

智能优化算法的量子理论纲要

王鹏¹ 辛罡^{2,3}

摘要 针对一些智能优化算法缺乏完备数学物理理论基础的现状, 利用优化问题和量子物理在概率意义上的相似性, 建立优化问题的薛定谔方程, 将优化问题转化为以目标函数为约束条件的基态波函数问题, 同时利用波函数定义了算法的能量、隧道效应和熵, 实现了以波函数为中心的优化问题量子模型. 这一纲要利用了量子物理完备的理论框架, 建立起了优化问题与量子理论广泛的内在联系. 从量子物理的角度回答了优化问题解的概率描述, 邻域采样函数的选择, 算法演化的过程设计, 多尺度过程的必要性等问题. 智能优化算法的量子理论纲要可以作为研究与构造算法的理论工具, 其有效性已得到初步验证.

关键词 优化问题, 优化算法, 量子理论, 波函数, 基态

引用格式 王鹏, 辛罡. 智能优化算法的量子理论纲要. 自动化学报, 2023, 49(11): 2396-2408

DOI 10.16383/j.aas.c190761

Quantum Theory of Intelligent Optimization Algorithms

WANG Peng¹ XIN Gang^{2,3}

Abstract Many metaheuristics, which are based on the metaphor of natural phenomena, are lacking the support of a complete mathematical or physical theory. In this paper, the Schrödinger equation of optimization problems is proposed based on the basic probabilistic similarities between an optimization problem and a quantum system. It transforms the optimization problem into a quantum physics problem with the objective function as the constrained conditions. Meanwhile, the algorithm energy, quantum tunnel effect, and entropy are defined based on the probabilistic interpretation of the wave function. This wave function based quantum physical model establishes a comprehensive internal relationship between optimization problems and quantum theory, and theoretically answers the following questions: probabilistic description of solution of optimization problems, selection of neighborhood function, evolutionary process of optimization, necessity of multi-scale process, etc. Those quantum theories of intelligent optimization algorithm can be used as a theoretical means for the analysis of optimization systems. Its effectiveness has been preliminarily verified.

Key words Optimization problem, optimization algorithm, quantum theory, wave function, ground state

Citation Wang Peng, Xin Gang. Quantum theory of intelligent optimization algorithms. *Acta Automatica Sinica*, 2023, 49(11): 2396-2408

从上世纪 70 年代开始, 智能优化算法经历了持续近 30 年的黄金时期, 其标志就是大量经典的优化算法被提出, 并得到广泛的应用. 1975 年 Holland^[1] 提出了遗传算法 (Genetic algorithm, GA), 这一算法演变为一类重要的算法簇——进化算法.

1983 年 Kirkpatrick 利用 Metropolis 准则^[2] 提出了模拟退火算法 (Simulated annealing, SA)^[3]; Dorigo 于 1992 年在他的博士论文中提出了蚁群算法 (Ant colony optimization, ACO)^[4-5], 这一算法对解决组合优化问题具有较好的效果; 1995 年 Eberhart 等^[6] 基于鸟类的飞行和社会行为提出了粒子群算法 (Particle swarm optimization, PSO), 这些算法的提出为优化算法的研究打下了基础. 进入 21 世纪, 智能优化算法发展的黄金时代已逐渐过去, 但各种新的算法仍然不断出现, 特别是自然启发式算法, 这些优化算法在不同的模型掩盖下存在着“核心”操作同质化的问题. 新提出的优化算法模型大都与现有的算法存在或多或少的相似之处, 这使优化领域的研究者们逐渐意识到, 当前优化算法所面临的挑战不是提出新的算法, 而是为优化问题和优化算法寻找一种合适的统一模型, 智能优化算法

收稿日期 2019-11-03 录用日期 2020-02-07

Manuscript received November 3, 2019; accepted February 7, 2020

西南民族大学中央高校基本科研业务费专项资金项目 (2020NYB18) 资助

Supported by Fundamental Research Funds for the Central Universities, Southwest Minzu University (2020NYB18)

本文责任编辑 王占山

Recommended by Associate Editor WANG Zhan-Shan

1. 西南民族大学计算机科学与技术学院 成都 610225 2. 中国科学院成都计算机应用研究所 成都 610041 3. 中国科学院大学北京 100049

1. School of Computer Science and Technology, Southwest Minzu University, Chengdu 610225 2. Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu 610041 3. University of Chinese Academy of Sciences, Beijing 100049

的研究进入了一个新的时代.

对于优化算法核心规则和策略的研究早已被一些研究者所关注, 粒子群算法的提出者 Kennedy 也意识到优化算法存在统一的简单模型和通用的基本迭代操作, 他针对粒子群算法提出了相应的“骨架”(Bare bone) 算法, Kennedy 认为: “希望利用这种骨架算法揭示随机群体算法之间相似性的奥秘, 并发现新的方法”^[7]. 随后差分进化算法^[8]和烟花算法^[9]的骨架算法也相继提出. 2016 年蚁群算法的提出者 Dorigo 在一份声明中也表达出对于当前不断提出的新的自然算法的谨慎态度. 其他一些研究者, 如 Sörensen 也认为, “现在优化计算领域的研究, 重要的不是提出新的算法而是为优化算法建立普遍适用的规则和策略, 研究优化问题和优化算法中的共性问题”^[10]. 萤火虫算法、布谷鸟算法、蝙蝠算法及鲜花授粉算法等多种优化算法的提出者 Yang 在自己的专著中提醒读者: “不鼓励大家再提出新的优化算法, 这些新算法可能只会分散解决优化中真正具有挑战性和真正重要的问题的注意力”^[11]. 相似的看法在其他一些学者的论述中也有表述, 如文献 [12] 认为和声优化算法本质上就是一种进化策略. 广义上讲遗传算法、粒子群算法、蚁群算法等算法的基本模型都是基于人对世界的主观认识对优化问题进行的建模, 这些模型往往缺乏完备的数学、物理体系的支持, 一定程度上也阻碍了优化算法这一学科的发展.

优化算法的统一模型需要同时具备通用性、深刻性和数学完备性. 20 世纪初量子理论在一大批杰出科学家的努力下获得了快速的发展, 建立起了完备的理论体系, 并在实践中得到了反复的检验. 与其他理论相比, 量子理论深刻地反映了大自然的基本运行规律, 特别是量子理论中对波函数的概率解释与基于概率采样的优化算法具有很大的相似性^[13].

事实上, 量子理论的思想在优化算法中早已得到了应用, 早在 1996 年 Narayanan 等^[14]利用量子理论中的“多宇宙”(Multi-universe) 的观点提出多宇宙衍生遗传算法 (Quantum-inspired genetic algorithm, QIGA), 该算法尝试把量子特性融入到进化算法中去, 后来量子遗传算法的思想也应用于解决组合优化问题^[15]. 随后量子蚁群算法 (Quantum ant colony optimization, QACO)^[16], 量子粒子群 (Quantum-behaved particle swarm optimization, QPSO)^[17] 等受量子理论启发的算法相继提出. 本文作者也于 2013 年提出了一种基于量子谐振子波函数构造的多尺度量子谐振子算法 (Multiscale quantum harmonic oscillator algorithm, MQHOA)^[18],

通过对 MQHOA 算法的研究, 我们发现了薛定谔方程和波函数在优化算法的描述中具有重要作用^[19]. 这些结果可以认为是量子理论与优化问题在概率特性上存在内在一致性的证据.

基于随机行为的智能优化算法通常都放弃了对最优解准确性的要求, 这与量子理论中采用波函数对粒子进行概率化描述的思想是一致的. 粒子群算法的提出者 Kennedy 在著作 *Swarm Intelligence* 中指出“随机性的程度决定了智能的高低”, 这一论断指出了智能的本质是随机性^[20]. 2018 年 11 月李国杰院士在《中国计算机学会通讯》上发表了题为“计算机科学基础理论需要重塑”的卷首语, 他指出“量子力学改变了传统的逻辑定义, 把概率看成了逻辑的内在组成. 在计算机领域, 构造一台完全公理化驱动的自动机也不现实, 而对复杂环境, 我们需要放弃严格逻辑而改用概率逻辑”.

由于量子理论已经过长期的充分发展, 其理论体系及数学表达已非常完备和丰富, 建立了如扩散蒙特卡罗法、路径积分法、微扰理论等基态波函数的求解方法. 本文基于对优化问题概率特性的认识, 采用量子物理理论对优化问题进行研究, 从量子理论的角度尝试解答优化问题的波函数表达、概率采样函数的选择、算法演化过程和多尺度过程的必要性等问题, 希望能为优化算法的分析与研究提供新的视角.

1 优化问题解的概率化定义

智能优化算法, 其主要指基于人类对自然规律的认知与借鉴而形成的智能优化系统, 这些系统大部分都是概率性算法, 其包含了大量概率化的迭代操作, 通过放弃对解的确定性的要求, 从而获得在可接受时间内巨大解空间上的搜索能力. 智能优化算法的角度看, 优化问题解的获得是一个不确定性的问题, 甚至算法无法对得到的解给出任何确定性保证. 优化问题的解采用概率形式来表达更能有效地反映解的特性.

首先来看数学上对优化问题的定义, 为方便讨论, 在本文中以一维目标函数 $f(x)$ 优化问题为例进行讨论. 从数学的角度一维函数优化问题可以定义如下:

针对目标函数 $f(x)$ 在实数定义域 $[a, b]$ 上的优化问题, 就是要找到一个实数 $x_0 \in [a, b]$, 对于任意 $x \in [a, b]$, 满足 $f(x_0) \leq f(x)$, 则称 x_0 为目标函数在实数定义域 $[a, b]$ 全局最优解.

从以上定义可知, 全局最优解是一个确定性的定义, 优化过程理论上需要找到 x_0 这一个确定的全

局最优解. 而对于智能优化算法, 为确保算法的通用性, 优化问题通常视为黑盒问题, 黑盒问题认为: 优化算法对目标函数的表达式 $f(x)$ 是未知的, 优化算法只能通过采样确定从目标函数定义域集合到值域集合中的每一个映射. 优化算法的每一次采样就是对黑盒进行的一次测量, 测量的次数越多所获得的目标函数信息越多.

智能优化算法对目标函数 $f(x)$ 全局最优解的位置同样是不确定的. 事实上算法在运行过程中并不知道它是否找到了全局最优解. 所以采用当前采样点的概率密度函数 $g(x)$ 来描述优化问题的结果更能准确地反映优化问题的概率特性, $P(a \leq x \leq b)$ 表示全局最优解出现在区间 $[a, b]$ 的概率.

$$P(a \leq x \leq b) = \int_a^b g(x) dx \quad (1)$$

对于智能优化算法, 优化问题的解由数学上确定的全局最优解的位置 x_0 转变为描述算法当前采样点的概率密度函数 $g(x)$.

在算法运行前全局最优解 x_0 从算法的角度是可行域中的任意位置, 而在优化算法运行后, 最优解附近的 $g(x)$ 会逐步增加, 最终收敛为一个很小的区域. 这一区域被认为是算法最优解可能存在的区域. 针对最优化问题, 我们通常只关心在最优化解附近的概率分布, 所以通常 $g(x)$ 为紧支撑函数. 这样概率分布会更集中在中心位置, 例如高斯分布、柯西分布、拉普拉斯分布等. 后面本文将采用量子模型说明 $g(x)$ 的概率形式与对目标函数进行势阱逼近的势阱选择有关.

优化问题解的概率定义从理论上建立起了优化问题的解与波函数概率解释之间的联系, 为采用量子波函数对优化问题进行描述提供了可能.

2 优化问题的量子模型

智能优化系统中采样点的运行方式是复杂的, 鉴于不同的算法模型, 很难有一套理论可以完整地给出对于不同智能优化算法动力学行为的描述. 在量子理论中, 粒子的运动状态是由其对应的波函数 $\psi(x)$ 所决定的. 决定波函数的动力学方程是由粒子的质能方程和波动方程联立得到的, 本文希望利用量子力学中的确定的动力学方程代替智能优化算法中复杂且无法确定的运动学方程, 从而为讨论其核心迭代过程和搜索行为提供研究基础.

早在 1926 年 Born^[21] 给出了波函数 $\psi(x)$ 的概率解释. Born 认为波函数模的平方 $|\psi(x)|^2$ 描述了粒子在空间出现的概率分布. 如果我们构造一个描述优化问题的量子化系统, 将全局最优解的概率密

度函数 $g(x)$ 用波函数模的平方 $|\psi(x)|^2$ 来进行表示, 则优化问题就可以转化为求量子系统波函数 $\psi(x)$ 的量子化问题.

2.1 优化问题的薛定谔方程

在量子系统中, 描述势阱中粒子运动状态的薛定谔方程可以用作描述粒子在势阱函数约束下的运动状态, 通过求解薛定谔方程的基态波函数, 可以获得最低能量下的粒子概率分布函数, 含时薛定谔方程为

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = \left(-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x) \right) \psi(x, t) \quad (2)$$

其中, i 是虚数单位, $V(x)$ 为约束粒子的势阱, m 为粒子的质量, \hbar 为普朗克常数, $\hbar = h/(2\pi)$. 利用薛定谔方程就可以求出粒子在被势阱 $V(x)$ 约束下的波函数.

从量子物理的观点看, 为了建立优化问题解的概率密度函数 $g(x)$ 和量子波函数 $\psi(x)$ 关系, 优化问题可以看作是一个能量最低的问题, 是以目标函数 $f(x)$ 为势阱约束下的量子系统, 即 $V(x) = f(x)$. 通过势阱等效, 并令 $S = \hbar^2/(2m)$, 则优化问题的薛定谔方程定义为

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = \left(-S \frac{\partial^2}{\partial x^2} + f(x) \right) \psi(x, t) \quad (3)$$

其中, S 在优化问题中决定着波函数的尺度, 它的值越大波函数的尺度越大, 对应的量子效应越明显, 反之则量子效应越弱, 所以称 S 为尺度系数.

通过薛定谔方程的转换, 优化问题转化为求解在势阱 $f(x)$ 约束下粒子的波函数 $\psi(x, t)$ 的问题, 许多优化问题关注的是 $f(x)$ 的全局最小值, 在物理系统中最小值对应于能量的最小值, 在量子力学中被称为基态. 对于优化问题而言, 其薛定谔方程的基态波函数 $\psi_0(x, t)$ 就描述了全局最优解的概率分布. 优化问题的解不再是一个确定的值, 而是一个由基态波函数所描述的概率分布, 这更符合随机优化算法的特点.

优化问题的薛定谔方程就是对优化问题的量子化描述, 也是优化问题量子模型的理论基础, 它从理论上建立起了基态波函数与优化问题解的概率密度函数之间的关系.

2.2 优化问题的波函数及相关物理量定义

波函数是量子理论中的一个重要的概念, 优化问题的薛定谔方程和优化问题解的量子化定义使波函数成为了描述和研究优化问题的核心理论工具.

波函数与算法的迭代过程紧密相连, 决定着当前算法的采样函数. 通过对波函数的定义, 可以进一步对优化算法的能量、量子隧道效应和熵进行定义和研究.

2.2.1 优化算法波函数的定义

从理论上讲, 优化算法的波函数就是优化算法薛定谔方程的解 $\psi(x, t)$, 它的模方 $|\psi(x, t)|^2$ 就代表了当前解的概率分布. 考虑到薛定谔方程下, 波函数的叠加性, 在实际应用中对于群体算法我们可以将个体采样概率的叠加作为算法当前的波函数, 这样定义的波函数的数学含义为当前全局最优解的分布和当前算法采样分布, 即

$$\psi(x, t) = \sum_{i=1}^k c_i \phi_i(x, t) \quad (4)$$

其中, $\phi_i(x, t)$ 为 t 时刻第 i 个可能薛定谔方程的解, c_i 为该解出现的几率, k 为解的个数. 这种表示方法在 QPSO 算法中得到应用, 其对应的波函数就是 k 个拉普拉斯分布的叠加. 算法波函数在迭代过程中会不断变化, 从而可以直观地观察到算法的收敛过程, 波函数收敛过程的示例参见文献 [18–19].

2.2.2 优化算法的能量

优化问题在量子模型下转变为求量子系统最低能量的问题, 优化算法在迭代过程中能量会逐步减小. 在经典模型下能量就是目标函数的值, 如模拟退火算法当前的最优解就是当前算法的能量, 而在量子模型下优化算法当前的能量是由波函数的概率分布所决定的, 其计算方法为

$$E = \int |\psi(x, t)|^2 f(x) dx - E_0 \quad (5)$$

其中, $\psi(x, t)$ 为算法当前的波函数, E_0 为目标函数值的理论最小值. 由于在量子模型下优化算法的解被波函数以概率分布的形式所描述, 所以优化算法的能量 E 就是算法当前解的数学期望. E_0 在黑盒模型下对于优化算法来说是未知的, 因此在算法实际应用时往往只跟踪前面部分的值 $\int |\psi(x, t)|^2 f(x) dx$ 来研究算法的收敛过程, 这一能量称为相对能量.

当算法迭代到某一尺度的基态时, 算法能量达到这一尺度的最小值, 这时的能量称为零点能, 尺度 S 下的零点能 E_0^S 为

$$E_0^S = \int |\psi_0^S(x, t)|^2 f(x) dx - E_0 \quad (6)$$

其中, $\psi_0^S(x, t)$ 为尺度 S 下的基态波函数, 通常尺度越小零点能会越小, 理论上讲零点能不能等于 0, 零

点能的存在是优化算法量子特性的表现, 也再一次说明优化问题的解只能用概率形式的波函数来描述.

2.2.3 优化算法的量子隧道效应

在量子系统中粒子能通过势垒贯穿出现在经典禁区, 我们称之为隧道效应, 这是量子系统所特有的现象. 如图 1 所示, 在一个双阱势能函数中, 不同量子特性的系统, 对应了不同尺度的波函数. 而粒子从 X 移动到 X' 出现在禁区的几率为图中阴影部分的面积.

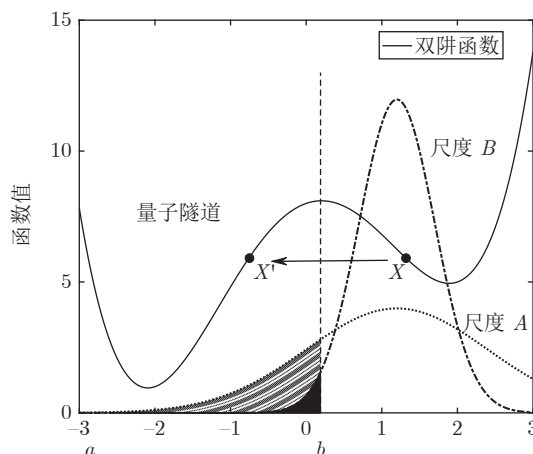


图 1 量子隧道效应示意图

Fig. 1 Schematic diagram of quantum tunneling effect

在优化问题的量子模型下隧道效应是算法跳出局部最优的一个基本机制^[22], 也是优化问题概率性和波动性的表现, 优化算法针对任意区域 $[a, b]$ 发生隧道效应的概率 P_{ab} 可以利用波函数进行计算, 即

$$P_{ab} = \int_a^b |\psi(x, t)|^2 dx \quad (7)$$

在一些经典的优化算法中通常会包含一些差解接受机制, 如模拟退火算法的 Metropolis 准则、遗传算法的变异操作等, 其目的都是为了使算法有机会跳出局部最优区域. 优化问题在量子模型下由波函数的概率解释所引起的隧道效应为优化算法跳出局部最优区域提供了一个“自然”机制, 优化算法并不需要刻意地采用某种跳出局部最优的机制, 而只需要按照波函数的概率分布进行迭代演化.

2.2.4 优化算法的熵

优化问题的波函数可以对优化算法的迭代过程进行跟踪, 但对于高维目标函数其波函数无法直观地表达; 同时对于迭代过程通常希望能给出一个具有明确物理意义的量来研究算法的收敛过程. 所以根据波函数的概率解释, 类比熵的定义可以定义优化算法的熵 H , 即

$$H = - \int |\psi(x, t)|^2 \log |\psi(x, t)|^2 dx \quad (8)$$

优化算法熵的物理意义为当前全局最优解的确定度, 优化算法熵在每一次迭代时都有一个确定的值, 它的值是由当前波函数所决定的, 通常随着算法的收敛, 算法熵的值会逐步减小, 熵的值越小说明算法针对全局最优解的确定度越大, 因此算法熵代表了优化算法对当前最优解确定性的认识, 也可以认为是对高维波函数的降维参数。

优化算法的熵随迭代次数变化的曲线称为熵收敛曲线, 熵收敛曲线能反映算法对解确定度的收敛情况, 与传统收敛曲线相比具有更深刻的物理意义. 熵收敛曲线可以作为优化问题量子模型下研究算法收敛过程的数学工具。

3 优化问题在量子理论下的探索

3.1 薛定谔方程下的随机搜索过程

在将优化问题引入量子系统之后, 就可以使用量子物理的手段对优化问题的求解过程进行研究. 我们将通过量子理论中求解系统基态波函数的一种重要方法, 扩散蒙特卡罗法 (Diffusion Monte Carlo, DMC)^[23], 分析在量子模型下群体优化算法的随机行为。

3.1.1 优化问题薛定谔方程与扩散方程的同构

扩散蒙特卡罗的基础是基于薛定谔方程与扩散方程的同构. 为了构造方程的同构, 将式 (3) 中的时间 t 替换为 τ , $\tau = it/\hbar$, 此时含时薛定谔方程转化为实域的扩散方程, 即

$$\frac{\partial \psi(x, \tau)}{\partial \tau} = S \frac{\partial^2 \psi(x, \tau)}{\partial x^2} - f(x) \psi(x, \tau) \quad (9)$$

同时, 一维无热源的热传导方程可以写为

$$\frac{\partial u(x)}{\partial \tau} = v \frac{\partial^2 u(x)}{\partial x^2} \quad (10)$$

其中, v 为扩散系数, $u(x)$ 为温度的分布. 对照含时的薛定谔方程和一维热传导方程可以发现, 这两个方程具有相同的结构. 特别地, 当系统未被约束时 $f(x) = 0$, 此时优化问题的约束态量子模型退化为自由粒子模型, 即

$$\frac{\partial \psi(x, \tau)}{\partial \tau} = S \frac{\partial^2 \psi(x, \tau)}{\partial x^2} \quad (11)$$

此时优化问题的薛定谔方程从形式上与扩散方程完全相同, 这说明波函数 $\psi(x, \tau)$ 随时间的演化过程和温度的分布 u 随时间的演化具有相似的行为. 而扩散蒙特卡罗方法则是通过构造随机粒子的扩散

行为, 模拟薛定谔方程下粒子的行为, 实现波函数的演化过程。

3.1.2 优化问题薛定谔方程的求解

在扩散蒙特卡罗方法中, 因为优化算法的时间演化过程可以采用粒子的扩散过程来进行描述, 我们可以通过向薛定谔方程引入一定的能量位移来影响量子模型下优化算法的收敛迭代过程. 通常用含时薛定谔方程来表示, 即

$$\frac{\partial \psi(x, \tau)}{\partial \tau} = S \frac{\partial^2 \psi(x, \tau)}{\partial x^2} - [f(x) - E_R] \psi(x, \tau) \quad (12)$$

这里 $\psi(x, \tau)$ 为随时间演化的波函数, E_R 为参考能量, 参考能量可以认为是优化算法当前获得的最低能量, 随着算法迭代的进行参考能量会逐步下降。

假设 $f(x)$ 在 $x \rightarrow \infty$ 是有限的, 粒子被限制在一个有限的空间内, 通过分离变量法求解, 方程的解为

$$\psi(x, \tau) = \sum_{n=0}^{\infty} c_n \phi_n(x) e^{-(E_n - E_R)\tau} \quad (13)$$

其中, E_n 为量子化的能级. 当 $E_R = E_0$ 时, E_0 为系统的基态能量, 其他非基态波函数所对应的指数部分都会随时间演化变为无穷小, 只保留基态所对应的指数项, 所以方程变为

$$\lim_{\tau \rightarrow \infty} \psi(x, \tau) = c_0 \phi_0(x) \quad (14)$$

式 (14) 说明当 $E_R = E_0$ 时对应量子系统的初始波函数 $\psi(x, 0)$ 一定会演化收敛为基态波函数 ϕ_0 , 这意味着当参考能量等于基态能量时无论选择什么波函数作为初始波函数, 最终都能收敛到基态波函数. 而当 $E_R > E_0$ 时, 波函数会随着时间的演化不断发散, 当 $E_R < E_0$ 时, 波函数会随着时间的演化不断耗散. 扩散蒙特卡罗据此设计了基于参考能量的“生灭过程”, 让不符合条件 (会使得波函数发散或耗散) 的采样点不继续繁殖, 通过参考能量调节随机行走的粒子的扩散行为, 最终影响波函数的演化结果, 从而获得低能量约束条件对应的基态波函数 (粒子分布状态)。

量子模型下波函数的时间演化从理论上说明了薛定谔方程下优化算法的收敛行为, 优化算法的迭代过程就是参考能量不断向基态能量的逼近过程, 当参考能量等于基态能量时就能确保收敛到基态波函数。

3.1.3 优化问题波函数的构造

从热力学的观点来看, 扩散过程是大量粒子的一种群体行为, 优化问题在量子模型下与扩散方程

的相似性从量子理论的角度解释了大量优化算法都采用群体策略的原因. 薛定谔方程的时间演化说明, 通过随机生成初始的波函数, 可以通过模拟群体的扩散行, 随时间向系统的基态波函数演化. 因此参考式 (4) 和式 (13) 优化问题的初始波函数可以采用如下的构造方法:

$$\psi(x, \tau) = \sum_{i=1}^k c_i \psi_0(x - x_i) \quad (15)$$

这一构造方法采用在 k 个不同的位置以 x_i 为中心生成 k 个预测的可能基态波函数 $\psi_0(x - x_i)$, 其出现几率为 c_i , 并以 $|\psi_0(x - x_i)|^2$ 为邻域采样函数分别生成 k 个新的可能位置 x'_i . 在演化迭代过程中, 通过比较采样值与 E_R 的关系, 逐步调整波函数演化的过程, 使参考能量向 E_0 逼近, 在扩散蒙特卡罗中, 参考能量是种群数量相关的一个参考量.

相似地, 在许多优化算法中, 可以将当前最好的解作为对参考能量的估算. 例如 PSO 算法在迭代过程中参考群体历史上的最优解和个体历史上最优解, 事实上就是以群体历史上的最优解和个体历史上最优解为参考能量, 对种群个体进行调节的过程.

扩散蒙特卡罗方法是可以借鉴研究优化算法的随机过程的一种物理方法, 除此之外许多量子力学中求解基态波函数的方法, 都可以用于分析随机优化过程.

3.2 简单势阱对目标函数的逼近

虽然可以利用扩散蒙特卡罗之类的方法对优化问题的薛定谔方程进行求解, 但基于以下原因我们无法直接获得一个复杂目标函数 $f(x)$ 所对应的基态波函数, 需要通过目标函数进行近似来获得基态波函数的解.

1) 对于复杂目标函数 $f(x)$ 其对应的基态波函数也是一个复杂的概率密度函数.

2) 薛定谔方程非常难解, 只有对一些简单的势阱才能准确地获得波函数的解析解.

3) 对于优化问题来说, 只需要知道在全局最优解附近的概率分布, 因此可以采用较为简单的势阱模型对复杂目标函数进行逼近.

Taylor 展开是关于某一个点 x_0 展开的, 它可以在点 x_0 附近对目标函数进行逼近. 假设目标函数 $f(x)$ 满足 Taylor 展开的条件, 在全局最优极值点 x_0 的 Taylor 展开为

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \quad (16)$$

对于满足 Taylor 展开的任意目标函数 $f(x)$ 其零阶、一阶和二阶 Taylor 展开分别为 (函数在极值点的一阶导数的值 $f'(x_0) = 0$)

$$f^0(x) = f(x_0) \quad (17)$$

$$f^1(x) = f(x_0) + f'(x_0)(x - x_0) = f(x_0) \quad (18)$$

$$f^2(x) = f(x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 \quad (19)$$

目标函数 $f(x)$ 的零阶和一阶 Taylor 展开均为常数, 对应的量子模型为自由粒子. 目标函数的二阶 Taylor 展开为谐振子势能, 这种势阱经常作为平衡位置附近复杂振动的一种近似, 其对应的基态波函数为高斯分布. 如 MQHOA 算法就是在这种势阱逼近模型下所构造的算法, 采用目标函数的二阶近似已能获得较好的优化性能^[24-29]. 更高阶的 Taylor 展开所对应势能的薛定谔方程很难解出一个简单的概率分布作为基态波函数, 所以通常不采用更高阶的 Taylor 展开作为目标函数的逼近.

在采用简单势阱对优化问题近似逼近后, 相应地, 可以通过式 (15), 利用不同势阱下的基态波函数构造算法的邻域采样函数, 以获得不同的算法性能. 同时 Taylor 展开也不是唯一的近似方法, 一些算法也会采用其他的逼近方法, 例如 QPSO 算法采用 δ 势阱对目标函数进行逼近^[30]. 也可以采用方势阱对目标函数进行逼近, 则优化算法的邻域采样分布为余弦分布.

不同优化算法邻域采样函数的选择, 可以在量子理论下解释为对目标函数的某种势阱逼近的基态波函数. 不同的邻域函数的构造连同不同的参考能量的选择策略共同影响了优化算法的性能.

3.3 多尺度过程的量子化解释

多尺度过程在优化算法中往往是一个必不可少的过程, 大多数优化算法都会引入采样步长收缩策略实现从全局搜索向局部搜索的变化. 例如模拟退火算法中温度的逐步下降过程; 粒子群算法中惯性权重的变化过程; 蚁群算法路径构建时信息素和距离项本质上是在实现全局搜索和局部搜索的平衡. 多尺度过程是一个性能良好的全局优化算法所必须的过程, 我们希望采用量子理论中的不确定性原理对优化算法中的多尺度过程加以证明.

不确定性原理 (Uncertainty principle) 是 Heisenberg 在 1927 年首次提出的, 不确定性原理认为粒子的位置和动量是一对不相容的观察量, 我们不能同时对其进行精确测量. 量子理论中不确定

性原理的一般性表述如下: 对于两个可观察物理量 (Observables) A 和 B 所对应的算符 \hat{A} 和 \hat{B} 不对易 (Non commuting).

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0 \quad (20)$$

则这两个物理量不能同时精确测量, 即

$$\sqrt{(\Delta A)^2(\Delta B)^2} \geq \frac{1}{2} |[\hat{A}, \hat{B}]| \quad (21)$$

其中, ΔA 和 ΔB 为测量精度.

类似地, 为了证明优化问题的不确定性原理, 需要首先定义位置算符和尺度算符. 优化问题的量子模型认为优化问题的解是采用波函数 $\psi(x)$ 来描述的, 在这一前提下优化问题的解不再是一个确定的值而是一个概率分布, 这使得优化问题的解具有位置和尺度 (尺度为频率的倒数) 两个特性. 在量子理论中物理观察量算符是通过其平均值来定义的, 优化问题的波函数代表了全局最优解的概率分布, 因此全局最优解位置的平均值 \bar{L} 可以利用波函数得出, 即

$$\bar{L} = \int \psi^*(x) \hat{L} \psi(x) dx = \int \psi^*(x) x \psi(x) dx \quad (22)$$

因此优化问题的位置算符 \hat{L} 可以表示为

$$\hat{L} = x \quad (23)$$

同理, 全局最优解的频率算符 $\hat{\omega}$ 可以通过解的频率平均值 $\bar{\omega}$ 得到:

$$\bar{\omega} = \int \hat{\omega}^* \omega \hat{\omega}(\omega) d\omega \quad (24)$$

其中, $\hat{\omega}(\omega)$ 为优化问题的解 $\psi(x)$ 的傅利叶变换, 即

$$\hat{\omega}(\omega) = \int \psi(x) e^{-i\omega x} dx, \quad \psi(x) = \int \hat{\omega}(\omega) e^{i\omega x} d\omega \quad (25)$$

为了得到优化问题解的频率算符, 利用优化算法的波函数的傅利叶变换对频率平均值计算式进行如下的数学变换:

$$\begin{aligned} \bar{\omega} &= \int \hat{\omega}^*(\omega) \omega \hat{\omega}(\omega) d\omega = \\ &= \iint \psi^*(x) e^{i\omega x} \omega \hat{\omega}(\omega) dx d\omega = \\ &= \iint \psi^*(x) \left(-i \frac{\partial}{\partial x}\right) e^{i\omega x} \hat{\omega}(\omega) dx d\omega = \\ &= \int \psi^*(x) \left(-i \frac{\partial}{\partial x}\right) \psi(x) dx \end{aligned} \quad (26)$$

因此, 优化问题解的频率算符 $\hat{\omega}$ 为

$$\hat{\omega} = -i \frac{\partial}{\partial x} \quad (27)$$

根据量子理论中不确定性关系的一般定义, 优化算法的解的位置算符和频率算符之间的对易关系为

$$\begin{aligned} [\hat{\omega}, \hat{L}] \psi(x) &= (\hat{\omega} \hat{L} - \hat{L} \hat{\omega}) \psi(x) = \\ &= \left[\left(-i \frac{\partial}{\partial x}\right) x - x \left(-i \frac{\partial}{\partial x}\right) \right] \psi(x) = \\ &= -i \frac{\partial(x\psi(x))}{\partial x} + ix \frac{\partial(\psi(x))}{\partial x} = \\ &= -i\psi(x) \end{aligned} \quad (28)$$

因此,

$$[\hat{\omega}, \hat{L}] \neq 0 \quad (29)$$

这一结论说明在量子模型的解释下, 任意优化算法在一个尺度下都不能同时精确获得优化问题的解 $\psi(x)$ 的频率信息和位置信息. 根据傅利叶分析的原理, 大尺度下的搜索可以有效获得目标函数的频率信息 (全局信息), 但位置精度会较低, 小尺度下的搜索具有较低的频率精度, 但具有较高的位置精度 (局部信息).

优化问题的不确定性关系可以表述为: 单一尺度的搜索不可能同时获得精确的全局搜索能力和局部搜索能力, 多尺度过程是优化算法的一个必要过程. 多尺度过程在一些量子算法, 如量子退火算法中也是一个重要的过程, 这类方法构造的物理系统能不断降低体系的量子效应, 从而实现多尺度优化^[31-32], 这也证明了优化算法中广泛存在的多尺度过程的必要性.

4 量子理论下智能优化算法的实现

为验证量子理论框架对智能优化算法研究的有效性, 本节将基于前面内容中的理论内容简述一种基于扩散蒙特卡罗方法的智能优化算法. 在第 2 节中, 智能优化系统中采样点的运动行为被假设能够被薛定谔方程所描述, 建立起了智能优化算法的量子力学基础; 在第 3 节中, 我们提到量子力学中的蒙特卡罗方法可以用作模拟扩散方程中粒子的运动行为, 从而求解薛定谔方程. 蒙特卡罗方法是一种基于薛定谔方程和扩散方程的同构性, 利用随机行走模拟粒子受扩散方程约束下的物理运动过程对优化问题进行求解的优化方法, 其基础迭代过程可以认为是量子理论下智能优化系统的基础迭代过程. 在此基础上, 通过算符的方法证明了算法多尺度过程的必要性, 通过目标函数的 Taylor 二阶近似, 将

扩散蒙特卡罗这样求解基态问题的方法转换为求解全局最优问题. 因此, 在本节中将首先简述其基础迭代过程, 在此基础上融入前述内容中的多尺度过程和 Taylor 二阶近似下的中值替换策略, 构造出新的算法, 并通过 CEC2013 测试集对该算法进行简单的测试.

4.1 基于蒙特卡罗方法的基础迭代过程

在扩散蒙特卡罗方法中, 经 Wick 转动 ($\tau = it/\hbar$) 之后, 时变薛定谔方程下粒子的概率幅度可以通过无数条可能轨迹上的积分来计算. 通过路径积分的方式, 薛定谔方程的解可以表达为

$$\Psi(x, \tau) = \lim_{N \rightarrow \infty} \int_{-\infty}^{\infty} \left(\prod_{j=0}^{N-1} dx_j \right) \prod_{n=1}^N W(x_n) = P(x_n, x_{n-1}) \Psi(x_0, 0) \quad (30)$$

其中, 从时刻 0 到时刻 τ 平均分为 N 个部分, $\Delta\tau = \tau/N$ ($N \rightarrow \infty$); 其中与空间中粒子动能项相关的蒙特卡罗采样方程为

$$P(x_n, x_{n-1}) = \left(\frac{m}{2\pi\hbar\Delta\tau} \right)^{\frac{1}{2}} \exp \left[-\frac{m(x_n - x_{n-1})^2}{2\hbar\Delta\tau} \right] \quad (31)$$

其中, 与粒子扩散过程中出现概率相关的权重函数为

$$W(x_n) = \exp \left[-\frac{[f(x_n) - E_R] \Delta\tau}{\hbar} \right] \quad (32)$$

路径积分的公式描述了粒子在扩散反应过程中的运动和密度的变化, 因其可以被蒙特卡罗方法求解, 故称之为扩散蒙特卡罗方法. 为了获得 $\Psi(x, \tau)$, 扩散蒙特卡罗方法包含了扩散方程采样和权重函数计算两个阶段

$$\underbrace{\Psi(x_0, 0)}_{\text{初始态}} \rightarrow \cdots \underbrace{P(x_n, x_{n-1})}_{\text{蒙特卡罗采样}} \underbrace{W(x_n)}_{\text{权重函数计算}} \cdots \rightarrow \underbrace{\Psi(x_N, \tau)}_{\text{终态}} \quad (33)$$

其中, 当存在初态 $\Psi(x_0, 0)$ 时, 终态 $\Psi(x_N, \tau)$ 可以被由随机扩散 $P(x_n, x_{n-1})$ 和权重函数 $W(x_n)$ 所组成的随机过程所获得; 在随机扩散环节, 新的粒子根据蒙特卡罗采样方程产生; 在概率函数分配环节, 较差的粒子“死亡”, 较好的粒子产生更多的新粒子, 这便是大家所熟知的扩散蒙特卡罗方法中的“生灭”过程.

扩散蒙特卡罗方法将求解含时薛定谔方程约束

态的问题转化为一个计算机可执行的随机算法过程. 可以认为薛定谔方程约束下智能优化系统的随机迭代过程由粒子的蒙特卡罗采样行为和权重函数计算两个部分所组成.

4.2 基于多种群的多尺度量子谐振子算法实现

文献 [33] 实现了多种群的叠加态量子谐振子算法 (Multiple-population-based superposition state MQHOA, MPSS-MQHOA), 其利用多种群的思想对蒙特卡罗迭代过程进行扩展, 并针对数值优化中的全局优化问题进行调整, 在算法中融入多尺度过程和 Taylor 二阶近似下的中值替换策略. 现简述其主要策略如下.

1) 参考能量的选择: 在蒙特卡罗方法中, E_R 参考能量会随着时间不断的动态调整, 其目的是通过分支过程平衡种群的规模. MPSS-MQHOA 为了方便计算将种群规模 ps 设置为常数, 在优化过程中, 往往希望表现较差的粒子被缓慢地排除, 以维持整个种群的多样性并避免算法早熟的问题. 因此将整个种群中最差的粒子, 设置为参考能量. 与此同时, 算法为了避免早熟, 在每轮迭代中有且仅有一个最差的子种群会被淘汰. 其基于适应度值的参考能量 E_R 可计算为

$$E_R = \max \{f_b(x_1^b), f_b(x_2^b), \dots, f_b(x_g^b)\} \quad (34)$$

其中, $f_b(x_g^b)$ 是第 g 个子种群中适应度值最好的粒子的适应度值.

2) 多种群的蒙特卡罗采样: 在该过程中, 每个子种群中的粒子根据式 (31) 产生, 来确定其在 n 时刻的位置, 由此, 可以给出 g 个子种群的蒙特卡罗采样方程 $P_{mp}(x'_i, x_i)$ 为

$$P_{mp}(x'_i, x_i) = \sum_{i=1}^g \sum_{j=1}^{ps_i} \frac{1}{\sqrt{2\pi\sigma_s}} e^{-\frac{(x'_i - x_i^b)^2}{2\sigma_s^2}} \quad (35)$$

其中, g 是子种群的个数, ps_i 是第 i 个子种群的规模, 其可以通过权重函数的计算动态调整, σ_s 是当前的搜索尺度, x_i^b 是在第 i 个子种群的最优粒子 x_i^b 附近产生的第 j 个粒子.

3) 子种群规模的调整: 在该过程中, 根据式 (32), 算法中子种群的权重函数 $W_{mp}(x_g^b)$ 算法通过第 g 个种群的最优值 x_g^b 计算. 为实现“生灭过程”所获得在优势区域的粒子数量的积聚, 在多种群粒子系统中, 为缓解数值优化问题中适应度值的波动会导致权重函数剧烈的变化这一问题, 适应度值的排序将被用于计算子种群的规模, 算法的子种群规模可由下式计算:

$$ps_i^{fr} = ps \cdot \frac{W_{mp}(x_i^b)}{\sum_{i=1}^g W_{mp}(x_i^b)} = ps \cdot \frac{[(Rank^f(x_i^b) - Rank^f(E_R))]^\alpha}{\sum_{i=1}^g [(Rank^f(x_i^b) - Rank^f(E_R))]^\alpha} \quad (36)$$

其中, $ps = \sum_{i=1}^g ps_i$ 是所有种群中粒子数的和, α 是一个增加种群规模差异性的参数, 通常设置为 [1, 2, 3]. 对于最差的子种群, $p_i^{fr} = 1$. 适应度值最差的子种群会在下一轮迭代中被其他种群的平均位置所替代, 可以认为是一种利用谐振子势阱的基态分布中心对算法的子种群进行修正的操作, 该操作如下:

$$\arg \max \{f_b(x_1^b), f_b(x_2^b), \dots, f_b(x_g^b)\} \leftarrow \left(\frac{1}{g-1} \sum_{i=1}^{g-1} x_i^1, \dots, \frac{1}{g-1} \sum_{i=1}^{g-1} x_i^l \right) \quad (37)$$

MPSS-MQHOA 算法的伪代码如下:

- 1: 初始化 $g, ps, LB, UB, \sigma_s = UB - LB$
- 2: 在 $[LB, UB]$ 范围内随机产生 g 个子种群粒子位置 $x_i (i = 1, \dots, ps_g)$
- 3: 计算所有粒子位置 x_i 的标准差 σ_{ps}
- 4: **while** (未达到终止条件)
- 5: **while** ($\sigma_{ps} > \sigma_s$ 基态判据)
- 6: 根据式 (34) 选择参考能量 E_R
- 7: 根据式 (36) 计算并调整子种群的粒子数 ps_i
- 8: **for** $i = 1$ to g **do**
- 9: 根据式 (35) 为第 i 个子种群产生 ps_i 个粒子 x'_i
- 10: 当 $f(x'_i) < f(x_i)$ 时更新种群中的部分粒子
- 11: **end for**
- 12: 根据式 (37) 淘汰最差的子种群
- 13: **end while**
- 14: 缩小搜索尺度 $\sigma_s = \sigma_s / C_r$
- 15: **end while**
- 16: 输出 $x^{\text{best}}, f(x^{\text{best}})$

在初始化步骤中, 在整个搜索空间中随机设置 g 个子种群的位置, 并将搜索规模 σ_s 设置为 UB 和 LB 之间的值 (搜索空间的上界和下界). 算法的迭代过程主要包括粒子密度调节和蒙特卡罗采样两个阶段. 首先, 根据式 (36) 计算每个子种群的种群大小 ps_i . 随后, 使用式 (35) 生成 ps_i 粒子 x'_i , 当 $f(x_i) < f(x'_i)$ 时, 粒子被更新. 粒子系统的最差子种群在该循环结束时被消除. 在尺度调整步骤中, 搜索尺度 σ_s 随着系数 C_r 逐渐减小. 迭代过程不断循环直到满足终止标准, 输出最优值.

4.3 实验及分析

为了验证本文理论对算法设计的有效性, 本节选取了基于种群方差的自适应微分进化 (Population's variance-based adaptive differential evolution, PVADE)^[34]、标准粒子群优化 (SPSO2011)^[35]、QPSO^[17]、失败者出局竞标赛烟花算法 (Loser-out tournament based fireworks algorithm, LoTFWA)^[36] 作为对照组与 MPSS-MQHOA 进行比较.

所有的算法将在 CEC2013 标准函数测试集^[37] 上进行实验, 每组实验重复 51 次, 单次运算的最大迭代次数为 $10000 \times D$, 其中 D 为维度, 实验维度设置为 30. 测试集一共包括 28 个测试函数, 根据函数结构不同可划分为单模函数 ($F1 \sim F5$), 多模函数 ($F6 \sim F20$), 组合函数 ($F21 \sim F28$). 所有函数的定义域为 $[-100, 100]$. 实验环境为 Intel(R) Core(TM) i7-1160G7 CPU @ 1.20 GHz 2.11 GHz, 16.0 GB, 程序采用 MATLAB2019 编写. 其中 MPSS-MQHOA 的粒子数设置为 $ps = 300$, 子种群个数 $g = 30$, $\alpha = 3$, $C_r = 1.2$; 其他算法参数根据其文献设置. 实验结果见表 1, 其中 Mean 为误差均值, Std. 为解的标准差, 最优的实验结果在表中用粗体标出.

在单模函数上, PVADE 表现出了优异的性能, 在除 $F5$ 外的所有函数上都排名第一. SPSO2011 的表现略好于 MPSS-MQHOA, 在单模函数上排名第二, MPSS-MQHOA 排名第三. QPSO 和 LoTFWA 除了在 $F1$ 上表现较好, 在其他四个函数上都落后于其他三种算法, 并列排名第四. 在多模函数和组合函数上, 根据平均排名, MPSS-MQHOA 排名第一, LoTFWA 排名第二, PVADE 表现出的性能较为均衡, 取得第三的排名成绩. SPSO2011 和 QPSO 表现较差, 分别排名第四和第五.

总的来说, 如表 1 所示, PVADE 在 30 维问题上 28 个函数中的 11 个函数获得了最小的平均误差. 在单模函数上, MPSS-MQHOA 和 LoTFWA 的平均误差排名远落后于 PVADE. 但这三种算法在所有 28 个函数上的平均排名却比较接近, 说明后两种方法在多模函数和组合函数上具有良好的性能. 同时, 具有多种群的 MPSS-MQHOA 在 30 维问题中获得了 13 个函数的最小平均误差, 其平均排名为第一. 该实验结果证明了基于量子框架所设计的算法相比于其他智能优化算法具有一定的竞争力, 也间接证明了本文理论对算法设计的有效性.

5 优化问题与量子理论相关性的讨论

在前面的各节中, 我们从不同角度分析了优化

表 1 PVADE, SPSO2011, QPSO, LoTFWA 和 MPSS-MQHOA 在 30 维 CEC2013 测试集下平均误差和标准差的对比实验, 所有算法的终止条件为 $\text{MaxFES} = 300\,000$, 所有实验独立重复 51 次

Table 1 Comparison of mean errors and standard deviations of PVADE, SPSO2011, QPSO, LoTFWA, MPSS-MQHOA, under 30D CEC2013 benchmark functions. The stopping condition for all schemes is set at $\text{MaxFES} = 300\,000$. The experiments are repeated 51 times individually

F.	PVADE		SPSO2011		QPSO		LoTFWA		MPSS-MQHOA	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
1	0.00E + 00	0.00E + 00	0.00E + 00	0.00E + 00	0.00E + 00	0.00E + 00	0.00E + 00	0.00E + 00	0.00E + 00	0.00E + 00
2	4.24E + 04	2.74E + 04	1.06E + 05	4.90E + 04	1.90E + 07	7.65E + 06	2.13E + 06	7.21E + 05	8.96E + 05	2.14E + 05
3	1.39E + 06	4.58E + 06	5.64E + 07	8.68E + 07	3.98E + 08	8.77E + 08	3.94E + 07	7.23E + 07	3.01E + 06	3.21E + 06
4	1.11E + 02	9.05E + 01	1.97E + 03	7.77E + 02	1.20E + 04	2.77E + 03	7.03E + 03	1.97E + 03	1.69E + 04	3.56E + 03
5	1.12E - 04	3.35E - 04	4.72E - 04	3.75E - 05	1.77E - 05	1.51E - 05	1.60E - 02	1.41E - 02	2.40E - 03	3.61E - 04
AR. 1 ~ 5	1.20	1.60	2.40	2.20	3.20	3.20	3.20	3.20	3.00	2.80
6	4.73E + 01	2.49E + 01	1.46E + 01	1.86E + 01	3.87E + 01	1.95E + 01	1.94E + 01	1.25E + 01	2.42E + 01	2.18E + 01
7	7.33E + 00	5.53E + 00	4.32E + 01	1.44E + 01	4.79E + 01	1.57E + 01	5.40E + 01	1.23E + 01	2.63E + 01	1.04E + 01
8	2.09E + 01	4.64E - 02	2.09E + 01	6.31E - 02	2.10E + 01	4.03E - 02	2.09E + 01	6.20E - 02	2.09E + 01	6.50E - 02
9	1.05E + 01	2.37E + 00	2.37E + 01	4.41E + 00	2.55E + 01	8.37E + 00	1.55E + 01	2.03E + 00	1.59E + 01	1.78E + 00
10	5.36E - 01	6.07E - 01	2.47E - 01	1.32E - 01	2.86E + 00	1.61E + 00	5.60E - 02	2.89E - 02	4.48E - 02	3.76E - 02
11	6.12E + 01	1.07E + 01	6.96E + 01	2.36E + 01	1.66E + 02	1.97E + 01	6.59E + 01	1.47E + 01	3.41E + 01	9.77E + 00
12	1.01E + 02	1.53E + 01	5.65E + 01	1.52E + 01	2.04E + 02	1.49E + 01	7.71E + 01	1.64E + 01	3.36E + 01	8.56E + 00
13	1.22E + 02	1.59E + 01	1.19E + 02	2.76E + 01	2.03E + 02	1.21E + 01	1.45E + 02	2.62E + 01	7.36E + 01	1.69E + 01
14	3.20E + 03	3.77E + 02	4.34E + 03	7.89E + 02	6.20E + 03	5.36E + 02	2.69E + 03	3.13E + 02	2.57E + 03	3.32E + 02
15	5.29E + 03	3.90E + 02	4.30E + 03	5.99E + 02	7.25E + 03	3.07E + 02	2.81E + 03	3.89E + 02	2.50E + 03	3.16E + 02
16	2.32E + 00	3.06E - 01	1.74E + 00	3.06E - 01	2.50E + 00	2.78E - 01	6.74E - 02	2.44E - 02	1.36E - 01	3.84E - 02
17	9.39E + 01	1.24E + 01	1.32E + 02	3.20E + 01	2.36E + 02	1.61E + 01	6.60E + 01	1.05E + 01	5.56E + 01	5.96E + 00
18	1.67E + 02	1.13E + 01	1.20E + 02	1.65E + 01	2.42E + 02	1.42E + 01	6.82E + 01	1.01E + 01	5.55E + 01	6.94E + 00
19	6.48E + 00	2.62E + 00	6.95E + 00	3.36E + 00	1.76E + 01	1.74E + 00	3.48E + 00	7.68E - 01	2.24E + 00	5.63E - 01
20	1.34E + 01	1.93E + 00	1.30E + 01	1.94E + 00	1.24E + 01	4.08E - 01	1.32E + 01	9.99E - 01	1.36E + 01	1.04E + 00
21	3.26E + 02	8.62E + 01	3.35E + 02	6.63E + 01	2.86E + 02	7.88E + 01	1.98E + 02	1.41E + 01	3.02E + 02	3.46E + 01
22	2.53E + 03	5.87E + 02	3.97E + 03	8.07E + 02	6.37E + 03	4.79E + 02	3.08E + 03	3.59E + 02	3.07E + 03	4.17E + 02
23	5.21E + 03	4.82E + 02	4.32E + 03	7.14E + 02	7.25E + 03	3.01E + 02	3.30E + 03	4.26E + 02	3.32E + 03	5.49E + 02
24	2.24E + 02	1.23E + 01	2.46E + 02	9.15E + 00	2.46E + 02	9.01E + 00	2.43E + 02	9.81E + 00	2.19E + 02	4.58E + 00
25	2.49E + 02	2.12E + 01	2.77E + 02	9.40E + 00	2.59E + 02	8.05E + 00	2.80E + 02	8.07E + 00	2.66E + 02	1.12E + 01
26	2.49E + 02	6.12E + 01	2.69E + 02	7.02E + 01	2.96E + 02	6.60E + 01	2.00E + 02	3.48E - 02	2.00E + 02	3.55E - 02
27	5.21E + 02	8.32E + 01	7.31E + 02	1.03E + 02	7.52E + 02	1.03E + 02	7.21E + 02	1.15E + 02	5.38E + 02	4.34E + 01
28	2.73E + 02	1.80E + 02	3.22E + 02	1.57E + 02	3.93E + 02	3.18E + 02	2.84E + 02	5.47E + 01	3.00E + 02	8.96E - 09
AR. 6 ~ 28	2.87	3.43	3.35	4.13	4.48	3.04	2.43	2.30	1.87	2.09
AR. 1 ~ 28	2.57	3.11	3.18	3.79	4.25	3.07	2.57	2.46	2.07	2.21

问题与量子理论的关系. 本节中将进一步阐明这些理论的相互关系, 见表 2 所示.

正如 1976 年图灵奖获得者 Rabin 所认为的: “应该放弃的只是以完全确定的方式获得结果, 这种结果可能出错, 然而出错的可能性微乎其微”. 不确定性不仅是概率算法构造的一个重要概念, 在本文中, 它同时也是概率算法量子特性的基础.

从优化问题的不确定性出发, 优化问题解的表达方式从一个确定的位置 x 变为为了一个概率分布

$\psi(x)$. 在可行域内某区域的解的分布的积分代表了在该区域解出现的几率, 这与量子理论中波函数的概率解释是一致的, 所以我们把这样的概率分布的表达方式称之为优化问题解的波函数. 这也是我们进一步尝试利用量子理论对优化问题进行讨论与研究的基础. 在优化算法的研究中, 使用数学的方法对随机过程进行研究是困难的, 而且通常需要对算法框架进行极大的简化. 所以本文建立优化问题的量子基础的核心思想是通过优化问题与势阱的等效

表 2 优化算法与量子理论的对应关系
Table 2 The relationship between optimization algorithm and quantum theory

优化算法	量子理论	说明
优化问题的解	波函数 $\psi(x)$	优化问题的概率描述
优化问题	薛定谔方程的约束条件	通过势阱等效将优化问题引入量子系统
优化系统的动态模型	含时薛定谔方程	含时薛定谔方程描述了随机搜索过程
多尺度过程	不确定性关系	通过算符的概念证明了多尺度过程的必要性
群体策略	波函数的叠加性	通过波函数的叠加性解释了算法种群的必要性
跳出局部最优	量子隧道效应	可以通过波函数来计算隧道效应概率
解的确定性	波函数熵	通过波函数定义优化算法的熵

对比, 把优化问题作为薛定谔方程的约束条件引入量子系统, 从而建立起薛定谔方程对随机搜索过程的动态描述. 在量子物理领域中, 存在许多计算量子体系势能的优化问题, 这类问题与计算机领域中的优化问题并无本质区别, 这也是我们可以借用量子系统对优化问题做类比研究的基础.

在此基础上, 我们从量子理论的角度出发尝试对优化问题进行了以下几方面的探索与讨论:

1) 薛定谔方程下的随机搜索过程: 在这一部分中, 我们通过类比量子力学与计算机领域的优化方法来研究随机搜索过程. 本文中讨论了扩散蒙特卡罗方法这一比较有代表性的随机优化方法. 该方法首先基于薛定谔方程与扩散方程的同构, 把薛定谔方程转换至实数域中, 再对薛定谔方程的势能约束条件 (目标函数) 引入了参考能量的概念, 通过参考能量调节随机行走的粒子的扩散行为, 最终影响波函数的演化结果, 从而获得低能量约束条件对应的基态波函数 (粒子分布状态). 单纯从优化算法的角度来看, 扩散蒙特卡罗之类的算法与优化领域的算法没有任何区别, 因此这类的方法中采用的数学物理手段, 都可以用于研究分析优化算法, 这也是建立优化问题与量子理论联系的优势所在.

2) 目标函数的近似: 由于计算机领域和量子物理领域对优化问题求解的不同需求, 我们首先需要把求解复杂目标函数转换成求解一个简单势阱的问题. 本文在选择势阱逼近模型时采用二阶 Taylor 近似逼近目标函数时的量子模型为量子谐振子模型, 其对应的基态波函数为高斯分布. 相似地, 在 QPSO 算法中, δ 势阱用于对目标函数进行逼近, 其对应的基态波函数为 Laplace 分布; 同时根据薛定谔方程下波函数的叠加特性, 利用简单势阱的基态波函数, 构造出优化系统的邻域函数. 这与概率优化系统中群体的概念是类似的, 在一定程度上印证了优化系统的量子特性. 同时, 这一概念的有效性已被我们的前期研究所证明. 并在这个框架下引入了更

精细的算法机制来提高算法的性能^[25-28].

3) 多尺度过程: 在许多优化算法中, 搜索尺度策略是影响算法性能的一个重要因素. 我们通过量子系统中优化问题的波函数的解释, 类比地定义了优化问题物理量的算符. 并通过位置算符和频率算符的相互关系, 证明了优化问题的测不准关系, 从而从量子物理的角度证明了优化问题多尺度策略必要性.

上述讨论说明优化问题与优化算法在量子理论的框架下可以得到部分的解释, 这说明优化问题与量子理论存在一定的相关性. 为分析优化算法提供数学物理理论的支持.

6 结束语

本文针对优化问题量子特性展开讨论, 其目的并不是为了提出一种新的优化算法, 而是希望为优化问题和优化算法的研究和分析提供一种新的视角.

本文通过对比研究表明, 量子理论与优化问题之间确实存在广泛的内在联系, 量子理论给出了描述优化问题的较为完备的数学框架, 为研究优化问题提供了一种有明确物理意义的数学工具. 虽然本文未对细节进行展开讨论, 但我们现有的研究表明, 在这一模型指导下的算法设计是有效的, 并且可以得到一些不错的结果.

从优化领域的发展趋势看提出新的优化算法不再是现在的研究焦点, 探索为优化问题建立统一的理论模型在现阶段变得尤为重要. 量子理论作为优化问题统一模型的理论基础是一个不错的选择, 本文的研究结果也支持这一结论. 但相关的研究还有待进一步的深入, 事实上的确还有一些重要问题并没有得到解决, 例如优化算法的隐含并行性问题 (Implicit parallelism)^[38-39] 的量子解释, 我们希望在日后的工作中完成.

References

- 1 Holland J H. *Adaptation in Natural and Artificial Systems*. Michigan: University of Michigan Press, 1975.
- 2 Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H, Teller E. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 1953, **21**(6): 1087–1092
- 3 Kirkpatrick S, Gelatt Jr C D, Vecchi M P. Optimization by simulated annealing. *Science*, 1983, **220**(4598): 671–680
- 4 Dorigo M. *Optimization, Learning and Natural Algorithms* [Ph.D. dissertation], Politecnico di Milano, Italy, 1992.
- 5 Dorigo M, Gambardella L M. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1997, **1**(1): 53–66
- 6 Eberhart R, Kennedy J. A new optimizer using particle swarm theory. In: Proceedings of the 6th International Symposium on Micro Machine and Human Science. Nagoya, Japan: IEEE, 1995. 39–43
- 7 Kennedy J. Probability and dynamics in the particle swarm. In: Proceedings of the Congress on Evolutionary Computation. Portland, USA: IEEE, 2004. 340–347
- 8 Omran M G H, Engelbrecht A P, Salman A. Bare bones differential evolution. *European Journal of Operational Research*, 2009, **196**(1): 128–139
- 9 Li J Z, Tan Y. The bare bones fireworks algorithm: A minimalist global optimizer. *Applied Soft Computing*, 2018, **62**: 454–462
- 10 Sörensen K. Metaheuristics — The metaphor exposed. *International Transactions in Operational Research*, 2015, **22**(1): 3–18
- 11 Yang X S. *Nature-inspired Optimization Algorithms*. Amsterdam: Elsevier, 2014.
- 12 Weyland D. A rigorous analysis of the harmony search algorithm: How the research community can be misled by a “novel” methodology. *International Journal of Applied Metaheuristic Computing*, 2010, **1**(2): 50–60
- 13 Griffiths D J, Schroeter D F. *Introduction to Quantum Mechanics*. Cambridge: Cambridge University Press, 2018.
- 14 Narayanan A, Moore M. Quantum-inspired genetic algorithms. In: Proceedings of the IEEE International Conference on Evolutionary Computation. Nagoya, Japan: IEEE, 1996. 61–66
- 15 Han K H, Kim J H. Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. *IEEE Transactions on Evolutionary Computation*, 2002, **6**(6): 580–593
- 16 Wang L, Niu Q, Fei M R. A novel quantum ant colony optimization algorithm. In: Proceedings of the International Conference on Life System Modeling and Simulation. Shanghai, China: Springer, 2007. 277–286
- 17 Sun J, Xu W B, Feng B. A global search strategy of quantum-behaved particle swarm optimization. In: Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems. Singapore, Singapore: IEEE, 2004. 111–116
- 18 Wang Peng, Huang Yan, Ren Chao, Guo You-Ming. Multi-scale quantum harmonic oscillator for high-dimensional function global optimization algorithm. *Acta Electronica Sinica*, 2013, **41**(12): 2468–2473
(王鹏, 黄焱, 任超, 郭又铭. 多尺度量子谐振子高维函数全局优化算法. 电子学报, 2013, **41**(12): 2468–2473)
- 19 Wang P, Ye X G, Li B, Cheng K. Multi-scale quantum harmonic oscillator algorithm for global numerical optimization. *Applied Soft Computing*, 2018, **69**: 655–670
- 20 Kennedy J. Swarm intelligence. *Handbook of Nature-Inspired and Innovative Computing*. Boston, USA: Springer, 2006. 187–219
- 21 Born M. Quantenmechanik der stoßvorgänge. *Zeitschrift für physik*, 1926, **38**(11–12): 803–827
- 22 Finnila A B, Gomez M A, Sebenik C, Stenson C, Doll J D. Quantum annealing: A new method for minimizing multidimensional functions. *Chemical Physics Letters*, 1994, **219**(5–6): 343–348
- 23 Kosztin I, Faber B, Schulten K. Introduction to the diffusion Monte Carlo method. *American Journal of Physics*, 1996, **64**(5): 633–644
- 24 Wang Peng, Huang Yan, Yuan Ya-Nan, Du Zheng, An Jun-Xiu. The convergence characteristics of multi-scale quantum harmonic oscillator algorithm. *Acta Electronica Sinica*, 2016, **44**(8): 1988–1993
(王鹏, 黄焱, 袁亚男, 都政, 安俊秀. 多尺度量子谐振子算法的收敛特性. 电子学报, 2016, **44**(8): 1988–1993)
- 25 Li B, Wang P, Jin J. Multiscale quantum harmonic oscillator algorithm with strict metastability constraints for multi-modal optimization. *IEEE Access*, 2019, **7**: 17377–17388
- 26 Ye X G, Wang P, Xin G, Jin J, Huang Y. Multi-scale quantum harmonic oscillator algorithm with truncated mean stabilization strategy for global numerical optimization problems. *IEEE Access*, 2019, **7**: 18926–18939
- 27 Li B, Wang P. Multiscale quantum harmonic oscillator algorithm with multi-harmonic oscillators for numerical optimization. *IEEE Access*, 2019, **7**: 51159–51170
- 28 Cheng K, Wang P. Analysis of multiscale quantum harmonic oscillator algorithm based on a new multimode objective function. *IEEE Access*, 2019, **7**: 46295–46305
- 29 Lu Zhi-Jun, An Jun-Xiu, Wang Peng. Partition-based MQHOA for multimodal optimization. *Acta Automatica Sinica*, 2016, **42**(2): 235–245
(陆志君, 安俊秀, 王鹏. 基于划分的多尺度量子谐振子算法多峰优化. 自动化学报, 2016, **42**(2): 235–245)
- 30 Sun J, Fang W, Wu X J, Palade V, Xu W B. Quantum-behaved particle swarm optimization: Analysis of individual particle behavior and parameter selection. *Evolutionary Computation*, 2012, **20**(3): 349–393
- 31 Brooke J, Bitko D, Rosenbaum T F, Aeppli G. Quantum annealing of a disordered magnet. *Science*, 1999, **284**(5415): 779–781
- 32 Stella L, Santoro G E, Tosatti E. Optimization by quantum annealing: Lessons from simple cases. *Physical Review B*, 2005, **72**(1): Article No. 014303
- 33 Xin G, Wang P. Exploring superposition state in multi-scale quantum harmonic oscillator algorithm. *Applied Soft Computing*, 2021, **107**: Article No. 107398
- 34 dos Santos-Coelho L, Ayala H V H, Freire R Z. Population’s variance-based adaptive differential evolution for real parameter optimization. In: Proceedings of the IEEE Congress on Evolutionary Computation. Cancun, Mexico: IEEE, 2013. 1672–1677
- 35 Zambrano B M, Clerc M, Rojas R. Standard particle swarm optimisation 2011 at CEC-2013: A baseline for future PSO improvements. In: Proceedings of the IEEE Congress on Evolutionary Computation. IEEE, 2013. 2337–2344
- 36 Li J Z, Tan Y. Loser-out tournament-based fireworks algorithm for multimodal function optimization. *IEEE Transactions on Evolutionary Computation*, 2018, **22**(5): 679–691
- 37 Liang J, Qu B Y, Suganthan P N, Hernández-Díaz A G. Problem Definitions and Evaluation Criteria for the CEC 2013 Special Session on Real-parameter Optimization, Technical Report 201212, Computational Intelligence Laboratory, Zhengzhou University, China, Nanyang Technological University, Singapore, 2013.
- 38 Bertonni A, Dorigo M. Implicit parallelism in genetic algorithms.

Artificial Intelligence, 1993, **61**(2): 307-314

- 39 Wang Peng, Chang Zheng. Physical model of implicit parallelism in algorithms. *Journal of University of Electronic Science and Technology of China*, 2009, **38**(4): 588-591
(王鹏, 常征. 算法隐含并行性的物理模型. 电子科技大学学报, 2009, **38**(4): 588-591)



王 鹏 西南民族大学计算机科学与技术学院教授. 2001 年获四川大学核技术及应用专业硕士学位. 2004 年获中国科学院成都计算机应用研究所计算机软件与理论专业博士学位. 主要研究方向为量子理论, 量子启发式算法, 计算智能与高性能计算. 本文通

信作者. E-mail: qhoalab@163.com

(WANG Peng Professor at the School of Computer Science and Technology, Southwest Minzu University. He received his master degree in nuclear technology and applications from Sichuan University in 2001 and Ph.D. degree in computer software and theory from

Chengdu Institute of Computer Application, Chinese Academy of Sciences in 2004. His research interest covers quantum mechanics, quantum inspired algorithm, computational intelligence, and high performance computing. Corresponding author of this paper.)



辛 罡 中国科学院成都计算机应用研究所博士研究生. 2011 年获德累斯顿工业大学微电子与固体电子专业硕士学位. 主要研究方向为量子启发式算法, 高性能计算.

E-mail: xin_gang@vip.126.com

(XIN Gang Ph.D. candidate at

Chengdu Institute of Computer Application, Chinese Academy of Sciences. He received his master degree in microelectronics and solid state electronics from Dresden University of Technology, Germany in 2011. His research interest covers quantum inspired algorithm and high performance computing.)