

·“双清论坛”专题：理论化学家视角中的仪器创制·

# 大规模原子模拟软件包 (LSASP) 程序设计和进展

刘智攀\* 商 城

(复旦大学化学系,上海市分子催化和功能材料重点实验室,  
物质计算科学教育部重点实验室,上海 200433)

**[摘要]** 本文介绍了笔者课题组开发的针对复杂体系精确计算的软件包 LSASP,包括核心算法,软件架构以及开发进展。与传统经验力场的程序不同,LSASP 软件整合现有的通用电子结构计算软件,提供统一数据接口,以通用型全局势能面神经网络势函数作为主要势能面计算工具,涵盖全局结构搜索,自动反应路径取样,过渡态寻找,分子动力学,蒙特卡洛模拟等势能面采样功能的大尺度计算平台。LSASP 软件的核心算法,随机势能面行走(SSW),具有自有知识产权,在势能面全局搜索,过渡态反应路径搜索,神经网络势函数模拟等方面具有国际竞争力。

**[关键词]** 势能面;全局优化;反应路径采样;LSASP 软件

物质所处的量子态(波函数)是其所包含原子集几何结构的函数。为了求解几何结构与电子波函数的关联,得到体系总能量,理论化学利用现代计算设备和电子结构计算软件,迭代自洽求解薛定谔方程(Schrodinger equation)。然而,基于量子力学方法计算体系的势能面函数,精度较高的同时,计算代价十分高昂,故而在复杂、大体系计算方面显得力不从心。如何进一步提高势能面计算精度,并有效对复杂体系势能面采样,是理论和计算化学的核心问题。

传统上对于复杂大体系的势能面采样,一般通过基于经验力场的分子动力学方法研究,代表性的软件包,包括 GULP, AMBER, LAMMPS 等<sup>[1-3]</sup>。由于经验力场的精度相对低,基本不能研究化学反应,这些软件研究对象一般较为专一,如简单合金, SiO<sub>2</sub> 等简单氧化物材料,以及生物蛋白等体系。本文介绍的 Large-Scale Atomic Simulation Package (LSASP) 程序是针对复杂体系精确计算的软件包。与传统经验力场的程序不同,LSASP 软件整合现有的通用电子结构计算软件,提供统一数据接口,以通用型全局势能面神经网络势函数为主要势能面计算工具,涵盖全局结构搜索,反应路径取样,过渡态寻找,分子动力学,蒙特卡洛模拟等势能面采样功能的

大尺度计算平台。LSASP 软件的核心算法具有自有知识产权,在势能面全局搜索,过渡态反应路径搜索,神经网络势函数模拟等方面具有国际竞争力。下文,我们简要介绍 LSASP 软件包的核心算法,软件架构以及开发进展。

## 1 LSASP 软件包核心算法

### 1.1 偏置势—约束双子算法(BP-CBD)

势能面的二阶梯度是区分极小点和鞍点的关键数学指标。原则上,通过对角化二阶梯度矩阵, Hessian, 即可以得到原子振动的本征模式和频率,然而计算 Hessian 矩阵代价昂贵,使得实际过渡态搜索和势能面取样中难于采用严格 Hessian 对角化方法。由 Henkelman 和 Jonsson 等人提出<sup>[4]</sup>并被其他一些课题组改进<sup>[5]</sup>的双子(Dimer 方法),可以在不需要 Hessian 矩阵的前提下寻找体系的最小本征模式。双子算法的操作单元是势能面上的两个结构像点(即像点 1 和 2)由这两个像点可以定义一个双子体系。这两个像点之间的距离固定为  $\Delta R$  ( $\Delta R$  的值一般取  $0.005 \text{ \AA}$ ),指向的方向由单位向量  $\hat{N}$  确定,通过旋转双子,改变双子的力( $F$ ),利用前向差分公式(1),来确定双子所在势能面上的局部曲率

收稿日期:2017-10-10;修回日期:2017-10-28

\* 通信作者,Email:zpliu@fudan.edu.cn

(Curv),即本征值。Dimer方法已经在VASP等程序中应用于过渡态寻找。

$$C = \frac{(F_0 - F_1) \cdot N}{\Delta R} \quad (1)$$

我们2012年提出了偏置势—约束双子算法(bias potential-constrained Broyden method, BP-CBD)<sup>[6]</sup>,该方法改进了原有双子算法的两个问题:(1)固定反应模式的问题。在极小点附近原有的双子算法优化给定反应模式,最终会寻找到体系频率最小的平动和转动模式,丢失掉给定反应模式的基本信息,容易导致过渡态寻找失败;(2)效率问题。我们提出的CBD方法采用约束Broyden准牛顿算法,相比于传统的CG方法,提高了寻找体系最小本征模式的效率。

BP-CBD方法的核心是在双子体系的像点1的势能面上引入偏置势函数 $V_N$ ,如公式(2)所定义,这样做的目的是为了能够将振动模式限制在我们初始给定的振动模式 $N_{init}$ 附近。我们一般采用二次函数作为偏置势函数,如公式(3),其中的参数 $a$ 用于调节函数值的大小,具体数值由程序中自动设定为预限制模式的曲率。

$$V_{R1} = V_{real} + V_N \quad (2)$$

$$V_N = -\frac{a}{2} \cdot [(R_1 - R_0) \cdot N_{init}]^2 \\ = -\frac{a}{2} \cdot (\Delta R \cdot N_t \cdot N_{init})^2 \quad (3)$$

## 1.2 随机势能面行走方法(SSW)和反应路径取样(SSW-RS)

势能面的全局搜索是一种非常重要的搜索技术,在物理和化学的许多领域都有广泛的应用,比如研究生物大分子的构象变化,生物体及纳米材料的自主装,固体材料不同压强下的晶体结构等。

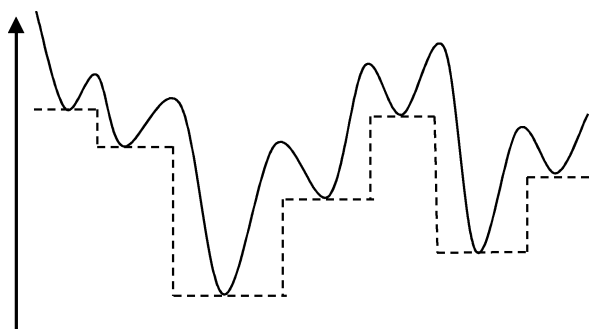


图1 全局最优化方法示意图

虚线表示使用盆地跳跃方法(BH<sup>[7]</sup>)以及拉马克型遗传算法<sup>[8]</sup>通过改变势能面来寻找全局极小。

传统的基于分子动力学的模拟退火方法(simulated annealing, SA<sup>[9]</sup>)可以通过升高体系温度来克

服结构变化所需的能垒,并通过间断的降温动作找到稳定结构。然而,模拟退火方法在搜索那些在自由能面上不是很稳定,甚至在高温区域不会出现的稳定结构时,其效率非常的低。与模拟退火方法相似的,能量极小值跳跃方法(minima hopping, MH)<sup>[10]</sup>同样是基于分子动力学的势能面全局搜索方法,并通过短时间升高温度克服势能面上的能垒。其他被称为全局最优化(GM)的方法中,比如盆地跳跃(basin-hopping, BH<sup>[7]</sup>)以及遗传算法(genetic algorithm, GA<sup>[11]</sup>)以找到势能面上的能量最小点为目的,完全不关心搜索过程中的反应通道。这类方法将原始的势能面变形为只存在一系列能量极小值的势能面,如图1中的虚线所示。这类方法的共同点是需要进行一些剧烈的破坏结构的动作,从而达到从当前的能量最小点跳出去的目的。比如BH方法中通过在任意一个找到的能量极小值点进行一次沿随机方向的剧烈的结构改变(沿体系的每一维度变化平衡键长的40%左右),而遗传算法经常是将两个结构分别拆分再从两个碎片中各取一部分进行重新组合。因此,这类方法的弊端就是会毫无选择地破坏当前结构中所有的化学键,包括那些已经成型的化学键结构。并且需要在初始就提供更多的限制条件来处理共价键分子体系。

我们在2013年提出了一种新型的全局势能面行走方法,即随机势能面行走方法(stochastic surface walking method, SSW)<sup>[12]</sup>。这种新方法是基于偏置势函数靶向分子动力学的思想<sup>[13]</sup>以及Metropolis蒙特卡洛<sup>[14]</sup>发展起来的。通过在势能面上将体系平滑地从一个稳定的结构像点变化为另一个稳定的像点,并通过Metropolis蒙特卡洛方法判断是否接收当前一步的变化。这一方法可以在不需要体系任何先验知识的前提下(比如体系的对称性,材料的组分分布等信息)迅速地搜索复杂的势能面。

SSW方法中的每一个蒙特卡洛步(MC步)包含三个独立的计算部分,分别为(1)沿势能面向上爬山;(2)释放所有偏置势,局部优化;(3)Metropolis蒙特卡洛(MC)。在图2中我们在一个简单的一维势能面上通过从一个稳定结构到另一个稳定结构的过程说明了这样的一个MC步所包含的爬山过程和优化过程。对于化学反应路径的取样(SSW-RS),我们只要稍微改变Metropolis MC部分,使得SSW过程记录下所寻找的反应过程,比如记录初态和终态,作为反应坐标。

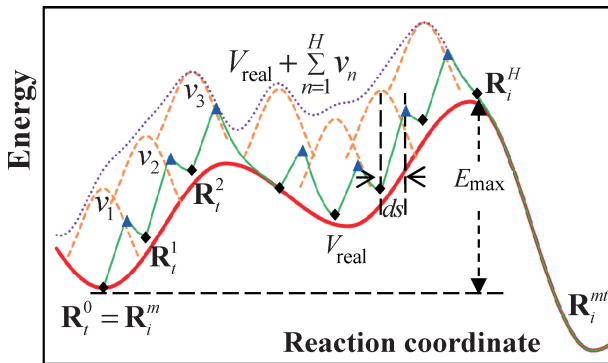


图 2 SSW 方法在一维势能面上的示意图

红、橙、紫、绿四种曲线分别代表真实势能面 ( $V_{\text{real}}$ ), 高斯势函数 ( $v_n, n=1, 2 \dots H$ ) 修改过的势能面, 以及搜索轨迹。

爬山过程是 SSW 算法的核心, 即从结构  $R_i^m$  移动到高能构型  $R_i^H$  的过程 (图 2)。我们使用了如公式 4 所示的一个修正过的势能面  $V_{m \rightarrow H}$ , 其中一系列偏置高斯势函数  $v_n$  一个接一个的沿着随机方向  $N_i^n$  叠加在真实势能面上 ( $n$  是高斯势函数的个数指标,  $n=1, 2 \dots H$ )。  $N_i^n$  总是使用偏置双子旋转方法 (BP-CBD) 根据初始的随机方向  $N_i^0$  进行更新。在公式 4 中,  $\omega$  和  $ds$  分别是高斯函数  $v_n$  的高度和宽度;  $R_i^n$  是在移动轨迹上的第  $n$  个局域极小值点, 即在势能面上添加  $n$  个高斯函数之后产生的新的局域极小值点。在  $R_i^n$  处的势能面由公式  $V_{m \rightarrow H} =$

$V_{\text{real}} + \sum_{k=1}^n v_k$  定义得到。

$$V_{m \rightarrow H} = V_{\text{real}} + \sum_{n=1}^H v_n$$

$$= V_{\text{real}} + \sum_{n=1}^H \omega_n \times \exp \left[ -\frac{((R_i^n - R_i^{n-1}) \cdot N_i^n)^2}{(2 \times ds^2)} \right] \quad (4)$$

总体来说, 从  $R_i^m$  运动到  $R_i^H$  的运动过程是如下三个步骤的一个重复操作: (1) 在第  $n-1$  个局域极小值点  $R_i^{n-1}$  更新振动模式  $N_i^n$ ; (2) 沿着振动模式  $N_i^n$  在势能面上所指向的方向添加新的高斯函数  $v_n$  并且将当前的像点  $R_i^{n-1}$  沿着  $N_i^n$  向前平移  $ds$  的距离 ( $R_i^{n-1} + N_i^n \cdot ds$ ); (3) 在偏置的势能面上优化得到新的局域极小值点  $R_i^n$ 。

高斯势函数的宽度  $ds$  是一个非常重要的参数, 控制着势能面行走的步长。一般来说这个变量的取值范围是 0.2 到 0.6 埃, 也就是化学键键长的 10% 到 40% 左右。如果使用比较大的  $ds$  值, 这一方法就可以迅速的扫描很大一片区域的势能面, 但是这是以损失对反应通道的分辨率为代价实现的。高斯势

函数的最大个数 ( $H$ ) 同样是一个体系相关的参数, 一般  $H$  参数取值范围为 10—15。这样就大约等于在每个 MC 步中总的发生了 4—6 Å 的构型变化。

### 1.3 双端势能面行走过渡态寻找方法 (DESW)

寻找化学反应的过渡态是理论预测化学反应活性的关键。与势能面的局部极小值点不同, 过渡态是势能面上的一阶鞍点, 有一个且仅有一个虚频。寻找过渡态, 一般有两类方法, (一) 单端方法: 从一个初始猜测结构开始, 通过预设的反应模式, 在势能面上逐步向过渡态靠近。这类方法包括拟牛顿方法<sup>[15,16]</sup>, 分区有理函数优化算法 (P-RFO)<sup>[17,18]</sup>, 杂化本征向量跟踪方法<sup>[19,20]</sup>, 双子算法 (Dimer)<sup>[4,5]</sup>, 以及键长限制优化算法<sup>[21]</sup> 等。(二) 双端方法: 同时需要反应物和产物的信息, 通过操作反应两端, 比如建立连接通道, 来逼近过渡态及反应通道。这类方法包括微扰橡皮筋模拟方法 (NEB)<sup>[22,23]</sup>, 双微扰橡皮筋模拟方法 (DNEB)<sup>[24]</sup> 以及绳状模拟方法 (String)<sup>[25,26]</sup>。

我们在 2013 年提出了一种新型的双端势能面行走方法 (double-ended surface walking method, DESW)<sup>[27]</sup>, 用于过渡态寻找。这种新方法和 SSW 方法相同, 都是基于偏置势函数推动的单端行走方法 (BP-CBD)。DESW 方法通过在势能面上的左右两端添加势函数, 将体系平滑地从两个稳定的结构像点朝着中间靠拢, 从而得到路径上的最高能量结构像点。最后, 在该最高能量结构位置, 利用单端算法约束双子算法 (CBD), 来定位反应过渡态。

在双端行走过程中, 更新偏置势函数方向的公式如下公式 (5) 和 (6), 其始终指向反应的方向, 其中  $R$  和  $P$  分别代表反应初态 (reactant) 和终态 (product) 两端的结构构象。

$$N_i^{\text{init}, R} = \frac{P_0^i - R_0^i}{\|P_0^i - R_0^i\|} \quad (5)$$

$$N_i^{\text{init}, P} = \frac{R_0^{i+1} - P_0^i}{\|R_0^{i+1} - P_0^i\|} \quad (6)$$

### 1.4 全局神经网络势函数方法

神经网络 (Neural network, NN) 方法拟合势能面传统上主要用于小分子反应动力学。为了解决复杂材料体系的大体系计算问题, 2007 年 Belher 和 Parrinello 提出了高维神经网络拟合方案<sup>[28,29]</sup> (图 3)。该方案提出了结构因子, 即对称性函数 symmetry function, 描述原子周围环境信息作为神经网络输入参量, 并利用能量原子均分近似, 实现高维神经网络拟合。Belher 和 Parrinello 的代表性工作主要在材料物理领域, 如拟合硅单质的势能面, 计算得到硅在

不同压力下的相图,从理论上预测了和实验相符的不同压力下单晶硅相变产物;拟合碳单质的势能面,研究高压石墨转化为金刚石的动力学行为<sup>[28,29]</sup>。近期,该方法也有开始应用到催化材料领域,如 Artirith 和 Kolpak 也采用 Belher-type 高维神经网络拟合铜金和二氧化碳势能面,结合蒙特卡洛 (Monte Carlo) 模拟,研究了铜金表面合金化能力以及 CO<sub>2</sub> 在不同的铜金团簇表面的还原活性<sup>[30]</sup>。

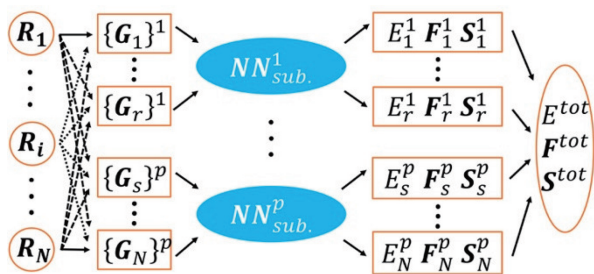


图3 高维神经网络架构

通过将直角坐标  $R$  转化为对称函数  $G$ , 作为输入参数, 再用神经网络, 拟合预测体系总能量  $E^{tot}$ , 力  $F^{tot}$ , 应力  $S^{tot}$ 。值得注意的是, 每一种元素单独形成一个神经网络, 体系总能量由每个原子的能量加和得到。该架构得到的神经网络可以预测不同大小的体系。

我们 2017 年提出了结合 SSW 全局势能搜索和 Belher-Parrinello 神经网络, 建立新型全局势能面神经网络势函数 SSW-NN 的系统方案<sup>[31]</sup>。SSW-NN 提出了全局数据—全局模拟 (global-to-global) 的技术方法, 即基于 SSW 全局势能面搜索技术产生训练大数据集, 通过拟合 SSW 全局势能面数据, 得到的神经网络势能面 (SSW-NN), 用于全局结构和反应预测。由于 SSW 数据集不仅包含稳定结构, 同时包括大量反应过渡态的化学环境, SSW-NN 得到的 NN 势能面具有良好的结构和化学反应的预测能力。目前该方案在 TiO<sub>2</sub> 固体全局势能面搜索和固体相变反应上得到了验证, 发现了动力学稳定的 TiO<sub>2</sub> 大孔材料 Phase-139<sup>[31]</sup>。值得注意的是, 该工作表明相比于 DFT 势能面, 对于大体系的计算模拟, 基于神经网络势函数的 SSW 可以提高 4 个数量级的计算效率, 并保持很好的能量、能垒预测能力。

## 2 软件架构

### 2.1 程序主架构

LSASP 软件的主架构如图 4 所示, 程序采用模块化设计, 包含能量计算和结构搜索两个主要部分, 也附带多种结果的后处理脚本工具。其中能量计算

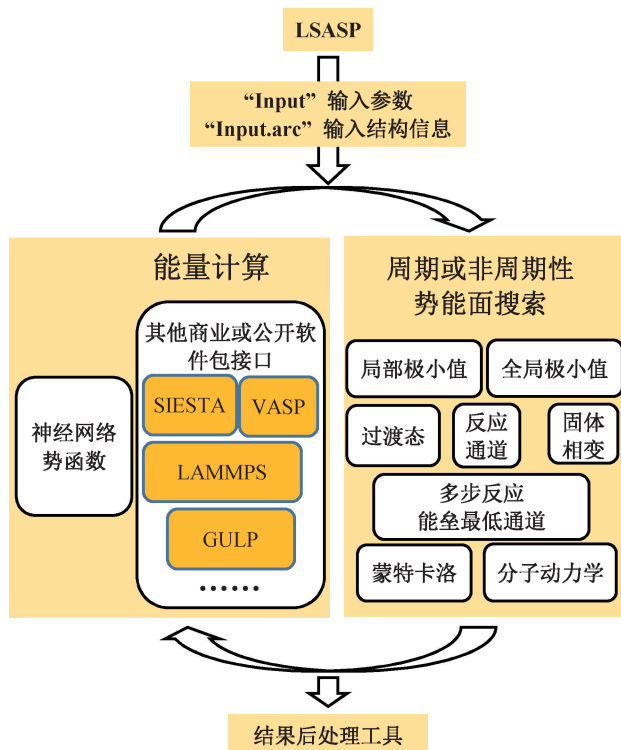


图4 LSASP 软件的主架构

模块, 以神经网络势函数计算模块为主, 也同时包含了与其他商业化或公开电子结构计算程序的接口。在结构搜索模块部分, 既包含了独立开发有自主知识产权的 SSW 全局结构优化, 多种过渡态程序 (CBD, DESW 等), SSW-RS 反应通道取样方法, 同时也包含了常用的几何结构局部优化 (如 BFGS 方法), 分子动力学 (恒温, 恒温恒压系综), 蒙特卡洛模拟模块。结构搜索模块能够处理周期性边界条件, 能自动将不合理的晶体晶胞转化为标准的晶胞 (如晶体晶胞角度大于 120 度或小于 60 度)。

LSASP 软件的输入参数由 input 文件定义, 采用自由格式的文本扫描关键词方式。关键词表包含能量计算模块选择, 神经网络势函数路径, 结构搜索模块选择和具体优化参数等。输入的初始结构由 input.arc 文件读入, 采用标准的 arc 结构文件格式, 其中包含晶体晶胞参数, 原子元素名称, 笛卡尔坐标, 预设的原子电荷 (神经网络势函数有可能采用)。

### 2.2 面向对象类模块化设计

LSASP 软件的神经网络能量计算部分和基于 SSW 的结构搜索部分采用 Fortran 90 面向对象的类 (class) 架构编写。下面以 SSW 核心类实现举例。SSW 的基础类包括结构类 Class\_str, 对结构类 Class\_pair, SSW 全局优化类 Class\_SSW, 单端寻找过渡态类 Class\_TS, 双端寻找过渡态类 Class\_

DESW, 以及自动化 SSW 反应路径搜索 Class\_SSW\_RS 类。他们之间的相互继承和依赖关系如图 5 所示。

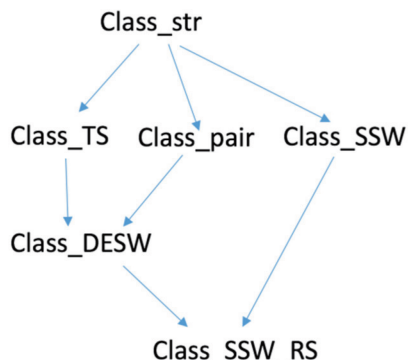


图 5 SSW 核心类的继承和依赖关系

举例来说, 结构类 Class\_str 主要实现了对于特定结构(含晶胞)的所有结构相关简单操作, 比如分数坐标和直角坐标转化, 倒空间计算, 折合坐标和力的转换, 结构的局部优化等。SSW 全局优化类 Class\_SSW, 主要实现了随机模式产生, BP-CBD 软化反应模式, 添加偏置势函数, 蒙特卡洛结构选取, 反应路径结构选取, 并行交换等关键功能。

### 3 进展和展望

LSASP 程序随着我们课题组研究工作的开展, 得到日益完善。目前核心算法相关的模块, 主要是 SSW 和神经网络计算, 均已经在面向对象类框架下重新编写。同时, 与其他大型商业软件包的接口程序, 后处理脚本程序接口也在逐步完善中, 目前已经可以和 VASP, SIESTA, GULP, LAMMPS 等程序对接。下一步的工作包括完善分子动力学模块, 蒙特卡洛模块, 自由能计算所涉及的增强抽样方法等。

总的来说, LSASP 作为大体系原子模拟的软件, 可以明显区别于传统的力场, 以及量化计算软件包, 具有如下特色:

1. LSASP 主要采用新一代高维神经网络全局势函数作为势能面计算工具, 具有精度可比于量化计算, 同时计算效率远远好于量化计算, 能够广泛研究, 包括化学反应等的势能面复杂问题。

2. LSASP 在势能面全局搜索, 反应路径取样等领域具有自有知识产权的计算算法, 核心 SSW 算法取得了国家发明专利, 在固体相变, 表界面反应取样等领域具有竞争优势。

3. LSASP 通过整合目前通用的商业和开源电子结构软件包, 统一了从原始量子力学大数据数据采

集, 到神经网络拟合, 再到基于神经网络势函数的大体系计算模拟的所需要的标准文件格式共享, 大大便利了神经网络势函数的构建和大体系模拟, 有利于未来标准材料数据库以及神经网络势函数库的建设。

**致谢** 本工作得到国家自然科学基金 (21533001、21603035), 上海市科学技术委员会 (08DZ2270500), 上海市浦江人才计划 (16PJ1401200) 的资助, 特此致谢。

### 参 考 文 献

- [1] Gale JD, Rohl AL. The general utility lattice program (GULP). *Molecular Simulation*, 2003, 29(5): 291–341.
- [2] Plimpton S. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 1995, 117(1): 1–19.
- [3] Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2013, 3(2): 198–210.
- [4] Henkelman G, Jónsson H. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *The Journal of Chemical Physics*, 1999, 111(15): 7010–7022.
- [5] Kästner J, Sherwood P. Superlinearly converging dimer method for transition state search. *The Journal of Chemical Physics*, 2008, 128(1): 014106.
- [6] Shang C, Liu ZP. Constrained Broyden dimer method with bias potential for exploring potential energy surface of multi-step reaction process. *Journal of Chemical Theory and Computation*, 2012, 8(7): 2215–2222.
- [7] Wales DJ, Doye JPK. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 1997, 101(28): 5111–5116.
- [8] Turner GW, Tedesco E, Harris KDM, et al. Implementation of Lamarckian concepts in a genetic algorithm for structure solution from powder diffraction data. *Chemical Physics Letters*, 2000, 321(3): 183–190.
- [9] Kirkpatrick S, Gelatt CD, Vecchi MP. Bayesian network reasoning with uncertain evidences. *Science*, 1983, 220: 671–680.
- [10] Goedecker S. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *The Journal of Chemical Physics*, 2004, 120(21): 9911–9917.
- [11] Oganov A R, Glass C W. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *The Journal of Chemical Physics*, 2006, 124(24): 244704.
- [12] Shang C, Liu ZP. Stochastic surface walking method for structure prediction and pathway searching. *Journal of Chemical Theory and Computation*, 2013, 9(3): 1838–1845.

- [13] Iannuzzi M, Laio A, Parrinello M. Efficient exploration of reactive potential energy surfaces using Car-Parrinello molecular dynamics. *Physical Review Letters*, 2003, 90(23): 238302.
- [14] Metropolis N, Rosenbluth AW, Rosenbluth MN, et al. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 1953, 21(6): 1087–1092.
- [15] Broyden CG. A class of methods for solving nonlinear simultaneous equations. *Mathematics of Computation*, 1965, 19(92): 577–593.
- [16] Peng C, Bernhard Schlegel H. Combining synchronous transit and quasi-newton methods to find transition states. *Israel Journal of Chemistry*, 1993, 33(4): 449–454.
- [17] Cerjan CJ, Miller WH. On finding transition states. *The Journal of chemical physics*, 1981, 75(6): 2800–2806.
- [18] Baker J. An algorithm for the location of transition states. *Journal of Computational Chemistry*, 1986, 7(4): 385–395.
- [19] Munro LJ, Wales DJ. Defect migration in crystalline silicon. *Physical Review B*, 1999, 59(6): 3969.
- [20] Kumeda Y, Wales DJ, Munro LJ. Transition states and rearrangement mechanisms from hybrid eigenvector-following and density functional theory. : application to C10H10 and defect migration in crystalline silicon. *Chemical Physics Letters*, 2001, 341(1): 185–194.
- [21] Wang HF, Liu ZP. Comprehensive mechanism and structure-sensitivity of ethanol oxidation on platinum; new transition-state searching method for resolving the complex reaction network. *Journal of the American Chemical Society*, 2008, 130(33): 10996–11004.
- [22] Henkelman G, Jónsson H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *The Journal of Chemical Physics*, 2000, 113(22): 9978–9985.
- [23] Henkelman G, Uberuaga BP, Jónsson H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*, 2000, 113(22): 9901–9904.
- [24] Trygubenko SA, Wales DJ. A doubly nudged elastic band method for finding transition states. *The Journal of Chemical Physics*, 2004, 120(5): 2082–2094.
- [25] Peters B, Heyden A, Bell AT, et al. A growing string method for determining transition states: Comparison to the nudged elastic band and string methods. *The Journal of Chemical Physics*, 2004, 120(17): 7877–7886.
- [26] Weinan E, Ren W, Vanden-Eijnden E. String method for the study of rare events. *Physical Review B*, 2002, 66(5): 052301.
- [27] Zhang XJ, Shang C, Liu ZP. Double-ended surface walking method for pathway building and transition state location of complex reactions. *Journal of Chemical Theory and Computation*, 2013, 9(12): 5745–5753.
- [28] Behler J, Parrinello M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters*, 2007, 98(14): 146401.
- [29] Behler J, Martoňák R, Donadio D, et al. Metadynamics simulations of the high-pressure phases of silicon employing a high-dimensional neural network potential. *Physical Review Letters*, 2008, 100(18): 185501.
- [30] Artrith N, Kolpak AM. Understanding the composition and activity of electrocatalytic nanoalloys in aqueous solvents: A combination of DFT and accurate neural network potentials. *Nano Letters*, 2014, 14(5): 2670–2676.
- [31] Huang SD, Shang C, Zhang XJ, et al. Material discovery by combining stochastic surface walking global optimization with a neural network. *Chemical Science*, 2017, 8(9): 6327–6337.

## Large-scale atomic simulation package (LSASP): program and progress

Liu Zhipan      Shang Cheng

(*Collaborative Innovation Center of Chemistry for Energy Material, Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433*)

**Abstract** Here we introduce a new software, large-scale atomic simulation package (LSASP), developed at Fudan University targeting for exploring the potential energy surface (PES) of complex systems. The key algorithm, the architecture and the latest progress will be described in detail. Different from the traditional force field programs, LSASP integrates common electronic structure calculation packages and novel neural network potential for force evaluation and can carry out a wide range of simulation tasks on potential energy surface exploration, including global structure optimization, automated reaction pathway sampling, the transition state searching, molecular dynamics and Monte Carlo simulation. The stochastic surface walking (SSW) method developed in Liu group play key roles in LSASP PES exploration and is competitive internationally, which can be utilized to achieve a range of functionalities, such as global optimization, transition state search and neural network potential simulation.

**Key words** potential energy surface; global optimization; reaction pathway sampling; LSASP software