arXiv:2301.08787v2 [cs.ET] 7 Sep 2023

数字忆阻计算机的自平均

Daniel Primosch、Yuan-Hang Zhang 和 Massimiliano Di Ventra*

Department of Physics, University of California San Diego, La Jolla, CA 92093

数字忆阻计算机（DMM）是一类新型计算机，它们利用带记忆的非量子动力系统来求解组合优化问题。在此，我们表明，DMM 的求解时间（TTS）服从逆高斯分布，且随着问题规模增大，TTS *self-averaging*，无论它们求解何种问题。我们通过求解 3-SAT（可满足性）问题的实例，从理论上阐释了这一现象，并提供了数值证据。DMM 的自平均特性意味着，随着问题规模增大，它们对所求解实例的具体特征越来越不敏感。这与应用于相同问题的传统算法形成鲜明对比，也体现了这种基于物理的方法在计算上的另一项优势。

I. 引言

组合优化问题广泛出现在学术界和工业界的各种应用中 [1]。这类问题要求找到一组变量赋值——这些变量通常取逻辑值1或逻辑值0——使其满足一组约束。给定一个此类问题，例如每个约束恰好包含三个变量的可满足性问题（称为3-SAT），可以使用不同的算法来求解其实例。由于同一问题的不同实例各不相同，即使变量数和约束数固定，而且算法的初始条件也可以不同，因此我们预计求解时间（TTS）——即算法求解这些实例所需的时间——会呈现某种分布。

如果待求解的问题很难，例如3-SAT，那么随着问题规模 N增大，求解时间的均值 μ 和标准差 σ 预计都会呈指数增长。因此，我们预计在极限 $N \to \infty$下， $\sigma^2(N)/\mu^2(N) \sim O(1)$。在这种情况下，我们会说求解时间（TTS）不具有自平均性：其相对方差在热力学极限 [2] 下不会趋于零。（下文将为传统算法明确展示这一现象。）

近年来，有人提出了一种新的计算范式，称为忆阻计算 [3–5]，，它利用时间非局域性 （记忆）来解决计算问题。尤其是，其数字版本（数字忆阻计算机或 DMM） [6], 旨在解决组合优化类问题。与传统算法不同，DMM 将原始问题映射为非线性（非量子）动力系统，其点吸引子就是该问题的解（如果解存在）。由于 DMM 不会产生混沌 [7]（也不会产生准周期轨道），因此，即使在硬件实现之前，也可以在传统计算机上高效地模拟和检验其常微分方程（ODE）。这些模拟已经表明，DMM 在广泛的应用中显著优于最先进的算法适用于更广泛的应用领域（例如， [5] 中列举了这种计算范式已解决的一些问题）。

本文表明，DMM 的物理机制意味着，无论其求解何种问题，其求解时间都服从逆高斯分布 [8], 。最重要的是，随着问题规模增大，求解时间会自平均*</style}*，即对于，有 *()() as*；对于物理噪声，有 （“强自平均”），而对于数值噪声，有 （“弱自平均”）。除了对这一现象进行解析理解外，我们还通过在存在物理噪声的情况下数值求解 *3-SAT* 的可满足实例，验证了这一预测，并将结果与若干著名传统算法的求解时间进行比较–。*DMM* 的求解时间随问题规模增大而自平均，这意味着它们对所求解实例的具体特征越来越不敏感。这些结果既阐明了这些机器背后的物理机制，也展示了它们在计算方面的优势。

II. 用于 3-SAT 的 DMM

不失一般性，我们将重点讨论 3-SAT（实例取自 [9]），它由一组子句（或门）构成，每个子句恰好包含三个布尔变量，y_i ($i = 1, 2, 3$)，这些子句通过逻辑合取（与门）关联起来。对于此类问题，可以按如下方式构造一个 DMM（不一定唯一） [13]。首先将变量 y_i 转换为连续变量，$v_i \in [-1, 1]$ （在 DMM 的实际电路实现中，它们可以是电压 [6]），其中 $v_i > 0$ 对应于 $y_i = 1$，而 $v_i < 0$ 对应于 $y_i = 0$。第 m 个子句表示为（$l_{m,i} \vee l_{m,j} \vee l_{m,k}$），其中 $l_{m,i} = \bar{y}_i$或 y_i，具体取决于 y_i 是否取反，而 $\vee$ 是或门的符号。随后，每个布尔子句都会映射为一个连续约束函数，

$$C_m(v_i, v_j, v_k) = \frac{1}{2} \min\left(1 - q_{m,i} v_i, 1 - q_{m,j} v_j, 1 - q_{m,k} v_k\right), \tag{1}$$

其中，如果 $l_{m_i} = y_i$ 且 $q_{m,i} = -1$ 如果 $l_{m_i} = \bar{y}_i$，则 m项的值为真，当且仅当

*电子邮箱：diventra@physics.ucsd.edu

如果 $C_m < 1/2$。

时间非局域性通过引入额外的“记忆自由度”来实现，这些自由度保证动力学的不动点仅对应于问题的解，且相空间中不存在鞍点以外的其他临界点 [5]。按照 [13]，，我们为每个或门引入两个额外的记忆变量：“短期”记忆 $x_{s,\,m}$ 和“长期”记忆 $x_{l,\,m}$。对于含有 N 个变量和 M 个子句的 3-SAT实例，DMM 的动力学为（$i = 1,\ \ldots\ldots,\ N;\ \ m = 1,\ \ldots\ldots,\ M$）：

$$\dot{v}_i = \sum_m x_{l,m} x_{s,m} G_{m,i} + (1 + \zeta x_{l,m})\,(1 - x_{s,m})\,R_{m,i}, \quad (2)$$

$$\begin{aligned} \dot{x}_{s,m} &= \beta\,(x_{s,m} + \epsilon)\,(C_m - \gamma), \quad x_{s,m} \in [0,1], \\ \dot{x}_{l,m} &= \alpha\,(C_m - \delta), \quad x_{l,m} \in [1, 10^4 M], \end{aligned} \quad (3)$$

其中，“类梯度”项为 $G_{m,\,i} = \frac{1}{2} q_{m,\,i} \min(1 - q_{m\,j} v_j, \ldots, 1 - q_{m,\,k} v_k)$；而“刚性”项为 $R_{m,\,i} = \frac{1}{2}(1 - q_{m\,i} v_i)$，若 $C_m = \frac{1}{2}$，则为 $R_{m\,i} = 0$，否则。与 [13] 中一样，我们选取参数$\alpha = 5,\ \beta = 20,\ \gamma = 0.25,\ \delta = 0.05,\ \epsilon = 10^{-3},\ \zeta = 0.1$。

关于方程（ [5, 13]的推导过程，详见。这里，我们直接将其视为 DMM 的表示形式，并指出它们可以简洁地写成 $\dot{\mathbf{x}}(t) = F(\mathbf{x}(t))$，其中$\mathbf{x}$ 是连续变量的集合，v_i，以及记忆变量，$x_{s,\,m}$，$x_{l,\,m}$，而 F 是可从方程（2）和（3）右侧读出的流矢量场。

为与实验直接联系，正如 [14], 中所做的那样，我们向方程（ $l(t)$中的记忆变量加入高斯白噪声，。 $l(t)$ 满足 $\langle l(t)\rangle = 0$ 和$\langle l(t)l(t')\rangle = \Gamma\delta(t - t')$，其中 Γ 是噪声强度。当然，这些常微分方程数值积分所引入的误差也起着“噪声”的作用；关于无加性噪声时的结果，请参阅补充材料（SM） [15] 。

III. DMM的动力学

无论是通过解析方法（利用超对称拓扑场论）还是数值方法，都已证明 DMM 在有无加性噪声的情况下，都会沿着相空间中的特定轨迹（“瞬子”） [16, 17]寻找问题实例的解。具体而言，DMM 会从具有一定数量不稳定方向的临界点（鞍点，其 $F(\mathbf{x}(t)) = 0$ 满足方程（2）和（3））“跃迁”至更稳定的临界点，直至找到解。图1展示了一个含有 3,000 个变量和 21,000 个约束的 3-SAT实例的此类动力学示例，其中连续变量 v_i 的突然跃迁对应于连接临界点的瞬子。这些结果

通过对 DMM 的方程（2）和（3）进行积分得到

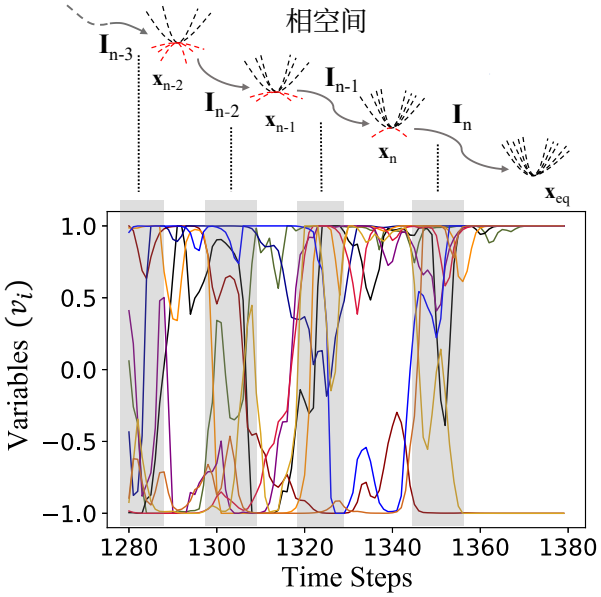

图 1。上图：DMM 状态动力学示意图。DMM 在相空间中从一个临界点（例如，$\mathbf{x}_{n-1}$）跃迁至更稳定的临界点（$\mathbf{x}_n$），每次瞬子跃迁（$\mathbf{I}_{n-1}$）后，不稳定方向的数量都会减少。下图：由方程（2）和（3）描述的 DMM 求解一个含有 3,000 个变量和 21,000 个约束的 3-SAT实例的数值示例。瞬子跃迁对应于阴影区域，在这些区域中，取值为负数（逻辑值0）的变量会突然变为正数（逻辑值1），反之亦然。两者之间的区域对应于一个临界点。

使用前向欧拉法 [18]，积分步长固定为（$\Delta t = 0.2$（全文均采用任意时间单位），且不额外添加噪声。

A. DMM 的求解时间分布

如预期，在存在噪声（无论是物理噪声和/或数值噪声）的情况下，我们预计求解时间会遵循某种分布。为了确定预期的分布类型——无论待求解的问题是什么——我们可以借助 DMM 动力学的物理类比，如图 1 所示。状态向量$\mathbf{x}(t)$ 在相空间中可以解释为“粒子的位置”，该粒子被 引导 朝向解，并受到某种噪声的影响。于是，该粒子具有“漂移速度”$\delta > 0$，这类似于 DMM 被驱动着朝向问题实例的解的程度。在这个物理系统中，由扩散常数 ν 建模的噪声，类似于 DMM 所经历的噪声。

接下来，我们可以询问这样的粒子从其起点（初始条件）到达任意距离 L （问题的解）需要多长时间。我们可以将其称为到达势垒时间（TTB），它是 DMM 的 TTS 的物理类比。这种到达势垒时间早已

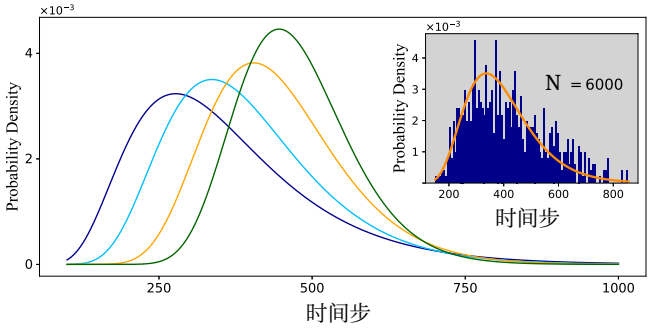

图 2. 对 1,000 个 3-SAT 实例（取自 [9]）中不同变量数的 DMM 的求解时间（以积分步数计）进行的各种拟合（逆高斯分布），N，子句变量比固定为 7。噪声强度为 $\Gamma = 0.12$（以时间倒数单位计）。插图：具有$N = 6{,}000$ 个变量和 42,000 个子句的 DMM 的求解时间直方图（以积分步数计）。拟合采用逆高斯分布，见式（5），其中 $\lambda = 3720 \pm 420$ 且$\mu = 390 \pm 10$ 。

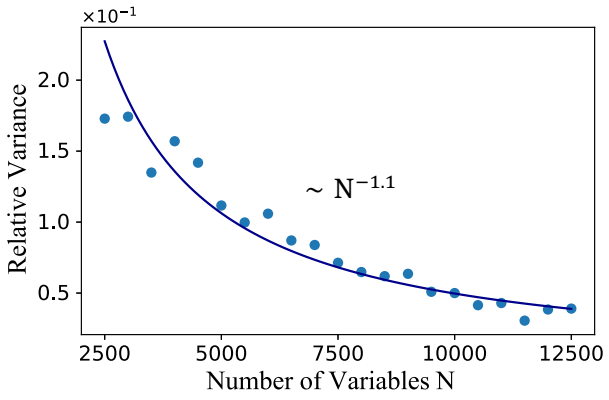

图 3. 图 σ^2/μ^2 2插图中的求解时间分布的相对方差 N，随变量数量的变化。相对方差可用幂律$N^{-1.10\pm0.07}$ σ^2/μ^2很好地拟合，这与自平均行为一致。

为文献所熟知，其概率密度函数服从如下形式的逆高斯分布 [8]

$$P_{\delta,L,\nu}(t) = \sqrt{\frac{L^2}{2\pi\nu t^3}} e^{-\delta^2(t-\frac{L}{\delta})^2/2\nu t}, \tag{4}$$

其中 $t > 0$ 表示粒子到达距离 L所需的时间。通过代入 $\mu = L/\delta$，$\lambda = L^2/\nu$，我们可以用两个独立参数重写式（4）。由此得到简化后的方程

$$P_{\lambda,\mu}(t) = \sqrt{\frac{\lambda}{2\pi t^3}} e^{-\lambda(t-\mu)^2/2\mu^2 t}. \tag{5}$$

该分布的均值为 μ，而其方差为$\sigma^2 = \mu^3/\lambda$ [8]。如果我们刚才所作的类比成立，那么 DMM 的求解时间应当也呈现类似的分布。

我们通过数值积分验证这一解析预测。我们再次使用前向欧拉法，并在方程 (2) 和 (3) 中的记忆变量上加入强度为 $\Gamma = 0.12$（以时间倒数单位计） [19]的高斯白噪声。求解时间（TTS）可以简单地定义为 DMM 找到解所需的积分步数。图 2 的插图报告了 DMM 求解 1,000 个 3-SAT 实例时的求解时间直方图 [9]，每个实例包含 $N = 6{,}000$ 个变量和 42,000 个子句。该直方图与参数为 5 的逆高斯分布拟合良好，参数为 $\lambda = 3720 \pm 420$和$\mu = 390 \pm 10$ [20]。图 2的主图展示了变量数不同、子句与变量比恒定时的拟合结果。这表明，随着变量数增加，逆高斯分布会变得更加尖锐。（关于采用不同积分方法以及不同子句与变量比时的更多统计数据，请参见 SM [15] 。）

B. 求解时间的自平均

为了更好地理解这一点，我们需要确定该分布如何随问题规模（变量数 N，假设子句变量比固定）变化。具体而言，我们想了解比值 σ^2/μ^2 如何随 N增加而变化。为此，我们再次运用刚刚建立的物理类比。

我们可以很容易地将 L 动力学初始点到 N的距离与变量数量联系起来，因为 L对应于相空间中一个 $N-$球体的平均半径，$S^N(L)$。在 N较大的极限下，这一数值就是 $L \sim \sqrt{N}$ [21]。

扩散常数 ν 与系统中的噪声量有关。对于物理噪声（例如我们在此考虑的加性噪声），它仅由某些外部参数决定，例如温度 [2]。换言之，对于物理噪声，ν 独立于 N。

最后，漂移速度 δ 对 N 的依赖关系可表示如下。若我们有 N 连续变量v_i，则注入系统的初始能量为 $E = N \times E_0$，其中 E_0 是与单个变量相关的能量（且在初始时刻 $t = 0$，各变量彼此独立；它们立即通过记忆变量耦合起来，耦合时间为 $t > 0$）。假设 N变量的质量均为单位质量（只要质量为常数，其具体数值并不重要），则漂移速度对 N 的依赖关系为$\delta \sim \sqrt{E} = \sqrt{N \times E_0}$。这意味着方差 $\sigma^2(N) = (L\nu)/\delta^3 \sim \nu/N$。

综合以上内容，我们得出结论：受物理噪声影响的 DMM 的求解时间会（强烈地）自平均：$\sigma^2(N)/\mu^2(N) \to 0$ 当 N^{-1} 时，对于 $N \to \infty$。这与根据中心极限定理 [22] 对自平均观测量得到的结果相同。这表明，DMM 对其所求解实例的微观细节（不同变量在各个子句中的分布方式）越来越不敏感。另请注意

通过这一物理类比得到的结果无论 DMM 求解什么问题都成立。这仅与它们的一般动力学行为有关。

在图 3 中，我们展示了与前述分析相印证的数值结果：针对不同的变量数量 $\Gamma = 0.12$，我们求解方程（2）和（3）（后者加入了强度为 $\Gamma = 0.12$的噪声），N并保持子句与变量比固定。相对方差可由曲线 σ^2（N）$/\mu^2$（N）$\sim N^{-1.10\pm0.07}$很好地拟合；结果也表明，在这种情况下，数值噪声对随机动力学的影响并不大。

如果只有数值噪声，或者数值噪声与加性噪声强烈耦合，那么扩散系数 ν 与 N 无关这一假设便不再成立，因为误差会不断累积。在这种情况下，我们预计 ν（N）会随 N增大，其函数形式取决于所采用的数值方法。补充材料（SM）的图 S7 和 S8[15], 显示，求解时间仍具有自平均性，但对于前向欧拉法，其自平均性较弱 [σ^2（N）$/\mu^2$（N）$\sim N^{0.39\pm0.02}$；而对于四阶龙格–库塔法，$N^{0.29\pm0.02}$则为]。

C. 传统算法的非自平均性 hms

最后，我们通过展示传统算法在同一问题上的求解时间并不具备 DMM 的自平均性质，来作总结。我们在图4中对此进行了明确展示：图中采用了一种广泛使用且具有代表性的局部搜索算法（WalkSAT）[10]，并将其应用于此前使用的相同实例 [9]。（在补充材料中 [15]，我们还展示了使用两种不同的冲突驱动子句学习算法 MiniSAT [11] 和 Kissat [12], 所得的其他数值结果；正如预期，它们也不具备自平均性。）

该算法按照预先设定的规则，在每次迭代中翻转一个或几个变量。与 DMM 不同，这种局部搜索算法的求解时间概率分布很好地符合指数拟合，形式为 $P_\lambda(t) = \lambda e^{-\lambda t}$。（关于这种分布为何必然是指数分布的解析论证，请参见 SM [15] 。）这里，λ 是平均值（$\mu = 1/\lambda$）的倒数，而该分布的方差为 $\sigma^2 = 1/\lambda^2$。由此可见，相对方差为 $\sigma^2/\mu^2 = 1$，且无论问题规模大小如何都保持不变。正如预期，这种传统算法并不具有自平均性。

IV. 结论

我们已经证明，DMM 的求解时间服从逆高斯分布，其相对方差会随着问题规模增大而趋近于零。这

无论存在物理噪声（体现出强自平均），还是存在数值噪声（体现出弱自平均），这一点都成立。这意味着，随着问题规模增大，DMM 对其求解实例的微观细节越来越不敏感。这一性质并非

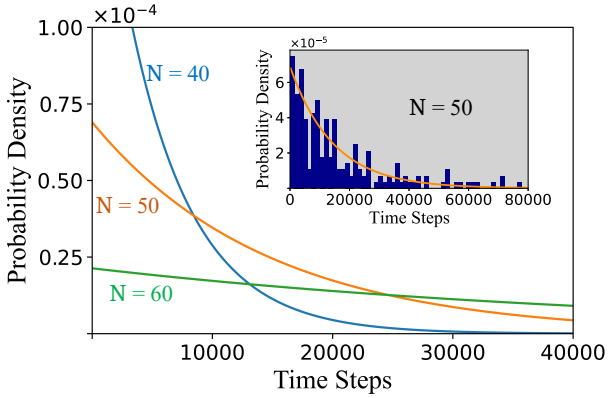

图4. 在固定子句变量比为 7 的情况下，局部搜索算法 WalkSAT 对 $N = 40$，50，60 个变量（200 多个实例）的求解时间（TTS）分布。对于所有考察的情况，TTS 均很好地符合指数概率分布。插图：变量数为 $N = 50$ 时的一个示例，展示了用于拟合主图数据的 TTS 直方图。

传统算法并不具备这一性质，这凸显了 DMM 在求解组合优化问题方面的显著优势。

致谢

本研究由美国国家科学基金会（NSF）第 2229880 号拨款资助。可在 [23] 获取用于复现本研究结果的演示代码。

[1]M. R. Garey 和 D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*（W. H. Freeman & Co.，纽约，1990年）。[2] N. van Kampen, *Stochastic Processes in Physics and Chemistry* 1992（爱思唯尔，阿姆斯特丹，）。[3] M. Di Ventra 和 Y. V. Pershin，《并行方法》，《自然·物理》**9**，200（2013年）。[4] F. L. Traversa 和 M. Di Ventra，通用 MemComputing 机器，《IEEE 神经网络与学习系统汇刊》与学习系统**26**，2702（2015年）。 [5] M. Di Ventra, *MemComputing: Fundamentals and Applications* 2022（牛津大学出版社，牛津，）。[6] F. L. Traversa 和 M. Di Ventra，利用数字忆阻计算机在多项式时间内求解素因数分解和 NP 完全问题，《混沌：非线性科学跨学科期刊》**27**，023107（2017）。

[7]M. Di Ventra 和 F. L. Traversa，具有解的数字忆阻计算机中不存在混沌，《Physics Letters A》381，3255（2017）。[8] G. A. Whitmore 和 V. Seshadri，逆高斯分布的启发式推导，《The American Statistician》41，280（1987年）。[9] W. Barthel、A. K. Hartmann、M. Leone、F. Ricci-Tersenghi、M. Weigt 和 R. Zecchina，在随机可满足性问题中隐藏解：一种统计力学方法，《Physical Review Letters》88，188701（2002年）。[10] S. Arora 和 B. Boaz，*Computational Complexity: AModern Approach*（剑桥大学出版社，剑桥，英国，2009年）。[11] N. Sorensson 和 N. Een，带冲突子句最小化功能的 MiniSAT v1.13 SAT 求解器，SAT 2005，1（2005年）。[12] A. B. K. F. M. Fleury 和 M. Heisinger，CaDiCaL、Kissat、Paracooba、Plingeling 和 Treengeling 参加 SAT 竞赛 2020，SAT 竞赛 2020 50，2020年（2020年）。[13] S. R. B. Bearden、Y. R. Pei 和 M. Di Ventra，利用数字忆阻计算高效求解布尔可满足性问题，《Scientific Reports》10，19741（2020年）。[14] S. R. B. Bearden、H. Manukian、F. L. Traversa 和 M. Di Ventra，自组织逻辑门中的瞬子，《Physical Review Applied》9，034029（2018年）。[15] 关于局部搜索算法指数分布的解析理解，[24]，采用传统冲突驱动子句学习算法得到的分布，[25, 26]，以及不同子句与变量比、不同噪声模型和不同初始条件下的更多数值结果，请参阅补充材料。[16] M. Di Ventra、F. L. Traversa 和 I. V. Ovchinnikov，拓扑场论与基于瞬子的计算，《Annalen der Physik》（柏林）529，1700123（2017）。[17] M. Di Ventra 和 I. V. Ovchinnikov，数字忆阻计算：从逻辑到动力学再到拓扑，《Annals of Physics》409，167935（2019）。[18] T. Sauer，*Numerical Analysis*（培生，纽约，2018年）。[19] 根据 [14]，，我们可以将噪声强度 Γ（以 s^{-1}为单位）与有效外部温度（以 K 为单位）联系起来。本研究所考虑的数值 $\Gamma = 0.12\ s^{-1}$，对应的温度将与室温 ~ 300 K 相当。[20]请注意，对于给定实例（即具有 *fixed* 个变量和子句），在 *ensemble* 组初始条件上，TTS 分布（5）同样会出现。参见补充材料中的图 S9 [15]。[21] H. Coxeter，*Regular Polytopes*（多佛出版社，1973年）。[22] A. Aharony 和 A. B. Harris，临界点附近随机系统中的自平均缺失与普适涨落，《Physical Review Letters》77，3700–3703（1996年）。[23] https://github.com/yuanhangzhang98/MemComputing3

SAT_演示。[24] H. Hoos，SAT 的 SLS 算法行为混合模型，AAAI-02 会议论文集（2002年）。[25] J. P. Marques Silva 和 K. A. Sakallah，GRASP——一种新的可满足性搜索算法，载于 *The Best of ICCAD*（施普林格，2003年），第 73–89 页。[26] D. H. Wolpert 和 W. G. Macready，优化的免费午餐定理，《IEEE进化计算汇刊》1，67（1997年）。

补充材料：数字存内计算机的自平均

局部搜索算法的指数分布

局部搜索算法（例如正文中使用的 WalkSAT）的求解时间（TTS）分布已知服从指数分布 [24]。这里，我们给出一个启发式推导，说明其原因。

由于问题描述中某个特定变量的翻转与之前的翻转以及其他变量相互独立，因此，每次决定翻转某个特定变量都是一个由离散泊松分布描述的事件：

$$P_{Poisson}(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \tag{S1}$$

其中，k 表示稀有事件的数量，即单个变量翻转后得到给定问题解的事件数量；λ 表示稀有事件发生率。变量翻转后未得到解的概率为 $P(k=0) = e^{-\lambda}$。变量翻转彼此独立这一事实使我们能够计算在某个大于 t 的时间 T 发生变量翻转并得到解的概率：

$$P(T > t) = \prod_{i=1}^{t} P(k=0) = e^{-\lambda t}. \tag{S2}$$

然后，我们通过取补得到累积分布函数（f_{CDF}），

$$P(T \le t) = 1 - e^{-\lambda t} = f_{CDF}(t). \tag{S3}$$

概率密度函数（f_{PDF}）就是 f_{CDF} 的导数：

$$f_{PDF}(t) = \lambda e^{-\lambda t}, \tag{S4}$$

这正是我们预期的求解时间分布。

使用冲突驱动子句学习算法的结果

许多现代 SAT 求解器都基于冲突驱动子句学习（CDCL）算法 [25]。与 WalkSAT 不同，CDCL 求解器通过为选定的变量赋予真值，利用布尔约束传播真值，并在发现冲突时回溯，来尝试找到满足条件的配置。

CDCL 求解器通常会使用许多启发式方法，而求解器的性能取决于子句的分布，因此很难从分析上理解其求解时间分布。相反，在本节中，我们仅报告使用两种流行的 CDCL 求解器得到的数值结果：MiniSAT [11]，在 SAT 2005 竞赛的所有工业类别中均获胜，也是被引用最多的 SAT 求解器之一；以及 Kissat [12]，赢得了 SAT 竞赛 2020 和 2021 的主要赛道。与正文一样，我们使用 Barthel 实例 [9]（子句与变量之比为 7）作为测试问题。

图 S1、S2 展示了这两种算法的求解时间直方图。数值拟合表明，它们的求解时间服从重尾分布，且尾部不受指数界限约束。从经验上看，我们可以用威布尔分布拟合其求解时间，

$$f(x) = \begin{cases} \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, & x \ge 0, \\ 0, & x < 0. \end{cases} \tag{S5}$$

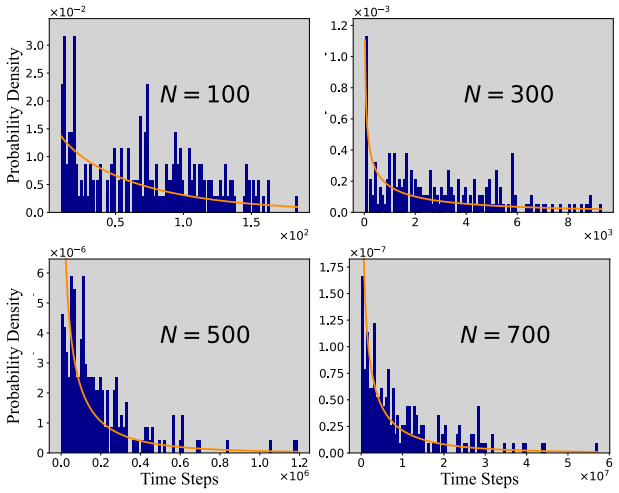

图 S1. MiniSAT算法的求解时间分布。对于每个直方图，我们尝试求解 200 个 Barthel 实例，子句变量比为 7，变量数各不相同 N。所有实例均已求解。使用威布尔分布（公式（S5））可以近似拟合这些直方图，其中 $k < 1$。请注意，随着问题规模增大，两个坐标轴的刻度都会大幅变化 N。

在图 S3中，我们绘制了两种算法的相对方差 $\sigma^2(N)/\mu^2(N)$。由于并非所有实例都能求解，相对方差是根据拟合分布计算的。对于两种算法，相对方差都先增大，然后在变量数足够大时趋于饱和。换言之，这些算法也不具有自平均性。

注意，Kissat算法的相对方差会增长到极端值。尽管Kissat能在短时间内求解大多数测试实例，但某些实例需要极长时间才能求解。我们认为，这种现象可能部分与SAT竞赛的目标有关：竞赛要求求解器在5000秒的时限内尽可能多地解决问题。宽泛地说，“免费午餐

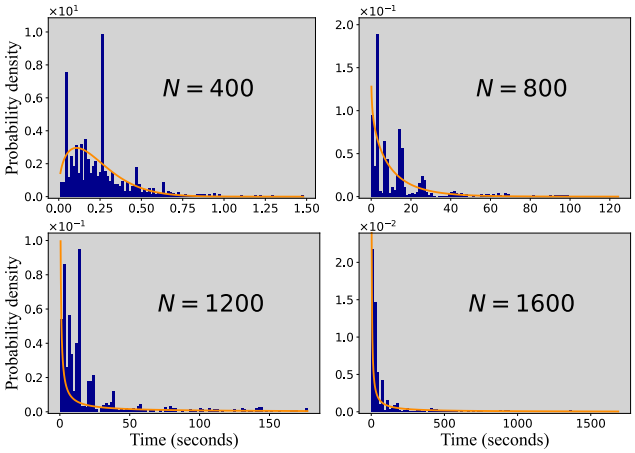

图 S2. Kissat算法的求解时间分布。对于每个直方图，我们在子句与变量比为 7 的情况下，对 999 个 Barthel 实例进行求解，变量数分别为 N。对于 $N = 400$ 和 800，所有实例均已求解。$N = 1200$ 个实例中有 995 个已求解，$N = 1600$ 个实例中有 997 个已求解。直方图可以用威布尔分布近似拟合。为便于观察，对于 $N = 800$，1200 和 1600，仅绘制已求解实例中最先求解的 90%，因为求解最难实例所需时间超过 10^5 秒。

定理” [26] 指出，如果没有利用任何结构，那么对某一类问题中的所有问题求解后，找到解的计算成本均值对于任何求解方法都相同。如果一个“优秀”的求解器能在极短时间内解决许多实例，这就意味着有一部分实例需要极长时间才能求解。因此，如果竞赛求解器需要在更短时间内解决更多实例，一种途径就是放大不同实例之间的这种差异，从而导致极大的方差。

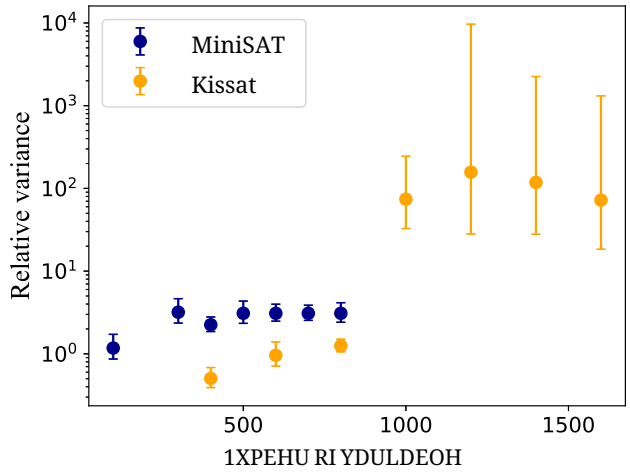

图 S3. MiniSAT 和 Kissat 的相对方差，$\sigma^2(N)/\mu^2(N)$。对于这两种算法，随着 N 增大，相对方差先迅速增长，随后达到平台期。这证实这些算法也不具有自平均性。

不同子句变量比下的自平均

正文中我们使用了子句变量比为 7 的 Barthel 实例 [9] 。这一选择是任意的，而不同类型的 3-SAT问题应当会得到类似的结果。本节给出子句变量比为 6 时的补充数值结果，以对此加以确认。

在图S4中，我们展示了不同变量数量 N下的 TTS 直方图。每个直方图均基于使用 DMM 求解的 1000 个问题实例，积分时采用固定时间步长 $\Delta t = 0.05$ 和噪声强度 $\Gamma = 0.12$。所有求解尝试均成功，其 TTS 均可用逆高斯分布很好地拟合。

不同 N 下 TTS分布的比较见图 S5，相对方差$\sigma^2(N)/\mu^2(N)$ 关于 N的函数关系绘于图 S6。幂律拟合表明 $\sigma^2(N)/\mu^2(N) \sim N^{-0.41\pm0.05}$，这表明存在弱自平均行为，说明在这种情况下，数值噪声与加性噪声强烈耦合。

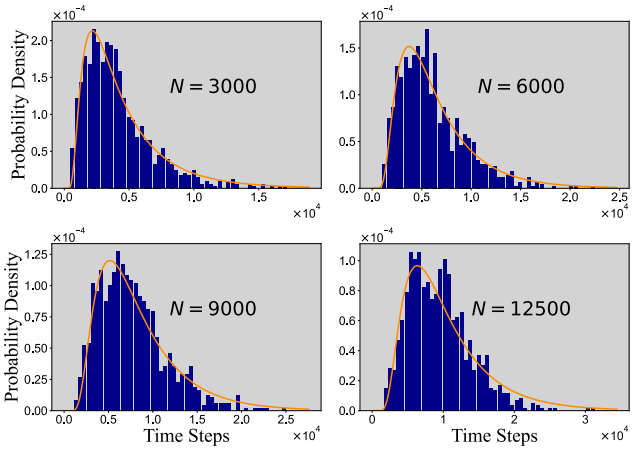

图 S4. 使用 DMM 求解子句变量比为 6 的 Barthel 实例时，求解时间的分布。每个直方图包含对 1000 个不同实例的求解尝试，且所有实例均成功求解。数值积分采用前向欧拉法，固定时间步长为 $\Delta t = 0.05$。

仅含数值噪声时求解时间的自平均

正文中方程（2）和（3）的数值积分会引入类似某种“噪声”的误差。事实上，这类噪声是非局域的，因为它会在积分过程中不断累积；而物理噪声通常在时间和空间上都是局域的。此外，由于数值噪声以非线性方式进入运动方程，其表现不同于我们在正文中所考虑的加性物理噪声。

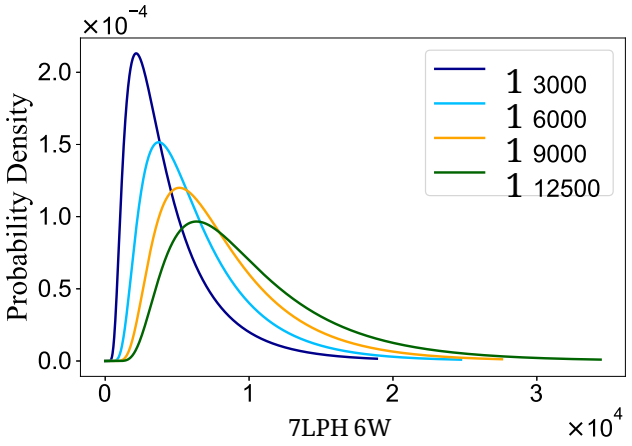

图 S5. 对变量数不同的逆高斯分布拟合结果进行比较， N，子句变量比为 6。拟合基于图 S4中所示的直方图。

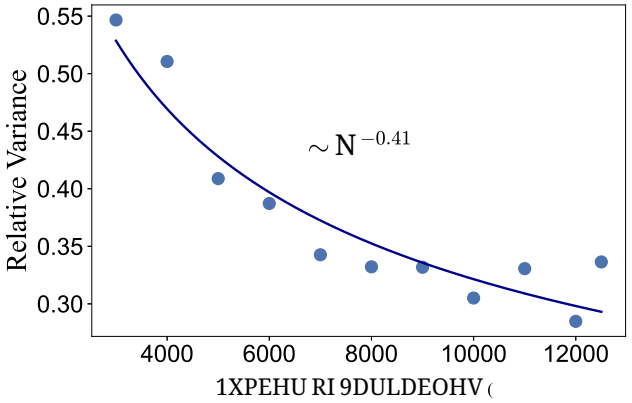

图 S6. 子句变量比为 6 时，相对方差 σ^2/μ^2 随变量数 N 的变化。相对方差可用幂律 $\sigma^2(N)/\mu^2(N) \sim N^{-0.41\pm0.05}$ 良好拟合，表明其具有弱自平均行为。

尽管如此，即使仅存在数值噪声，我们仍观察到求解时间的逆高斯概率分布，以及其自平均行为。不过，与存在物理加性噪声时相比，自平均效应较弱，且取决于所使用的积分方案。在图 S7 中，我们展示了采用前向欧拉积分方案时，一系列 3-SAT 实例（变量数从 2,500 到 20,000，子句变量比固定为 7）的相对方差。由此得到的相对方差为 $\sigma^2(N)/\mu^2(N) \sim N^{-0.39\pm0.02}$，表明存在弱自平均。

在图S8中，我们展示了采用四阶龙格-库塔积分方案进行的相同计算。求解时间仍然呈弱自平均，但相对方差不同，为 $\sigma^2(N)/\mu^2(N) \sim N^{-0.29\pm0.02}$。这证实了正文中的结论：仅存在数值噪声时，求解时间的相对方差

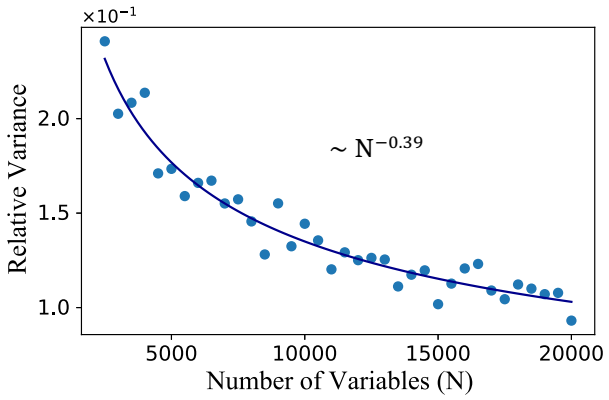

图 S7. 主文中的方程（2）和（3）所表示的 DMM 使用前向欧拉法积分时，其求解时间的自平均。对于每种规模，我们求解了 1,000 个 3-SAT 实例（来自 [9]），所用积分步长为 $\Delta t = 0.2$（任意单位）。相对方差 $\sigma^2(N)/\mu^2(N) \sim N^{-0.39\pm0.02}$，表明其呈现弱自平均。

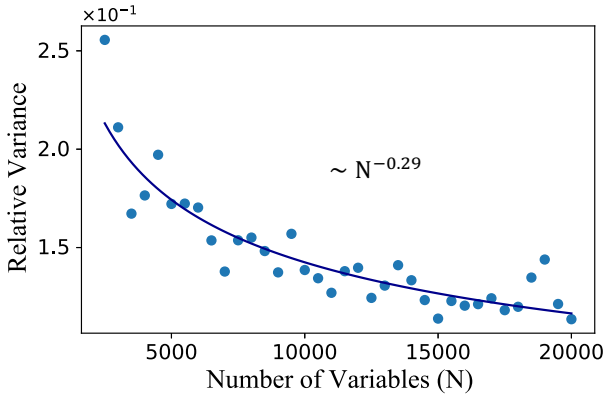

图 S8. 使用四阶龙格-库塔法对正文中的方程 (2) 和 (3) 所表示的 DMM 进行积分时，求解时间的自平均情况。对于每个规模，我们求解了 1,000 个 3-SAT 实例（来自 [9]），使用的积分步长为（任意单位）。相对方差 $\sigma^2(N)/\mu^2(N) \sim N^{-0.29\pm0.02}$，表明其弱自平均特性与使用前向欧拉积分方案得到的结果不同。

在 $N \to \infty$ 极限下趋于零，速率为 $N^{-\theta}$，且取决于积分方案。

随机初始条件下同一实例的求解时间分布

正文中的模拟针对每种问题规模采用不同实例，并为每个实例设置随机初始条件。在图S9中，我们展示了多次求解同一个实例并采用随机初始条件时，所得结果与每次尝试求解不同实例时观察到的逆高斯分布相同

。这表明，对 DMM 而言，不同 3-SAT 问题实例的“微观”特征仅仅表现为同一相空间中的不同初始条件。

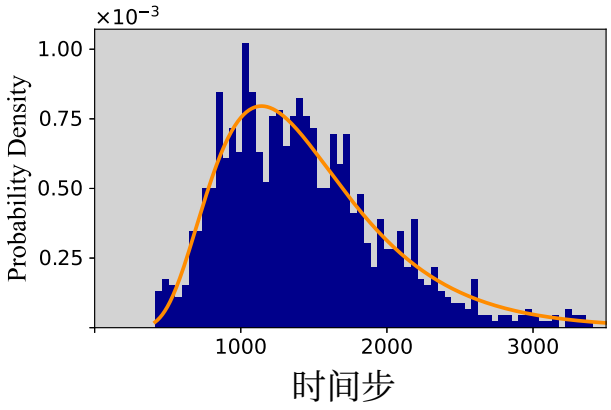

图 S9. DMM 求解取自 [9] 的一个$N = 6{,}000$ 个变量、42,000 个子句的 3-SAT实例时，求解时间（以积分步数计）的直方图；每次尝试均采用不同的随机初始条件。该直方图根据在固定超时时间内完成的 1,000 次 3-SAT 尝试中的 981 次生成。拟合采用逆高斯分布，即正文中的公式 (5)，其中 $\lambda = 6160 \pm 130$ 且 $\mu = 644 \pm 3$。