha^1$  in the selection process for  $\alpha^1, \alpha^2$ ,

$$\text{where } \alpha^1 \geq \max \left\{ 0, \max_{t_k^{L1} \leq t_k \leq t_k^{U1}} \left( -\frac{1}{2} \lambda_k^V(t_k) \right) \right\}$$

$$\text{and } \alpha^2 \geq \max \left\{ 0, \max_{t_k^{L2} \leq t_k \leq t_k^{U2}} \left( -\frac{1}{2} \lambda_k^V(t_k) \right) \right\}$$

**Proof:**

$$\begin{aligned} \forall t_k \in [t_k^{L1}, t_k^{U1}], \alpha^2 \geq \alpha^1 \geq 0, \\ (t_k^{L2} - t_k) \leq (t_k^{L1} - t_k) \leq 0, \\ (t_k^{U2} - t_k) \geq (t_k^{U1} - t_k) \geq 0 \\ \implies \alpha^1 \sum_{k=1}^K (t_k^{L1} - t_k)(t_k^{U1} - t_k) \geq \alpha^2 \sum_{k=1}^K (t_k^{L2} - t_k)(t_k^{U2} - t_k) \\ \implies \mathcal{L}(t_k, \alpha^1, t_k^{L1}, t_k^{U1}) \geq \mathcal{L}(t_k, \alpha^2, t_k^{L2}, t_k^{U2}) \quad \square \end{aligned}$$

**Remark:** The underestimators constructed over supersets of the current set are always *less tight* than the underestimator constructed over the current box constraints for every point within the current box constraints.

In summary, in this section we have proven that  $\mathcal{L}$  is a *convex, lower bounding* function of  $V$ . Also,  $\mathcal{L}$  *matches*  $V$  at all corner points of the box constraints inside which it has been defined. Its maximum separation from  $V$  is *bounded* and *proportional* to  $\alpha$  and to the square of the diagonal of the box constraints. Finally, the values of different definitions of  $\mathcal{L}$  at any point, if  $\mathcal{L}$  is constructed over a tighter box of constraints each time, define a *nondecreasing* sequence. In Appendix C it is shown that function  $\mathcal{L}$  corresponds to a relaxed dual bound of the original function  $V$  [10, 30, 31, 11, 32]. In the next section, based on the aforementioned properties, a branch and bound type global optimization algorithm is introduced for solving the global minimization of  $V$ .

## 5. Global Optimization Algorithm

By utilizing the properties introduced in the previous section a global optimization algorithm is proposed for locating the global minimum potential energy  $V^*$  by constructing converging lower and upper bounds on  $V^*$ . It is clear that the value of  $V$  at any point  $t_k$  is an upper bound on  $V^*$ . Lower bounds on  $V^*$  within some box constraints are derived by utilizing Properties (1) and (3). From Property (1),  $\mathcal{L}$  is a lower bounding function of  $V$ . Therefore, its global minimum within some box constraints will always underestimate the global minimum of  $V$  within the same box constraints. However, from Property (3) we know that, unlike  $V$ ,  $\mathcal{L}$  is convex and therefore it involves a single global minimum which can be routinely found with any commercially available local optimization algorithm (e.g. MINOS 5.3 [21]). An upper bound on  $V^*$  can then be obtained by simply calculating  $V$  at the global minimum point of  $\mathcal{L}$ . From property (4) we know that the gap between the upper

bound and the lower bound will be at most,

$$(V^U - V^L) \leq \frac{1}{4}\alpha \sum_{k=1}^K (t_k^U - t_k^L)^2$$

where  $V^U$  and  $V^L$  are the upper and lower bounds on the global minimum  $V^*$  inside the current box constraints,  $[t_k^L, t_k^U]$ ,  $k = 1, \dots, K$ .

The next step, after establishing an upper and a lower bound on the global minimum, is to tighten them by using Property (5). Property (5) implies that the value of  $\mathcal{L}$  at every point, and therefore at its global minimum, is increased by restricting the box constraints within which it has been defined. Tighter box constraints can be realized by partitioning the rectangle that the initial box constraints define into a number of smaller rectangles. One way of partitioning is to successively divide the current rectangle in two subrectangles by halving on the middle point of the longest side of the initial rectangle (bisection). Presumably, at each iteration the lower bound of  $V^*$  is simply the minimum over all the minima of  $\mathcal{L}$  in every subrectangle composing the initial rectangle. Therefore, a straightforward (bound improving) way of tightening the lower bound  $V^L$  is to halve at each iteration, only the subrectangle responsible for the infimum of the minima of  $\mathcal{L}$  over all subrectangles, according to the rules discussed earlier. This procedure generates a *nondecreasing* sequence for the lower bound  $V^L$  of  $V^*$ . Furthermore, we construct a *nonincreasing* sequence for the upper bound  $V^U$  by selecting it to be the infimum over all the previously recorded upper bounds. Clearly, if the global minimum of  $\mathcal{L}$  in any subrectangle is greater than the current upper bound  $V^U$  we can safely ignore this subrectangle because the global minimum of  $V$  cannot be situated inside it (fathoming step).

Property (4) answers the question of how small these subrectangles must become before the upper and lower bounds of  $V$  inside these subrectangles are within  $\epsilon$ . If  $\delta$  is the diagonal of the subrectangle,

$$\delta = \sqrt{\sum_{k=1}^K (t_k^U - t_k^L)^2}$$

and  $\epsilon$  is the convergence tolerance, from property (4) we have the following condition for convergence,

$$\epsilon \geq \frac{1}{4}\alpha \sum_{k=1}^K (t_k^U - t_k^L)^2 = \frac{1}{4}\alpha\delta^2 \geq V^U - V^L$$

which means that if the diagonal  $\delta$  of a subrectangle is,

$$\delta < \sqrt{\frac{4\epsilon}{\alpha}}$$

then  $\epsilon$ -convergence to the global minimum of  $V$  has been achieved. It is interesting to note that the required for convergence value of  $\delta$  is proportional to the square root

of  $\epsilon$ . Therefore, if for example  $\epsilon$  is set to be 0.0001,  $\delta$  suffices to be proportional to 0.01. Note also, that  $\delta$  is inversely proportional to the squared root of the parameter  $\alpha$  reflecting the fact that the smaller the value of  $\alpha$  is, the faster the convergence rate becomes.

The basic steps of the proposed global optimization algorithm are summarized in the following section.

## 6. Steps of the Global Optimization Algorithm

### STEP 1 - Initialization

A convergence tolerance  $\epsilon$  is selected and the iteration counter  $Iter$  is set to one. Appropriate global bounds  $t_k^{LBD}, t_k^{UBD}$  on  $t_k$  are chosen and local bounds  $t_k^{L,Iter}, t_k^{U,Iter}$  for the first iteration are set to be equal to the global ones. Lower and upper bounds on the global minimum  $V^{LBD}, V^{UBD}$  are initialized and an initial current point  $t_k^{c,Iter}$  is selected.

### STEP 2 - Update of Upper Bound $V^{UBD}$

$V$  is calculated at the current point  $V^{c,Iter}$  and the upper bound  $V^{UBD}$  is updated as follows,

$$V^{UBD} = \min(V^{UBD}, V(t_k^{c,Iter}))$$

### STEP 3 - Partitioning of Current Rectangle

The current rectangle  $[t_k^{L,iter}, t_k^{U,Iter}]$ ,  $k = 1, \dots, K$  is partitioned into the following two rectangles ( $r = 1, 2$ ):

$$\left[ \begin{array}{cc} t_1^{L,Iter} & t_1^{U,Iter} \\ \vdots & \vdots \\ t_{l^{Iter}}^{L,Iter} & \frac{(t_{l^{Iter}}^{L,Iter} + t_{l^{Iter}}^{U,Iter})}{2} \\ \vdots & \vdots \\ t_K^{L,Iter} & t_K^{U,Iter} \end{array} \right], \quad \left[ \begin{array}{cc} t_1^{L,Iter} & t_1^{U,Iter} \\ \vdots & \vdots \\ \frac{(t_{l^{Iter}}^{L,Iter} + t_{l^{Iter}}^{U,Iter})}{2} & t_{l^{Iter}}^{U,Iter} \\ \vdots & \vdots \\ t_K^{L,Iter} & t_K^{U,Iter} \end{array} \right]$$

where  $l^{Iter}$  corresponds to the variable with the scaled longest side in the initial rectangle,

$$l^{Iter} = \arg \max_k \frac{(t_k^{U,Iter} - t_k^{L,Iter})}{(t_k^{UBD} - t_k^{LBD})}$$

**STEP 4 - Solution of Convex Problems in two Subrectangles**

Update the parameter  $\alpha$  and solve the following convex nonlinear optimization problem in both subrectangles ( $r = 1, 2$ ) by using a commercially available nonlinear solver (e.g. MINOS 5.3 [21]).

$$\min_{t_k} \mathcal{L} = \sum_{j=4}^m \sum_{(i_1, \dots, i_j) \in \mathcal{B}^j} v(r_{i_1, i_j}(t_k)) + \alpha \sum_{k=1}^K (t_k^{L, Iter} - t_k) (t_k^{U, Iter} - t_k)$$

If a solution  $L_{sol}^{r, Iter}$  is less than the current upper bound,

$$L_{sol}^{r, Iter} \leq V^{UBD}$$

then it is stored along with the value of the variable  $t_k$  at the solution point  $t_{k, sol}^{r, Iter}$ .

**STEP 5 - Update Iteration Counter  $Iter$  and Lower Bound  $V^{LBD}$** 

The iteration counter is increased by one,

$$Iter \leftarrow Iter + 1$$

and the lower bound  $V^{LBD}$  is updated to the minimum solution over the stored ones from previous iterations. Furthermore, the selected solution is erased from the stored set.

$$V^{LBD} = L_{sol}^{r', Iter'}$$

$$\text{where } L_{sol}^{r', Iter'} = \min_{r, I} L_{sol}^{r, I}, \quad r = 1, 2, \quad I = 1, \dots, Iter - 1.$$

**STEP 6 - Update Current Point  $t_k^{c, Iter}$  and Current Bounds  $t_k^{L, Iter}, t_k^{U, Iter}$  on  $t_k$** 

The current point is selected to be the solution point of the previously found minimum solution in **Step 5**,

$$t_k^{c, Iter} = t_{k, sol}^{r', Iter'}, \quad k = 1, \dots, K$$

and the current rectangle becomes the subrectangle containing the previously found solution,

$$\left[ t_k^{L, Iter}, t_k^{U, Iter} \right] = \begin{bmatrix} t_1^{L, Iter'} & t_1^{U, Iter'} \\ \vdots & \vdots \\ t_{l_{Iter'}}^{L, Iter'} & \frac{(t_{l_{Iter'}}^{L, Iter'} + t_{l_{Iter'}}^{U, Iter'})}{2} \\ \vdots & \vdots \\ t_K^{L, Iter'} & t_K^{U, Iter'} \end{bmatrix}, \quad \text{if } r' = 1$$

$$\left[ t_k^{L,Iter}, t_k^{U,Iter} \right] = \left[ \begin{array}{cc} t_1^{L,Iter'} & t_1^{U,Iter'} \\ \vdots & \vdots \\ \frac{(t_{l^{Iter'}}^{L,Iter'} + t_{l^{Iter'}}^{U,Iter'})}{2} & t_{l^{Iter'}}^{U,Iter'} \\ \vdots & \vdots \\ t_K^{L,Iter'} & t_K^{U,Iter'} \end{array} \right], \quad \text{if } r' = 2$$

**STEP 7 - Check for Convergence**

IF  $(V^{UBD} - V^{LBD}) > \epsilon$ , then return to **STEP 2**

Otherwise,  $\epsilon$ -convergence has been reached and the global minimum solution, and solution point are:

$$\begin{aligned} V^* &\leftarrow V^{U,Iter''} \\ t_k^* &\leftarrow t_k^{c,Iter''}, \quad k = 1, \dots, K \\ \text{where } Iter'' &= \arg \{ V^{U,I} = V^{UBD} \}, \quad I = 1, \dots, Iter. \end{aligned}$$

In the following section, a mathematical proof that the proposed global optimization algorithm converges to the the global minimum is given based on the analysis of standard global optimization algorithm presented in [15].

**7. Proof of Convergence to the Global Minimum**

Convergence properties of a global optimization algorithm depend on:

- (i) the limit behavior of the difference  $V^U - V^L$  for unfathomed successively refined partitions,
- (ii) the subdivision process of the current partitions, and
- (iii) the employed selection process of the partition(s) that have to be further refined.

In the following, a In the employed global optimization algorithm for every partition element  $(r, Iter)$  a lower bound  $V_{r,Iter}^L$  is obtained as the solution of the convex minimization of  $\mathcal{L}$  inside the box constraints of the current partition element and an upper bound  $V_{r,Iter}^U$  as the value of  $V$  at the minimum point of  $\mathcal{L}$ . The partition element involving the minimum lower bound  $V_{r,Iter}^L$  is selected for further refining according to the bisection rule and partition elements whose lower bound  $V_{r,Iter}^L$  is greater than the current upper bound  $V_{Iter}^U$  are fathomed. A sufficient condition for a global optimization algorithm to be convergent to the global minimum, stated in [15], requires that the bounding operation must be *consistent* and the selection operation *bound improving*.

**Definition 1:** A bounding operation is called **consistent** if at every step any unfathomed partition can be further refined, and if any infinitely decreasing sequence of successively refined partition elements satisfies,

$$\lim_{Iter \rightarrow \infty} (V_{Iter}^U - V_{r,Iter}^L) = 0,$$

where  $V_{r,Iter}^L$  is the lower bound of  $V$  inside the  $(r, Iter)$  partition element and  $V_{Iter}^U$  is the best upper bound at iteration  $Iter$  not necessarily occurring inside the  $(r, Iter)$  partition element.

Clearly, any partition element  $(r^o, Iter^o)$  that is not fathomed eventually becomes the one involving the minimum current lower bound  $V_{iter}^L$  and is further refined according to the bisection rule. Property (5) implies that the subdivisions  $(r, Iter)$  of this partition element  $(r^o, Iter^o)$  involve lower bounds  $V_{r,Iter}^L$  which are at least as tight as  $V_{r^o, Iter^o}^L$ .

$$\min_r V_{r,Iter}^L \geq V_{r^o, Iter^o}^L$$

which implies that any infinitely refined partition sequence involves infinitely non-decreasing lower bounds.

In practice, the requirement  $\lim_{Iter \rightarrow \infty} (V_{Iter}^U - V_{r,Iter}^L) = 0$  for any infinitely decreasing sequence of successively refined partition elements is difficult to verify because  $V_{Iter}^U$  is not necessarily attained at the partition element  $(r, Iter)$ . Therefore, in view of the inequality  $V_{r,Iter}^U \geq V_{Iter}^U \geq V_{r,Iter}^L$ , the following, at least equally strong condition, suffices to be shown:

$$\lim_{Iter \rightarrow \infty} (V_{r,Iter}^U - V_{r,Iter}^L) = 0.$$

From Property (4) we know that,

$$\frac{1}{4}\alpha\delta^2 \geq (V_{r,Iter}^U - V_{r,Iter}^L) \geq 0.$$

where  $\delta$  is the diagonal of the current partition element corresponding to a rectangle. Therefore, it suffices to show that

$$\lim_{Iter \rightarrow \infty} \frac{1}{4}\alpha\delta^2 = 0,$$

or equivalently that

$$\lim_{Iter \rightarrow \infty} \delta = 0.$$

This says that the employed subdivision process for any unfathomed infinitely decreasing sequence of successively refined partition elements is **exhaustive** [15].

**Definition 2:** A subdivision process is **exhaustive** if  $\lim_{Iter \rightarrow \infty} \delta = 0$  for all decreasing subsequences of partition elements generated by the subdivision.

**Theorem 1:** *The subdivision process of the employed global optimization algorithm is exhaustive.*

**Proof:** The employed subdivision process is the *bisection* where every partition element  $(r, Iter)$  is subdivided into two subrectangles  $r = 1, 2$  of equal volume by halving at the midpoint of the longest scaled side. Because all the sides of any partition element are scaled with the sides of the rectangle that the initial global constraints define, the scaled sides of the initial rectangle are all equal to one. Therefore, the condition of always subdividing along the longest side can be satisfied by simply subdividing first along the side  $k = 1$ , then along the side  $k = 2$ , etc. until the last side  $k = K$  is encountered when the subdivision starts again from  $k = 1$ . By partitioning in this orderly manner each side of every successively refined partition element is halved *exactly* once every  $K$  subdivisions. Consequently, after  $K$  subdivisions the diagonal of the resulting partition elements is one half the diagonal of the original partition element. Therefore, as the number of successive subdivisions of a partition element goes to infinity, the diagonals of the resulting partition elements go to zero. This implies that the employed subdivision process is *exhaustive*.  $\square$

**Theorem 2:** *The bounding operation of the employed global optimization algorithm is consistent.*

**Proof:** We have shown that the subdivision process is *exhaustive* and therefore,

$$\lim_{Iter \rightarrow \infty} \delta = 0.$$

Furthermore, we have,

$$0 \leq \lim_{\delta \rightarrow 0} (V_{Iter}^U - V_{r,Iter}^L) \leq \lim_{\delta \rightarrow 0} (V_{r,Iter}^U - V_{r,Iter}^L) \leq \lim_{\delta \rightarrow 0} \frac{1}{4} \alpha \delta^2 = 0.$$

Consequently, for any infinitely decreasing sequence of successively refined partition elements

$$\lim_{Iter \rightarrow \infty} (V_{Iter}^U - V_{r,Iter}^L) = 0,$$

meaning that the employed bounding operation is *consistent*.  $\square$

**Definition 3:** *A selection operation is called bound improving if at least one partition element where the actual lower bound is attained is selected for further partition after a finite number of refinements.*

**Theorem 3:** *The selection operation of the employed global optimization algorithm is bound improving.*

**Proof:** The employed selection operation is clearly *bound improving* because the partition element where the actual lower bound is attained is selected for further

partition in the immediately following iteration.  $\square$

**Theorem 4:** *The employed global optimization algorithm is convergent to the global minimum.*

**Proof:** We have shown that the bounding operation is *consistent* and that the selection operation is *bound improving*, therefore according to Theorem IV.3. in [15] the employed global optimization algorithm is *convergent* to the global minimum.  $\square$

In the next section, upper and lower bounds on the total number of iterations required for  $\epsilon$ -convergence are obtained by analyzing the structure of the branch and bound tree resulting from the subdivision process.

## 8. Bounds on the Total Number of Iterations

In the previous section mathematical proof that the proposed global optimization algorithm indeed converges to the global minimum was provided. In practice, however, not mere  $\epsilon$ -convergence to the global minimum, but convergence in a finite number of iterations is required. By analyzing the structure (sparsity) of the branch and bound tree resulting from the subdivision process, finite upper and lower bounds on the total number of required for  $\epsilon$ -convergence are obtained. A similar analysis on bounds on the total number of required iterations for a concave global optimization algorithm is given in [22].

**Theorem 5:** *The maximum number of required iterations for  $\epsilon$ -convergence to the global minimum  $V^*$  for the proposed global optimization algorithm is,*

$$I^{max} = \left[ \frac{\sum_{k=1}^K (t_k^{UBD} - t_k^{LBD})^2}{\sqrt{\frac{4\epsilon}{\alpha}}} \right]^K - 1$$

**Proof:** The worst-case for convergence occurs when no partition element is fathomed until the last level of the branch and bound tree. In this case, the branch and bound tree is a complete binary tree whose end-nodes correspond to the partition elements of the last iteration. Because  $\epsilon$ -convergence is achieved only in the last level of the branch and bound tree, the lower and upper bound of  $V$  in all these partition elements are within  $\epsilon$ . Therefore, the diagonals of all these subrectangles are given by,

$$\delta = \sqrt{\frac{4\epsilon}{\alpha}}$$

The number of the partition elements corresponding to end-nodes in the branch and bound tree will then be,

$$N_{rec}^{I^{max}} = \left[ \frac{\sum_{k=1}^K (t_k^{UBD} - t_k^{LBD})^2}{\delta} \right]^K = \left[ \frac{\sum_{k=1}^K (t_k^{UBD} - t_k^{LBD})^2}{\sqrt{\frac{4\epsilon}{\alpha}}} \right]^K$$

The number of levels of the complete binary branch and bound tree is,

$$\#levels = \log_2 N_{rec}^{I^{max}} = K \log_2 \left[ \frac{\sum_{k=1}^K (t_k^{UBD} - t_k^{LBD})^2}{\sqrt{\frac{4\epsilon}{\alpha}}} \right]$$

The total number of iterations  $I^{max}$  is the total number of nodes in the complete binary tree excluding the end nodes. Therefore,

$$\begin{aligned} I^{max} &= \sum_{i=0}^{\#levels-1} 2^i = \frac{2^{\#levels} - 1}{2 - 1} = 2^{\#levels} - 1 \\ &= \left[ \frac{\sum_{k=1}^K (t_k^{UBD} - t_k^{LBD})^2}{\sqrt{\frac{4\epsilon}{\alpha}}} \right]^K - 1. \square \end{aligned}$$

By following the same line of reasoning we can find the minimum total number of required iterations.

**Theorem 6:** *The minimum number of required iterations for  $\epsilon$ -convergence to the global minimum  $V^*$  for the proposed global optimization algorithm is,*

$$I^{min} = K \log_2 \left[ \frac{\sum_{k=1}^K (t_k^{UBD} - t_k^{LBD})^2}{\sqrt{\frac{4\epsilon}{\alpha}}} \right] - 1$$

**Proof:** The best case for the proposed algorithm, in terms of efficiency of partitioning, occurs when always one partition element is fathomed in every subdivision step. In this case, each non-end node in the branch and bound tree will spawn a single child node. The levels of this tree will still be,

$$\#levels = K \log_2 \left[ \frac{\sum_{k=1}^K (t_k^{UBD} - t_k^{LBD})^2}{\sqrt{\frac{4\epsilon}{\alpha}}} \right]$$

but the total number of iterations is now,

$$I^{min} = \#levels - 1 = K \log_2 \left[ \frac{\sum_{k=1}^K (t_k^{UBD} - t_k^{LBD})^2}{\sqrt{\frac{4\epsilon}{\alpha}}} \right] - 1. \quad \square$$

**Remark:** It is interesting to note that although  $I^{max}$  is exponential in the total number of variables  $K$ , which is something that should have been expected since global optimization problems belong to the NP-complete [12],  $I^{min}$  is only linear in  $K$ . In practice, the total number of iterations is much closer to  $I^{min}$  rather than to  $I^{max}$ . In the next section the proposed global optimization algorithm is applied to a number of example problems.

## 9. Examples

Two different examples of increasing difficulty are considered in this section. Based on the analysis presented on previous sections, the problem of minimizing the total potential energy of a molecule in the coordinate space of independent dihedral (torsion) angles  $t_k$  is formulated as follows:

$$\min_{t_k} \sum_{j=4}^m \sum_{(i_1, \dots, i_j) \in \mathcal{B}^j} v(r_{i_1, i_j})$$

$$\text{where } r_{i_1, i_j} = r_{i_1, i_j}(\phi_{ijkl}, (i, j, k, l) \in \mathcal{B}^4), \\ \forall (i_1, \dots, i_j) \in \mathcal{B}^j, j = 4, \dots, m$$

$$\text{and } \phi_{ijkl} = \phi_{ijkl}(t_k, k = 1, \dots, K), \quad \forall (i, j, k, l) \in \mathcal{B}^4$$

$$t_k^L \leq t_k \leq t_k^U, \quad k = 1, \dots, K.$$

Note that different expressions for the pairwise potential function  $v(r)$  have been employed in the example problems. Also, the functionality between the non-bonded atomic Euclidean distances  $r_{ij}$  and the dihedral angles is given in Appendix A. Furthermore, the relations between the dihedral angles  $\phi_{ijkl}$  and the independent dihedral angles  $t_k$  are listed for every individual example.

### 9.1. Pseudoethane

The first example is the one-dimensional conformational problem of the fully substituted ethane molecule (see Figure 2) as shown in [23]. All three hydrogen atoms of both carbon atom are replaced with one C, one O, and one N atom. Clearly such

a molecule cannot exist; it simply corresponds to a prototype for testing methods that predict molecular conformations based on the energetics of non-bonded atom interactions.

The pseudoethane molecule involves four C, two N, and two O atoms. The number of non-bonded interactions is equal to the total number of unique chains of covalently bonded atoms involving at least four atoms. The longest chain in the pseudoethane molecule involves four atoms ( $m = 4$ ) and there are nine such chains. Two for each heteratomic pair (C-N, C-O, and N-O), and one for each C-C, N-N, O-O interaction. The Lennard-Jones potential function has been adopted for describing the non-bonded atom interactions.

$$v(r_{ij}) = \frac{B_{ij}}{r_{ij}^{12}} - \frac{A_{ij}}{r_{ij}^6}$$

The values of the interaction parameters  $A_{ij}$ ,  $B_{ij}$  are given in Table 1 [14]. Furthermore, all covalent bond lengths have been assumed to be equal to 1.54 Å and all covalent bond angles equal to 109.5°. For the current set of data, this problem has been reported [23] to have three local minima with similar objective values.

$$\begin{aligned} \theta = 61.42^\circ &\longrightarrow V_{min} = -0.79733156 \text{ kcal/mol} \\ \theta = 296.12^\circ &\longrightarrow V_{min} = -1.03989551 \text{ kcal/mol} \\ \theta = 183.45^\circ &\longrightarrow V_{min} = -1.07111459 \text{ kcal/mol} \end{aligned}$$

where  $\theta$  is the dihedral angle between the two C-C-C planes.

Based on the measure of the matrix approach, the value of the parameter  $\alpha$  is found to be 9.042908. The single independent internal variable is selected to be the dihedral angle  $t_1$  between the two C-C-C planes. All dihedral angles are then equal to  $t_1$ ,  $t_1 + \frac{2\pi}{3}$ , or  $t_1 - \frac{2\pi}{3}$ .

$$\begin{aligned} \phi_{C(1)-C(4)-C(5)-C(6)} &= t_1 \\ \phi_{C(1)-C(4)-C(5)-N(7)} &= t_1 + \frac{2\pi}{3} \\ \phi_{C(1)-C(4)-C(5)-O(8)} &= t_1 - \frac{2\pi}{3} \\ \phi_{N(2)-C(4)-C(5)-C(6)} &= t_1 + \frac{2\pi}{3} \\ \phi_{N(2)-C(4)-C(5)-N(7)} &= t_1 - \frac{2\pi}{3} \\ \phi_{N(2)-C(4)-C(5)-O(8)} &= t_1 \\ \phi_{O(3)-C(4)-C(5)-C(6)} &= t_1 - \frac{2\pi}{3} \\ \phi_{O(3)-C(4)-C(5)-N(7)} &= t_1 \\ \phi_{O(3)-C(4)-C(5)-O(8)} &= t_1 + \frac{2\pi}{3} \end{aligned}$$

The global bounds on  $t_1$  are  $[\sigma, \sigma + 2\pi]$  where  $\sigma$  is a positive parameter between zero and  $2\pi$  which allows us to observe the performance of the proposed algorithm by solving iteratively the same problem for different randomly selected values of  $\sigma$ . Three different choices for the parameter  $\alpha$  have been considered, ( $\alpha = 1$ ,  $\alpha = 5$ , and  $\alpha = 10$ ) and the tolerance  $\epsilon$  was set to  $10^{-4}$ . For each value of  $\alpha$  the same problem was solved for 100 different values of  $\sigma$ . For  $\alpha = 10$  the global optimization algorithm always converged to the global minimum,

$$t_1^* = 183.45^\circ, \quad V^* = -1.07111459 \text{ kcal/mol}$$

in about 20 iterations with an average CPU time of about 2.5 seconds on a HP-730 workstation. For  $\alpha = 5$  the algorithm still converged to the global minimum solution in all 100 runs in about 16 iterations and with an average CPU time of about 1.8 seconds. Note that for  $\alpha$  as small as one, 75% of the times convergence to the global minimum was achieved for only a fraction of the previously mentioned computational requirements.

## 9.2. 1,2,3-trichloro-1,4-bifluoro-butane

This second larger example addresses the problem of finding the global minimum total potential energy conformation of 1,2,3-trichloro-1,4-bifluoro-butane (CHClF-CHCl-CHCl-CH<sub>2</sub>F), (see Figure 3). This molecule involves 27 different chains composed of four covalently bonded atoms; 18 chains of five covalently bonded atoms; and 9 chains of six covalently bonded atoms. In this example all covalent bond angles are assumed to be tetrahedral (109.5°). The equilibrium bond lengths as well as the parameters  $A_{ij}$  and  $B_{ij}$  for the Lennard-Jones potential are listed in Table 2 and Table 3.

Clearly, only three dihedral angles are required to fully describe the rotational conformation of the molecule, and one possible selection is the following.

$$\begin{aligned} 0 \leq t_1 &\longleftarrow \phi_{F(1)-C(4)-C(5)-Cl(6)} \leq 2\pi \\ 0 \leq t_2 &\longleftarrow \phi_{Cl(6)-C(5)-C(7)-Cl(9)} \leq 2\pi \\ 0 \leq t_3 &\longleftarrow \phi_{Cl(9)-C(7)-C(10)-F(12)} \leq 2\pi \end{aligned}$$

The expressions for all 27 dihedral angles are then,

$$\begin{aligned} \phi_{F(1)-C(4)-C(5)-Cl(6)} &= t_1 & \phi_{Cl(6)-C(5)-C(7)-Cl(9)} &= t_2 \\ \phi_{F(1)-C(4)-C(5)-C(7)} &= t_1 + \frac{2\pi}{3} & \phi_{Cl(6)-C(5)-C(7)-C(10)} &= t_2 + \frac{2\pi}{3} \\ \phi_{F(1)-C(4)-C(5)-H(8)} &= t_1 - \frac{2\pi}{3} & \phi_{Cl(6)-C(5)-C(7)-H(11)} &= t_2 - \frac{2\pi}{3} \\ \phi_{Cl(2)-C(4)-C(5)-Cl(6)} &= t_1 - \frac{2\pi}{3} & \phi_{C(4)-C(5)-C(7)-Cl(9)} &= t_2 - \frac{2\pi}{3} \\ \phi_{Cl(2)-C(4)-C(5)-C(7)} &= t_1 & \phi_{C(4)-C(5)-C(7)-C(10)} &= t_2 \\ \phi_{Cl(2)-C(4)-C(5)-H(8)} &= t_1 + \frac{2\pi}{3} & \phi_{C(4)-C(5)-C(7)-H(11)} &= t_2 + \frac{2\pi}{3} \\ \phi_{H(3)-C(4)-C(5)-Cl(6)} &= t_1 + \frac{2\pi}{3} & \phi_{H(8)-C(5)-C(7)-Cl(9)} &= t_2 + \frac{2\pi}{3} \\ \phi_{H(3)-C(4)-C(5)-C(7)} &= t_1 - \frac{2\pi}{3} & \phi_{H(8)-C(5)-C(7)-C(10)} &= t_2 - \frac{2\pi}{3} \\ \phi_{H(3)-C(4)-C(5)-H(8)} &= t_1 & \phi_{H(8)-C(5)-C(7)-H(11)} &= t_2 \end{aligned}$$
  

$$\begin{aligned} \phi_{Cl(9)-C(7)-C(10)-F(12)} &= t_3 \\ \phi_{Cl(9)-C(7)-C(10)-H(13)} &= t_3 + \frac{2\pi}{3} \\ \phi_{Cl(9)-C(7)-C(10)-H(14)} &= t_3 - \frac{2\pi}{3} \\ \phi_{C(5)-C(7)-C(10)-F(12)} &= t_3 - \frac{2\pi}{3} \\ \phi_{C(5)-C(7)-C(10)-H(13)} &= t_3 \\ \phi_{C(5)-C(7)-C(10)-H(14)} &= t_3 + \frac{2\pi}{3} \\ \phi_{H(11)-C(7)-C(10)-F(12)} &= t_3 + \frac{2\pi}{3} \\ \phi_{H(11)-C(7)-C(10)-H(13)} &= t_3 - \frac{2\pi}{3} \\ \phi_{H(11)-C(7)-C(10)-H(14)} &= t_3 \end{aligned}$$

Table 1. Data for parameters  $A_{ij}$  and  $B_{ij}$  in the pseudoethane example

atom-atom	$B$ ( $10^4$ kcal $\text{\AA}^{12}/\text{mol}$ )	$A$ (kcal $\text{\AA}^6/\text{mol}$ )
C(sp <sup>3</sup> )-C(sp <sup>3</sup> )	28.58	372.5
N(sp <sup>3</sup> )-N(sp <sup>3</sup> )	17.82	344.7
O(sp <sup>2</sup> )-O(sp <sup>2</sup> )	14.49	367.2
C(sp <sup>3</sup> )-N(sp <sup>3</sup> )	16.82	348.7
C(sp <sup>3</sup> )-O(sp <sup>2</sup> )	20.52	367.2
N(sp <sup>3</sup> )-O(sp <sup>2</sup> )	15.75	349.5

Table 2. Covalent bond lengths  $r_o$  in the 1,2,3-trichloro-1,4-bifluoro-butane example

atom-atom	$r_o$ ( $\text{\AA}$ )
C-C	1.54
C-H	1.06
C-Cl	1.77
C-F	1.39

Table 3. Data for parameters  $A_{ij}, B_{ij}$  in the 1,2,3-trichloro-1,4-bifluoro-butane example

atom-atom	$A$ (kcal $\text{\AA}^6/\text{mol}$ )	$B$ ( $10^{-4}$ kcal $\text{\AA}^{12}/\text{mol}$ )
F-Cl	457.6	29.6700
F-H	79.8	0.9361
F-C	223.2	10.0700
H-H	76.0	0.7220
H-C	127.4	3.7430
H-Cl	272.8	8.0740
Cl-Cl	1562.0	200.5000
Cl-C	759.3	64.6400

For the current data a large number of solutions have been found which are summarized in Table 4. The proposed global optimization algorithm for  $\alpha = 5.0$  and  $\epsilon = 10^{-4}$  in all 100 runs converged to the global minimum,

$$t_1^* = 291.95^\circ, t_2^* = 180.42^\circ, t_3^* = 182.86^\circ \longrightarrow V^* = -5.271002 \text{ kcal/mol}$$

in about 1000 iterations with an average CPU time of 1400 seconds on a HP-730 workstation. For  $\alpha = 1$  convergence to the global minimum is achieved 70 % of the times for only a fraction of the previous computational requirements.

Table 4. Solutions for the 1,2,3-trichloro-1,4-bifluoro-butane example

$t_1^o$	$t_2^o$	$t_3^o$	$V_{min}^{local}$ kcal/mol
291.95	180.42	182.86	-5.271002
291.95	180.49	290.97	-5.159241
291.94	180.06	64.59	-5.127455
194.32	179.85	182.85	-4.975758
291.08	292.43	184.17	-4.944256
194.33	179.91	290.95	-4.865099
291.07	292.13	65.12	-4.861028
289.02	67.06	183.21	-4.841873
194.30	179.49	64.62	-4.830295
289.03	67.09	290.93	-4.697346
289.04	67.06	65.54	-4.691369
194.30	67.43	183.17	-4.603630
194.61	292.47	184.16	-4.565333
194.60	292.17	65.20	-4.485494
194.60	292.21	291.12	-4.473805
194.29	67.46	290.93	-4.461059
194.29	67.44	65.53	-4.452745
56.17	181.51	182.86	-4.451375
56.16	181.58	290.87	-4.334131
55.61	68.03	183.30	-4.296107
56.18	181.14	64.58	-4.290602
55.14	292.50	184.16	-4.192450
55.60	68.02	290.99	-4.160552
55.61	67.98	65.47	-4.153306
55.16	292.27	65.15	-4.116179
55.14	292.28	291.13	-4.102804

## 10. Summary and Conclusions

In this paper a global optimization algorithm was proposed for finding the global minimum potential energy conformation of small molecules based on the energetics of non-bonded interactions. First, explicit relations for the non-bonded atom Euclidean distances as function of bond lengths, covalent bond angles, and dihedral (torsion) angles were derived. Then, the problem was formulated as an unconstrained nonconvex optimization problem on a set of independent dihedral angles. Based on an eigenvalue analysis, a convex lower bounding function  $\mathcal{L}$  of the total potential energy function  $V$  was defined and various methods for estimating the required parameter  $\alpha$  were also discussed. The special properties of the function  $\mathcal{L}$  coupled with an efficient partitioning scheme enabled us to construct a global optimization algorithm which was then proved to always converge to the global minimum conformation through the solution of a series of convex nonlinear optimization problems. Lower and upper bounds on the total number of required iterations were also derived. Finally, the proposed approach was applied successfully to a number of example problems of which two are included in this paper. It should be emphasized that the global optimization algorithm is not restricted to

the specifics of the problem at hand. It can and has been applied to other classes of problems. Also, the presented analysis can be extended to more realistic interaction fields involving complex potential interaction terms (e.g. bending, stretching, torsion, bending-stretching, and out-of-plane distortion). Work in this direction is currently in progress.

### Acknowledgements

Financial support from the National Science Foundation under Grants CBT-8857013, 160-4099 as well as Amoco Chemicals Co., Mobil Co., and Tennessee Eastman Co. is gratefully acknowledged.

### References

- [1] N.L. Allinger, *Adv. Phys. Org. Chem.*, **13**, 1, (1976).
- [2] N.L. Allinger, M.A. Miller, F.A. Van Catledge, and J.A. Hirsch, Conformational Analysis. LVII. The Calculation of the Conformational Structures of Hydrocarbons by the Westheimer-Hendrickson-Wiberg Method, *J. Am. Chem. Soc.*, **89**, 4345, (1967).
- [3] U. Burkert and N.L. Allinger, *Molecular Mechanics*, (American Chemical Society, Washington, D.C. 1982).
- [4] W.A. Coppel, *Stability and Asymptotic Behavior of Differential Equations*, (Boston, Mass.: D.C. Heath, 1965).
- [5] T.H. Cormen, C.E. Leiserson, and R.L. Rivest, *Introduction to Algorithms*, (McGraw-Hill Book Company, 1989).
- [6] G. Dalquist, Stability and Error Bounds in the Integration of Ordinary Differential Equations, *Trans. Roy. Inst. Tech.*, **130**, (1959).
- [7] P. de la Mare, L. Fowden, E.D. Hughes, and C.K. Ingold, *J. Chem. Soc.*, 3200, (1955).
- [8] C.A. Desoer and H. Haneda, The Measure of a Matrix as a Tool to Analyze Computer Algorithms for Circuit Analysis, *IEEE Trans. Circ. Th.*, **19**, 480, (1972).
- [9] I. Dostrovsky, E.D. Hughes, and C.K. Ingold, *J. Chem. Soc.*, 173, (1946).
- [10] C.A. Floudas and V. Visweswaran, A Global Optimization Algorithm for Certain Classes of Nonconvex NLPs, Part 1: Theory, *Computers and Chemical Engineering*, **14**, 1397, (1990).
- [11] C.A. Floudas and V. Visweswaran, Primal-Relaxed Dual Global Optimization Approach, *Journal of Optimization Theory and its Applications*, **78**, Iss. 2, 187, (1993).
- [12] M.R. Garey and D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, (Freeman, San Francisco, 1979).
- [13] T.L. Hill, *J. Chem. Phys.*, **14**, 465, (1945).
- [14] A.J. Hopfinger, *Conformational Properties of Macromolecules*, (Academic Press, New York and London, 1973).
- [15] R. Horst and H. Tuy, *Global Optimization, Deterministic Approaches*, (Springer-Verlag, Berlin 1990).
- [16] E.J. Jacob, H.B. Thompson, and L.S. Bartell, *J. Chem. Phys.*, **49**, 2574, (1968).
- [17] A.C. Kokossis and C.A. Floudas, Stability Issues in Optimal Design: Synthesis of Complex Reactor Networks, *AIChE*, in press, (1994).

- [18] D.G. Luenberger, *Linear and Nonlinear Programming*, (Addison-Wesley, Reading, Mass., 1984).
- [19] C.D. Maranas and C.A. Floudas, A Global Optimization Approach for Lennard-Jones Microclusters, *J. Chem. Phys.*, **97**, 10, (1992).
- [20] C.D. Maranas and C.A. Floudas, Global Optimization for Molecular Conformation Problems, *Annals of Operation Research*, in press, (1993).
- [21] B.A. Murtagh and M.A. Saunders, *MINOS 5.3 User's Guide*, (Systems Optimization Laboratory, Department of Operations Research, Stanford University, 1987).
- [22] A.T. Phillips and J.B. Rosen, A Parallel Algorithm for Constrained Concave Quadratic Global Minimization *Math. Prog.*, **42**, 421, (1988).
- [23] L. Piela, J. Kostrowicki and H.A. Scheraga, The multiple-Minima Problem in the Conformational Analysis of Molecules. Deformation of the Potential Energy Hypersurface by the Diffusion Equation Method, *J. Phys. Chem.*, **93**, 3339, (1989).
- [24] M. Rieger and F.H. Westheimer, *J. Am. Chem. Soc.*, **72**, 19, (1950).
- [25] Z. Slanina, Does the Global Energy Minimum Always Also Mean the Thermodynamically Most Stable Structure?, *J. Mol. Str.*, **206**, 143, (1990).
- [26] R. Susnow, R.B. Nachbar, C. Schutt, and H. Rabitz, Sensitivity of Molecular Structure to Intramolecular Potentials, *J. Phys. Chem.*, **95**, 8585, (1991).
- [27] H.B. Thompson, Calculation of Cartesian Coordinates and Their Derivatives from Internal Molecular Coordinates, *J. Chem. Phys.*, **47**, 3407, (1969).
- [28] J.M. Troyer and F.E. Cohen, *Simplified Models for Understanding and Predicting Protein Structure*, VCH Publishers, New York, NY, 1991.
- [29] B. van der Graaf and J.M.A. Baas, The Implementation of Constraints in Molecular Mechanics to Explore Potential Energy Surfaces, *Rev. Trav. Chim. Pays-Bas*, **99**, 327, (1980).
- [30] V. Visweswaran and C.A. Floudas, A Global Optimization Algorithm (GOA) for Certain Classes of Nonconvex NLPs, Part 2: Applications of Theory and Test Problems, *Computers and Chemical Engineering*, **14**, 1419, (1990).
- [31] V. Visweswaran and C.A. Floudas, A Global Optimization Algorithm (GOA) for Certain Classes of Nonconvex NLPs, III. New Properties for Improved Computational Performances, *Proceedings of Process Systems Engineering, PSE '91*, I.6.1, (1991).
- [32] V. Visweswaran and C.A. Floudas, New Properties and Computational Improvement of the GOA Algorithm for Problems with Quadratic Objective Function and Constraints, *Journal of Global Optimization*, **3**, Iss. 2, in press, (1993).
- [33] A. Warshel and S.J. Lifson, Consistent Force Field Calculations. II Crystal Structures, Sublimation Energies, Molecular and Lattice Vibrations, Molecular Conformations, and Enthalpies of Alkanes, *J. Chem. Phys.*, **53**, 582, (1970).
- [34] F.H. Westheimer and J.E. Mayer, *J. Chem. Phys.*, **14**, 733, (1946).
- [35] J.E. Williams, P.J. Stang, P. Schleyer, Physical Organic Chemistry: Quantitative Conformational Analysis; Calculation Methods, *Ann. Rev. Phys. Chem.*, **19**, 531, (1968).

## Appendix

### A.1. Analytical Expressions for $r_{ij}$ 's

Based on the analysis presented in section (3), the following general expressions for the Euclidean distances between the end atoms in  $m$ -atom ( $m \leq 6$ ) chains are obtained:

$$r_{ik}^2 = r_{ij}^2 + r_{jk}^2 - 2 \cos(\theta_{ijk}) r_{ij} r_{jk}, \quad (i, j, k) \in \mathcal{B}^3$$

$$\begin{aligned} r_{il}^2 &= r_{ij}^2 + r_{jk}^2 + r_{kl}^2 \\ &- 2 \cos(\theta_{ijk}) r_{ij} r_{jk} - 2 \cos(\theta_{jkl}) r_{jk} r_{kl} \\ &- 2 [\sin(\theta_{ijk}) \sin(\theta_{jkl}) \cos(\phi_{ijkl}) - \cos(\theta_{ijk}) \cos(\theta_{jkl})] r_{ij} r_{kl}, \\ &\quad (i, j, k, l) \in \mathcal{B}^4 \end{aligned}$$

$$\begin{aligned} r_{im}^2 &= r_{ij}^2 + r_{jk}^2 + r_{kl}^2 + r_{lm}^2 \\ &- 2 \cos(\theta_{ijk}) r_{ij} r_{jk} - 2 \cos(\theta_{jkl}) r_{jk} r_{kl} - 2 \cos(\theta_{klm}) r_{kl} r_{lm} \\ &- 2 [\sin(\theta_{ijk}) \sin(\theta_{jkl}) \cos(\phi_{ijkl}) - \cos(\theta_{ijk}) \cos(\theta_{jkl})] r_{ij} r_{kl} \\ &- 2 [\sin(\theta_{jkl}) \sin(\theta_{klm}) \cos(\phi_{jklm}) - \cos(\theta_{jkl}) \cos(\theta_{klm})] r_{jk} r_{lm} \\ &- 2 [\cos(\theta_{ijk}) \cos(\theta_{jkl}) \cos(\theta_{klm}) \\ &\quad - \sin(\theta_{ijk}) \sin(\theta_{jkl}) \cos(\theta_{klm}) \cos(\phi_{ijkl}) \\ &\quad - \sin(\theta_{ijk}) \cos(\theta_{jkl}) \sin(\theta_{klm}) \cos(\phi_{ijkl}) \cos(\phi_{jklm}) \\ &\quad - \cos(\theta_{ijk}) \sin(\theta_{jkl}) \sin(\theta_{klm}) \cos(\phi_{jklm}) \\ &\quad - \sin(\theta_{ijk}) \sin(\theta_{klm}) \sin(\phi_{ijkl}) \sin(\phi_{jklm})] r_{ij} r_{lm}, \\ &\quad (i, j, k, l, m) \in \mathcal{B}^5 \end{aligned}$$

$$\begin{aligned} r_{in}^2 &= r_{ij}^2 + r_{jk}^2 + r_{kl}^2 + r_{lm}^2 + r_{mn}^2 \\ &- 2 \cos(\theta_{ijk}) r_{ij} r_{jk} - 2 \cos(\theta_{jkl}) r_{jk} r_{kl} \\ &- 2 \cos(\theta_{klm}) r_{kl} r_{lm} - 2 \cos(\theta_{lmn}) r_{lm} r_{mn} \\ &- 2 [\sin(\theta_{ijk}) \sin(\theta_{jkl}) \cos(\phi_{ijkl}) - \cos(\theta_{ijk}) \cos(\theta_{jkl})] r_{ij} r_{kl} \\ &- 2 [\sin(\theta_{jkl}) \sin(\theta_{klm}) \cos(\phi_{jklm}) - \cos(\theta_{jkl}) \cos(\theta_{klm})] r_{jk} r_{lm} \\ &- 2 [\sin(\theta_{klm}) \sin(\theta_{lmn}) \cos(\phi_{klmn}) - \cos(\theta_{klm}) \cos(\theta_{lmn})] r_{kl} r_{mn} \\ &- 2 [\cos(\theta_{ijk}) \cos(\theta_{jkl}) \cos(\theta_{klm}) \\ &\quad - \sin(\theta_{ijk}) \sin(\theta_{jkl}) \cos(\theta_{klm}) \cos(\phi_{ijkl}) \\ &\quad - \sin(\theta_{ijk}) \cos(\theta_{jkl}) \sin(\theta_{klm}) \cos(\phi_{ijkl}) \cos(\phi_{jklm}) \\ &\quad - \cos(\theta_{ijk}) \sin(\theta_{jkl}) \sin(\theta_{klm}) \cos(\phi_{jklm}) \\ &\quad - \sin(\theta_{ijk}) \sin(\theta_{klm}) \sin(\phi_{ijkl}) \sin(\phi_{jklm})] r_{ij} r_{lm} \\ &- 2 [\cos(\theta_{jkl}) \cos(\theta_{klm}) \cos(\theta_{lmn}) \\ &\quad - \sin(\theta_{jkl}) \sin(\theta_{klm}) \cos(\theta_{lmn}) \cos(\phi_{jklm}) \\ &\quad - \sin(\theta_{jkl}) \cos(\theta_{klm}) \sin(\theta_{lmn}) \cos(\phi_{jklm}) \cos(\phi_{klmn}) \\ &\quad - \cos(\theta_{jkl}) \sin(\theta_{klm}) \sin(\theta_{lmn}) \cos(\phi_{klmn})] \end{aligned}$$

$$\begin{aligned}
& -\sin(\theta_{jkl}) \sin(\theta_{lmn}) \sin(\phi_{jklm}) \sin(\phi_{klmn})] r_j k r_{mn} \\
+ & 2 [\cos(\theta_{ijk}) \cos(\theta_{jkl}) \cos(\theta_{klm}) \cos(\theta_{lmn}) \\
& -\sin(\theta_{ijk}) \sin(\theta_{jkl}) \cos(\theta_{klm}) \cos(\theta_{lmn}) \cos(\phi_{ijk}) \\
& -\sin(\theta_{ijk}) \cos(\theta_{jkl}) \sin(\theta_{klm}) \cos(\theta_{lmn}) \cos(\phi_{ijk}) \cos(\phi_{jklm}) \\
& -\cos(\theta_{ijk}) \sin(\theta_{jkl}) \sin(\theta_{klm}) \cos(\theta_{lmn}) \cos(\phi_{jklm}) \\
& -\sin(\theta_{ijk}) \cos(\theta_{jkl}) \cos(\theta_{klm}) \sin(\theta_{lmn}) \cos(\phi_{ijk}) \cos(\phi_{jklm}) \cos(\phi_{klmn}) \\
& -\cos(\theta_{ijk}) \sin(\theta_{jkl}) \cos(\theta_{klm}) \sin(\theta_{lmn}) \cos(\phi_{jklm}) \cos(\phi_{klmn}) \\
& -\cos(\theta_{ijk}) \cos(\theta_{jkl}) \sin(\theta_{klm}) \sin(\theta_{lmn}) \cos(\phi_{klmn}) \\
& +\sin(\theta_{ijk}) \sin(\theta_{jkl}) \sin(\theta_{klm}) \sin(\theta_{lmn}) \cos(\phi_{ijk}) \cos(\phi_{klmn}) \\
& -\sin(\theta_{ijk}) \sin(\theta_{klm}) \cos(\theta_{lmn}) \sin(\phi_{ijk}) \sin(\phi_{jklm}) \\
& -\sin(\theta_{ijk}) \cos(\theta_{klm}) \sin(\theta_{lmn}) \sin(\phi_{ijk}) \sin(\phi_{jklm}) \cos(\phi_{klmn}) \\
& +\sin(\theta_{ijk}) \sin(\theta_{lmn}) \sin(\phi_{ijk}) \cos(\phi_{jklm}) \sin(\phi_{klmn}) \\
& -\sin(\theta_{ijk}) \cos(\theta_{jkl}) \sin(\theta_{lmn}) \cos(\phi_{ijk}) \sin(\phi_{jklm}) \sin(\phi_{klmn}) \\
& -\cos(\theta_{ijk}) \sin(\theta_{jkl}) \sin(\theta_{lmn}) \sin(\phi_{jklm}) \sin(\phi_{klmn})] r_j r_{mn} \\
& (i, j, k, l, m, n) \in \mathcal{B}^6
\end{aligned}$$

## A.2. Estimation of the Parameter $\alpha$

As it has been shown in section (4),  $\alpha$  is a nonnegative parameter which controls by how much the eigenvalues of  $\mathcal{L}$  will be greater than the eigenvalues of  $V$ . If  $\mathcal{L}$  is to be a convex function then  $\alpha$  must be at least equal to minus one half of the minimum eigenvalue of  $V$ . Clearly, the smaller the value of  $\alpha$ , the more tightly the function  $\mathcal{L}$  will underestimate the original function  $V$ , and presumably the smaller the total number of iterations required for convergence. Therefore, ideally one would want the parameter  $\alpha$  to be exactly equal to minus one half of the minimum eigenvalue of  $V$ .

$$\alpha = \max \left\{ 0, \max_{t_k^L \leq t_k \leq t_k^U} \left( -\frac{1}{2} \lambda_k(t_k) \right) \right\}$$

However, explicit expressions for the eigenvalues  $\lambda_k$  as functions of  $t_k$ , in general, can not be derived. Therefore, the minimum eigenvalue  $\lambda_k(t_k)$  both over  $k$  and  $t_k$  can implicitly be obtained with the solution of the following optimization problem,

$$\min_{t_k^L \leq t_k \leq t_k^U} \lambda$$

$$\text{subject to} \quad \det [He(t_k) - \lambda I] = 0$$

where  $He(t_k)$  is the Hessian matrix of  $V$  as a function of  $t_k$ , and  $I$  the identity matrix. The equality constraint in the above formulation involves  $K$  solutions for  $\lambda$ . Each solution  $\lambda(t_k)$  must then be minimized over  $t_k^L \leq t_k \leq t_k^U$ , and the negative one half of this solution is the parameter  $\alpha$ . Because the lengthy expansion of the determinant corresponds to a nonconvex equality constraint, solutions of the above problem can be guaranteed to correspond to only local minima. However, by repetitively solving this problem from different initial points one increases the chances of locating the global minimum. Even if a mere local minimum is obtained, a safety margin can be used to compensate for the difference from the global minimum. In principle, however, only lower bounds on the required value of the parameter  $\alpha$  can be obtained with this formulation.

Upper bounds on  $\alpha$  can be derived by using the concept of a measure of a matrix. The measure  $\mu(A)$  of a square matrix  $A$  is a mapping from  $R^{K \times K}$  into  $R$  [6, 4, 8, 17],

$$\mu(A) = \lim_{\epsilon \rightarrow 0^+} \frac{\|I + \epsilon A\| - 1}{\epsilon}$$

defined as the directional derivative of the induced norm  $\|I + \epsilon A\|$  at the point  $I$  in the direction of  $A$ . By definition, the induced norm of a matrix  $A$  corresponds to the largest amount by which any vector in  $R^K$  is amplified when multiplied by  $A$ . Different induced norms give rise to different expressions for the measure of a matrix  $A$ .

$$\begin{aligned} \|A\|_1 &= \max_j \sum_i |a_{ij}| & \rightarrow \mu_1(A) &= \max_j \left[ a_{jj} + \sum_{j \neq i} |a_{ij}| \right] \\ \|A\|_2 &= \max_i \lambda_i [(A^*A)^{1/2}] & \rightarrow \mu_2(A) &= \max_i \lambda_i \left( \frac{A^*+A}{2} \right) \\ \|A\|_\infty &= \max_i \sum_j |a_{ij}| & \rightarrow \mu_\infty(A) &= \max_i \left[ a_{ii} + \sum_{i \neq j} |a_{ij}| \right] \end{aligned}$$

The measure of a matrix has been shown [6, 4, 8, 17] to satisfy a number of important properties. Among these properties are the following three:

**Property 1.** The measure  $\mu(A)$  of a square matrix  $A$  always exists.

**Property 2.**  $\mu(aA + (1 - a)B) \leq a\mu(A) + (1 - a)\mu(B), \forall a \in [0, 1]$ ,

**Property 3.**  $\mu(-A) \leq \text{Real} \{ \lambda_k \} \leq \mu(A), k = 1, \dots, K$ .

These properties imply that the measure  $\mu(A)$  of a matrix  $A$  is always uniquely defined, is convex, and can be used for deriving lower and upper bounds on the eigenvalues of matrix  $A$ . Based on these properties, an upper bound for the parameter  $\alpha$  can be obtained by solving the following optimization problem.

$$\alpha = -\frac{1}{2} \min_{t_k^L \leq t_k \leq t_k^U} -\mu(-He(t_k)) = \frac{1}{2} \max_{t_k^L \leq t_k \leq t_k^U} \mu(-He(t_k))$$

where  $He(t_k)$  is the Hessian matrix of  $V$  as a function of  $t_k$ . The next step is to select which measure  $\mu(1, 2$  or  $\infty)$  will be employed in the optimization problem. Clearly,  $\mu_1$  and  $\mu_\infty$  are more suitable because the maximization terms that they involve can be easily reduced to a set of inequality constraints. Furthermore, it is quite straightforward to show that because  $He(t_k)$  is a symmetric matrix, the expressions for  $\mu_1, \mu_\infty$  are identical and that the eigenvalues of  $He(t_k)$  are all real. Based on the definition of the measure of a matrix an upper bound on  $\alpha$  is given by the following formulation,

$$\alpha = \min_{\substack{t_k^L \leq t_k \leq t_k^U \\ m \geq 0}} \frac{m}{2}$$

subject to

$$m \geq h_{kk} + \sum_{\substack{k'=1 \\ k' \neq k}}^K |h_{kk'}(t_k)|, \quad k = 1, \dots, K$$

where  $h_{kk'}(t_k)$  is the  $(k, k')$  element of the Hessian matrix,

$$h_{kk'} = \frac{\partial^2 V}{\partial t_k \partial t_{k'}}$$

which can explicitly be written as,

$$h_{kk'} = \sum_{j=4}^m \sum_{(i_1, \dots, i_j) \in B^j} \frac{\partial^2 v(r_{i_1, \dots, i_j})}{\partial t_k \partial t_{k'}}$$

The partial derivative of  $v$  in terms of  $t_k$  and  $t_{k'}$  can be simplified to,

$$\frac{\partial^2 v(r_{i_1, \dots, i_j})}{\partial t_k \partial t_{k'}} = \frac{d^2 v(r_{i_1, \dots, i_j})}{dr_{i_1, \dots, i_j}^2} \frac{\partial r_{i_1, \dots, i_j}}{\partial t_k} \frac{\partial r_{i_1, \dots, i_j}}{\partial t_{k'}} + \frac{dv(r_{i_1, \dots, i_j})}{dr_{i_1, \dots, i_j}} \frac{\partial^2 v(r_{i_1, \dots, i_j})}{\partial t_k \partial t_{k'}}$$

where the first and second derivative of  $v$  in terms of  $r$  as well as the derivatives of  $r$  with respect to  $t_k$  and/or  $t_{k'}$  can easily be calculated based on expressions given in earlier sections.

Based on the implicit minimization of the minimum eigenvalue of  $V(t_k)$  and the maximization of the measure of the negative Hessian matrix of  $V(t_k)$  respectively, lower and upper bounds on the parameter  $\alpha$  can be obtained.

Alternatively, under certain conditions, one can derive a tight upper bound on  $\alpha$  by analyzing possible special structure of the expression for the pairwise potential interaction  $v(r)$  [19, 20]. Furthermore, in practice heuristic rules for the selection of  $\alpha$  can be derived and tested by applying the global optimization algorithm on problems where the solution is known. In fact, computational experience indicates that even if the value of  $\alpha$  is not large enough to maintain convexity of  $\mathcal{L}$  over the entire box region, the employed local optimization algorithm will most likely converge to the global minimum of the nearly convex function  $\mathcal{L}$ .

**A.3. Derivation of Dual Lower Bound  $\mathcal{L}$**

It is quite interesting that the convex lower bounding function  $\mathcal{L}$  can alternatively be obtained from the relaxed dual formulation [10, 30, 31, 11, 32] of a transformed problem equivalent with the minimization of  $V$ , as it has been performed earlier [19, 20].

With the introduction of a new set of variables  $T_k$  the minimization of  $V(t_k)$  over the box constraints defined by the bounds  $t_k^L, t_k^U$  can equivalently be reformulated as follows:

$$\begin{aligned} \min_{t_k^L \leq t_k, T_k \leq t_k^U} \quad & V(T_k) + \alpha \sum_{k=1}^K (t_k^2 - t_k T_k) \\ \text{subject to} \quad & t_k = T_k, \quad k = 1, \dots, K \end{aligned}$$

where  $t_k, T_k$  correspond to the x-type, y-type variables as presented in [10, 30, 31, 11]. The exact dual of the above problem will then be,

$$\sup_{\lambda_k} \min_{t_k^L \leq t_k, T_k \leq t_k^U} \quad V(T_k) + \alpha \sum_{k=1}^K (t_k^2 - t_k T_k + \lambda_k (t_k - T_k))$$

The exact dual can be relaxed by dropping the maximization over  $\lambda_k$  and substituting  $\lambda_k$ 's with the value derived from the KKT conditions,

$$\lambda_k = T_k^c, \quad k = 1, \dots, K$$

where  $T_k^c$  is the current point. The formulation for the relaxed dual then becomes,

$$\min_{t_k^L \leq t_k, T_k \leq t_k^U} \quad V(T_k) + \alpha \sum_{k=1}^K (t_k - T_k) (T_k^c - T_k)$$

The minimization over  $t_k$  is accounted by considering all different combinations of signs for the qualifying constraints  $T_k^c - T_k, \quad k = 1, \dots, K$  which are multiplied by  $t_k$ . More specifically,

$$\begin{aligned} t_k &= t_k^L \quad \text{if } \forall T_k, T_k^c - T_k \geq 0, \implies T_k^k = t_k^U, \quad k = 1, \dots, K \\ t_k &= t_k^U \quad \text{if } \forall T_k, T_k^c - T_k \leq 0, \implies T_k^k = t_k^L, \quad k = 1, \dots, K \end{aligned}$$

Note that both cases collapse to the following minimization problem over  $T_k$ ,

$$\min_{t_k^L \leq T_k \leq t_k^U} \quad V(T_k) + \alpha \sum_{k=1}^K (t_k^L - T_k) (t_k^U - T_k)$$

which is identical with the minimization of  $\mathcal{L}$ .

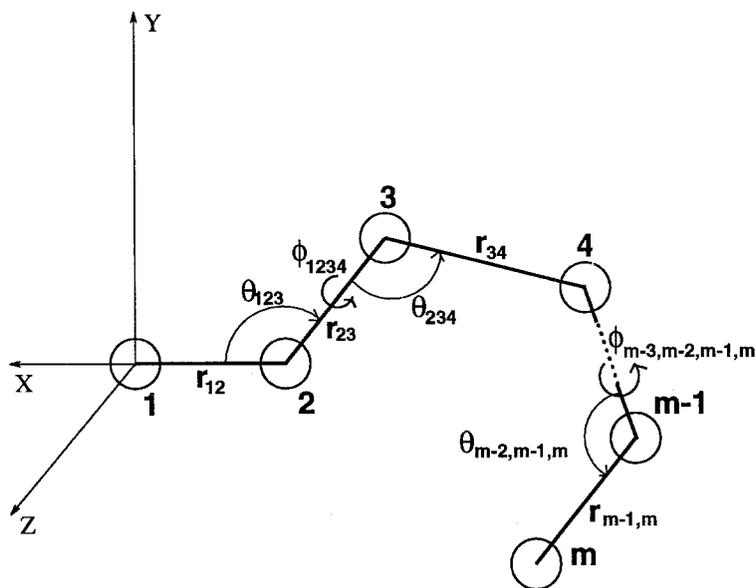


Figure A1. Coordinate set of atomic chain.

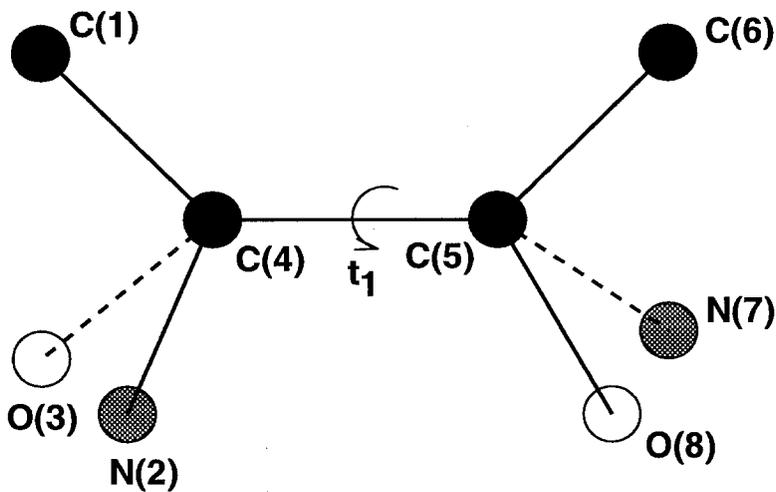


Figure A2. Pseudoethane molecule.

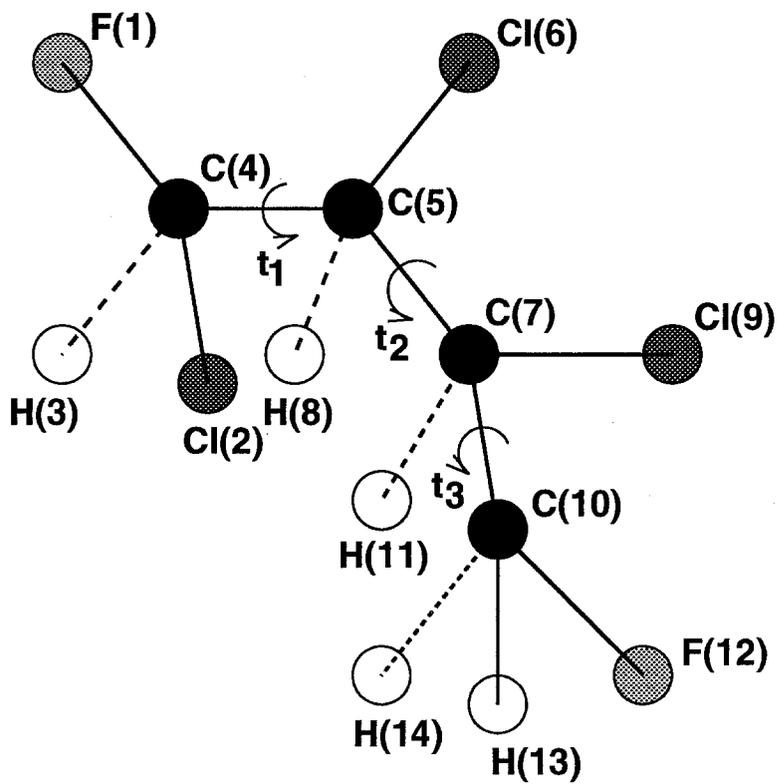


Figure A3. 1,2,3-trichloro-1,4-bifluoro-butane molecule