遍历周期：一种用于 2.5D/3D集成电路周期性热分析的高效随机求解器

Shengnan Cai[1], *, Qibo Xue[1], *, Tianshu Hou[1],Tianxiang Zhu[2],Zixiao Wang[3], †, Yibo Lin[2],Runsheng Wang[2],Bei Yu[1], †[1]香港中文大学，香港， [2]北京大学，中国北京[3]香港科技大学（广州），中国广州

arXiv:2611.32082v1 [cs.AR] 24 Nov 2026

摘要—对于在周期性工作负载下运行的 2.5D/3D集成电路，长期周期热分析至关重要。确定性全局求解器在重复进行热查询时成本高昂，而现有随机游走方法并非针对长期周期问题或多尺度 3D 集成电路结构设计，因此效率和准确率有限。我们提出了 *CycleWalker*，一种周期残差随机游走求解器。首先，我们将周期稳态（PSS）表述为无限历史反向游走，该方法无需初始温度或预热，但会将完整温度场的计算成本分摊到每条路径上。其次，为了降低这一成本，我们令游走仅估计周期先验的残差，并通过俄罗斯轮盘赌缩短长历史，同时保持无偏性。第三，由于估计量对先验的约束仅为其在时间上呈周期性且分段线性，我们通过粗网格傅里叶求解并随后进行细网格平滑，以极低的成本构建先验。最后，我们构建了能够感知硅通孔（TSV）的热图，在保留 TSV 阵列各向异性传导特性的同时，限制游走的跳跃率。在周期性 2.5D/3D 集成电路基准上的实验表明，在准确率相同的情况下，*CycleWalker* 相较于瞬态随机游走求解器的加速比最高可达 75.85× 。

I. 引言

在集成电路设计与优化过程中，热预测需要反复进行 [1]–[3]。对于 2.5D/3D集成电路，热量经由复杂结构散发，包括细长的硅通孔（TSV），这是设计中的一项关键考量 [4], [5]。这类集成电路（如 AI 加速器）通常在周期性工作负载下运行 [6],例如 周期性推理批次 [7] 和任务调度 [8]。持续运行期间，其温度会趋于周期稳态（PSS） [9], ，必须评估其在一个周期内的变化 [6]。

确定性方法（如有限体积法）广泛用于热分析 [10]。时域 [11],频域 [9], 和射击方法 [12] 均可求得周期稳态，但在高负载下，全场计算的成本可能很高

时空分辨率 [13]–[18]。然而，在许多设计评估中，只需要获取少数位置和相位处的温度，例如图 1 [19], [20] 中高功率阶段的热点温度。此时，大量全场计算都耗费在未被查询的位置和相位上，因此需要一种仅在需要的位置和相位评估周期稳态的求解器。

随机游走方法 [19], [20]无需求解完整场，即可估算选定位置的温度，从而降低稀疏热查询的成本。早期的浮点和混合随机游走求解器 [19] 用于求解稳态热传导问题，之后基于费曼–卡茨公式的求解器 [21]引入了罗宾边界处理，并通过预计算 [22] 和 GPU 加速 [23]提升了效率。先验辅助随机游走方法则通过使用截断点处的近似温度，进一步缩短游走路径。

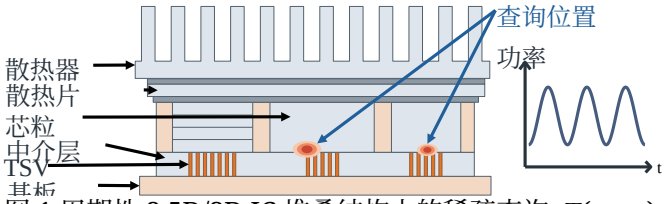


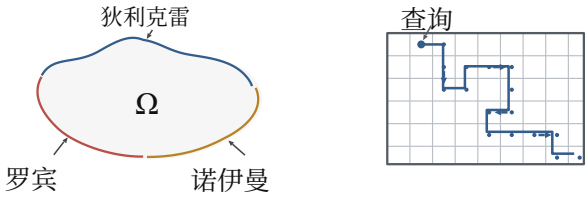
图 1 周期性 2.5D/3D IC 堆叠结构上的稀疏查询 $T(x,\ t)$。

这些稳态方法 [24],[25], 无法捕捉依赖阶段的周期稳态分析所需的时间演化。初值瞬态随机游走方法，包括时空场求解器和热网络上的时间反向游走 [20],，会显式纳入时间，并将指定的初始温度作为时间终点 [20]。在周期稳态分析中，并未给定这样的初始温度，因此反向游走必须考虑来自任意久远周期的贡献，或依赖人工初始化和预热。这促使我们开发一种直接针对周期稳态的随机游走求解器。

为应对这些挑战，我们提出*CycleWalker*，一种直接针对周期稳态热分析的随机游走求解器。我们首先将周期稳态 formulation 表述为无限历史反向游走，从而无需初始温度或预热，但代价是每条路径都要承载完整的温度场。因此，我们将温度分解为周期先验和误差场，并让游走仅估计误差，通过俄罗斯轮盘赌在不引入偏差的情况下终止过长的历史。该估计量要求先验在时间上具有周期性且为分段线性，但不要求其精确，因此我们先通过粗网格傅里叶求解，再进行细网格平滑，以极低成本构造先验。最后，由于游走运行于热图上，我们构造了一种热图，使其跳跃率在2.5D/3D集成电路的多尺度结构中保持有界。我们的主要贡献

总结如下：

- **周期随机 formulation**。 据我们所知，这是首个直接针对长期周期热分析的随机游走 formulation。
- **残差随机游走**。我们将温度分解为先验和误差场，并应用随机游走来估计误差。我们结合无偏的俄罗斯轮盘赌与自适应采样，以降低路径方差、避免截断偏差，并设定路径数。
- **周期性感知先验**。 为满足残差估计器所要求的周期性和分段线性特征，我们通过粗网格傅里叶求解，再进行轻量级细网格平滑，构建低成本先验，使残差能够通过同一周期游走进行采样，并采用时间积分闭式解。



（a）边界划分　　（b）网格上的游走

图 2 区域 Ω 及其不同边界，以及从查询单元出发、沿面相邻的反向游走。

•TSV 感知图构建。 我们结合特征对齐离散化与子网格 TSV 均质化。这样既能限制游走的 TSV诱导的跳跃率，又能保留 2.5D/3D集成电路中 TSV 阵列的各向异性传导特性。

II. 预备知识

我们首先回顾离散化的瞬态热传导模型，并说明其解如何表示为随机游走的期望，这为我们的周期性公式奠定了基础。

A. 瞬态热传导

令 $T(\boldsymbol{x},t)$ 表示计算区域 Ω 中的温度。它满足瞬态热传导方程 [26]

$$c_v\frac{\partial T}{\partial t}-\nabla\cdot(k\nabla T)=q(\boldsymbol{x},t), \tag{1}$$

其中 k 为热导率， c_v 为体积热容， q 为体积功率密度。边界 $\partial\Omega$ 被划分为三个互不相交的部分， $\partial\Omega = \Gamma_D\cup\Gamma_N\cup\Gamma_R$，其上施加了不同的边界条件 Fi 。2a 21：

$$\begin{aligned} &(\ \ g\ \ (\ ))\ [\ \ ] \\ &T=T_D \text{ on } \Gamma_D, \qquad k\frac{\partial T}{\partial n}=g_N \text{ on } \Gamma_N, \\ &-k\frac{\partial T}{\partial n}=h(T-T_a) \text{ on } \Gamma_R, \end{aligned} \tag{2}$$

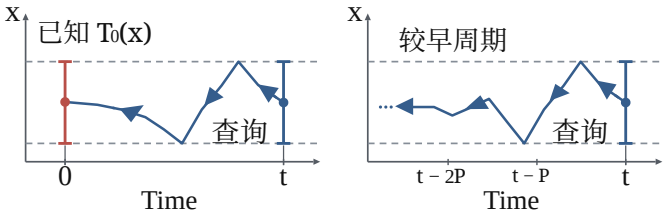
其中 n 是外向单位法向量。这三部分分别对应于指定的边界温度 T_D （狄利克雷边界条件）、指定的内向热通量 g_N （诺伊曼边界条件，其中 $g_N=0$表示绝热），以及以传热系数 T_a 向环境温度进行对流冷却 h （罗宾边界条件） [20]。对于 2.5D/3D集成电路，热量通过堆叠结构的顶部和底部散出，因此我们将散热器、热界面材料和印刷电路板的热阻归并为等效传热系数 [27]，并在这些位置施加 Robin边界 [28]；侧表面没有冷却结构，因此为绝热边界。我们用B_i 表示通过边界面从单元格 i 流向外部温度的热导。这些罗宾边界面是问题中唯一的吸收，因此，只有在 $B_i>0$ 这些位置，路径权重才会衰减，从而使第 III-A节中的无限历史得以收敛。

有限体积离散 [10] 给出了每个单元格的热平衡 i：储存能量的变化等于从相邻单元格流入的热量 j，减去通过边界散失的热量，再加上单元格内部产生的热量，

$$C_i\frac{dT_i}{dt}=\sum_{j\in\mathcal{N}_i}G_{ij}(T_j-T_i)-B_i\big(T_i-T_{a,i}\big)+q_i, \tag{3}$$

其中， $\mathcal{N}_i$ 是与单元格 i共享一个面的单元格集合， T_i 是单元格 i 的温度， $C_i=\int_{V_i}c_v\,dV$是体积为 V_i的单元格 i 的热容量， G_{ij} 是单元格 i 与 j之间的热导。对于位于 Γ_R 或 Γ上的单元格

$_D$,$T_{a,\ i}$ 是



(a) 有限历史　　(b) 周期稳态

Fi图 3 一维中的反向游走：有限历史与无周期稳态。

该面的外部温度。项 q_i 表示单元中产生的功率加上通过其各个面的诺伊曼通量。将未知量 T_i 移至左侧，将已知项移至右侧，得到 [29]

$$\boldsymbol{C}\dot{\boldsymbol{T}}(t)+\boldsymbol{A}\boldsymbol{T}(t)=\boldsymbol{q}(t), \tag{4}$$

其中 $\boldsymbol{C}$ 是对角电容矩阵， $\boldsymbol{A}$ 汇集胞间传导和边界耗散， $\boldsymbol{q}$ 汇集所有已知热输入，而 $\boldsymbol{T}(t)$ 堆叠单元温度$T_i(t)$。方程 (4) 是本文其余部分所依据的系统。我们的目标是求解 $\boldsymbol{T}(t)$。

B. 瞬态随机游走表示

在热图上进行游走。随机游走方法 [19],[20]将温度表示为随机路径上期望值。我们采用瞬态随机游走求解器的网格游走（WOG）公式 [30], ，其中游走者在热图的相邻单元格之间移动。根据方程（3），传导项定义了一个连续时间马尔可夫链，其中位于单元格 i的游走者以速率跳转至邻居 j，因此离开该单元格的总速率为 G_{ij}/C_i。 ν_i 。

$$\nu_i=\frac{\sum_j G_{ij}}{C_i}. \tag{5}$$

在单元格 i 中的停留时间服从指数分布，均值为 $1/\nu_i$，因此跳跃率较大的单元格每一步仅使游走沿反向时间前进很短的时间。令 X_s 表示在反向时间 s时所处的单元格。边界项以速率 B_i/C_i表示向环境的吸收。我们不对该事件进行逐步采样，而是为每条路径赋予生存权重

$$W_s=\exp\Big(-\int_0^s\frac{B(X_u)}{C(X_u)}\,du\Big), \tag{6}$$

其期望值相同，但方差更低 [20]。由于 $B_i=0$在内部单元中， W_s 仅当路径访问边界单元时才会衰减。我们始终使用WOG，而不是球上游走（WOS）。WOG步进采用面相邻方式，因此无论材料分布如何，游走都以相同方式移动（图2(b)），而Robin边界通过 B_i 直接引入，无需调整边界层 [21], [30]。在2.5D/3D集成电路中密集堆叠的薄层内，WOS球无论如何也无法跨越材料界面，因此其半径会受到层厚度的限制。

Feynman–Kac表示。从查询时间反向行走 t，Feynman–Kac公式20 给出

$$T_i(t)=\mathbb{E}_i\left[\int_0^t W_s\,\frac{q(X_s,\,t-s)}{C(X_s)}\,ds+W_t\,T_0(X_t)\right], \tag{7}$$

where $\mathbb{E}_i$ denotes the expectation over path starting from cell i， q（i， t）是 i的第 $\boldsymbol{q}$（t）项， T_0 是预先规定的初始温度。路径在 $t=0$处终止，此时温度已知（图3(a)） [20]。在长期周期运行下，该端点不再存在（图3(b)）。因此，行走没有自然终止点，这使得周期案例明显比有限历史

案例更为困难，而这正是我们在本文中要解决的情况。

III. 方法

CycleWalker 分四个阶段求解周期性热查询。首先，它将 PSS 温度表示为无限历史反向游走（第 III-A 节）。由于每条路径现在都携带完整温度场，方差非常大，因此需要大量路径。随后，构建该游走，使其仅估计周期先验留下的残差，并采用无偏提前终止（第 III-B 节）。采用残差随机游走时，先验无需十分准确，因此可以在频域中低成本构建，再在细粒度图上进行细化（第 III-C 节）。最后，游走在热图上运行，其跳跃率在2.5D/3D集成电路的多尺度结构中始终有界（第III-D节）。

A. 无限历史周期随机游走表示

对于周期为 P，$\boldsymbol{q}(t+P)=\boldsymbol{q}(t)$ 的工作负载，我们寻求满足 $\boldsymbol{T}(t+P)=\boldsymbol{T}(t)$ 的 PSS。公式（7）中的反向游走必须在温度已知的时刻停止，但 PSS 并未指定这样的时刻。因此，PSS 中的游走必须向前追溯多个周期。为此，我们将起始时间无限推向过去。权重 W_s随着游走沿路径累积传导损耗而衰减，因此起始温度场的贡献趋于消失，并从期望中消失：

$$T_i(t)=\mathbb{E}_i\left[\int_0^{\infty} W_s\,\frac{q\big(X_s,\,(t-s)\bmod P\big)}{C(X_s)}\,ds\right]. \tag{8}$$

方程（8）是精确的，无需预热，但它会将计算每个路径所需的完整周期稳态成本全部计入。

B. 残差随机游走

有三个问题使方程（8）难以实用：每条路径都会读取完整功率，因此方差较大（图 4(a)）；时间范围是无限的，因此任何固定截断都会使估计产生偏差；而且不同查询的方差差异很大，因此手动选定路径数要么浪费计算，要么在不易察觉的情况下达不到目标精度。因此，随机游走只估计廉价先验的误差，利用俄罗斯轮盘赌无偏地终止长路径，并通过试探路径设定预算。

残差表述。令 $\boldsymbol{T}_p(t)$ 为 PSS 温度的任意 P周期近似，并令 $\boldsymbol{e}(t)=\boldsymbol{T}(t)-\boldsymbol{T}_p(t)$为其误差。将 $\boldsymbol{T}=\boldsymbol{T}_p+\boldsymbol{e}$ 代入方程（4）的离散化系统，其中电容矩阵为 $\boldsymbol{C}$，传导矩阵为 $\boldsymbol{A}$，可得

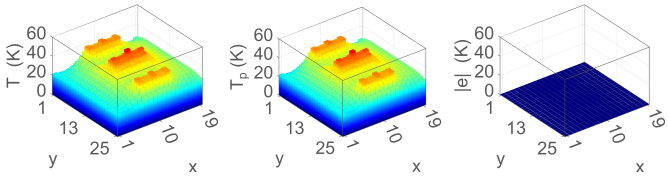
$$\boldsymbol{C}\dot{\boldsymbol{e}}+\boldsymbol{A}\boldsymbol{e}=\boldsymbol{r},\qquad \boldsymbol{r}=\boldsymbol{q}-\boldsymbol{C}\dot{\boldsymbol{T}}_p-\boldsymbol{A}\boldsymbol{T}_p, \tag{9}$$

其中 $\boldsymbol{r}$ 是先验未能解释的热输入。算子保持不变，且 $\boldsymbol{r}$ 继承了 P的周期性 $\boldsymbol{q}$和 $\boldsymbol{T}_p$，因此 $\boldsymbol{e}$ 是方程（9）的周期稳态，方程（8）可替换为 $\boldsymbol{r}$：

$$e_i(t)=\mathbb{E}_i\left[\int_0^{\infty} W_s\,\frac{r\big(X_s,\,(t-s)\bmod P\big)}{C(X_s)}\,ds\right]. \tag{10}$$

现在每条路径读取的是残差，而不是温度。通过将先验加回估计的误差，即可恢复温度。和 $\boldsymbol{T}$ 以及 $\boldsymbol{e}$ 的估计量形式相同，但方差不同。现在，各条路径仅因 $\boldsymbol{T}_p$ 的局部误差而不同（图4(c)），方差也随之降低。

无偏提前终止。我们对其应用俄罗斯轮盘赌



(a) $\boldsymbol{T}$ (b) $\boldsymbol{T}_p$ (c) $\boldsymbol{e}=\boldsymbol{T}-\boldsymbol{T}_p$

图 4 残差分解：$\boldsymbol{T}_p$ 使 $\boldsymbol{e}$ 保持较小且平滑。

公式（10）中的无限时域积分：如果 $W_s<\lambda$，

$$W_s\leftarrow\begin{cases}1, & \text{with probability } W_s,\\ 0, & \text{otherwise (the path terminates)},\end{cases} \tag{11}$$

这使 $\mathbb{E}[W_s]$保持不变，因此估计量仍然是无偏的，同时长路径会被提前终止。

自适应路径预算。少量试探路径即可估计所有查询中的最大样本方差 σ^2，并

$$N=\lceil\gamma\sigma^2/\varepsilon^2\rceil \tag{12}$$

通过 ε条路径达到目标标准误差，同时 $\gamma>1$ 防止低估。所有查询共享这一路径预算，因此没有任何查询达不到目标。用户只需指定目标精度 ε，N 便会自动确定。

估计量对先验只有两项要求：它必须是 P周期性的，这样 $\boldsymbol{r}$ 在方程（9）中便保持周期性。它必须关于时间分段线性，这样方程（10）中的残差积分就可以写成闭式。无需保证准确率：残差是根据给定的任意先验计算的，因此估计量仍然无偏；而较粗略的 $\boldsymbol{T}_p$ 只会增加所需路径数。

C. 周期性感知先验构建

先验还必须易于计算，而两种常规方法都不满足这一点。瞬态求解需要经过预热周期才能达到周期稳态，而稳态求解完全无法体现相位依赖性。因此，我们改用频域方法：周期源会使各谐波彼此解耦，从而可以通过粗图以较低成本达到周期稳态。

仅靠粗图并不足够，因为它会遗漏界面和 TSV 附近的局部细节。因此，我们通过在细粒度图上采样和平滑来恢复这些细节，并通过插值使结果在时间上连续。图 5展示了这四个阶段。先验构建如下

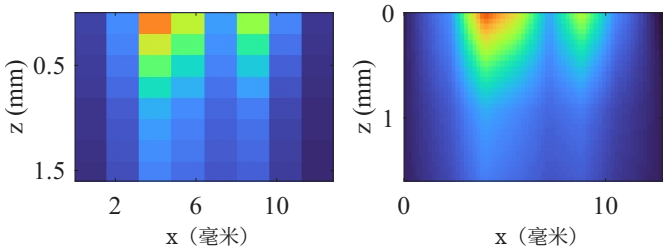
$$\boldsymbol{T}_c\xrightarrow{\text{sample}}\boldsymbol{T}_{\text{sp}}\xrightarrow{\text{smooth}}\boldsymbol{T}_{\text{sm}}\xrightarrow{\text{interpolate}}\boldsymbol{T}_p, \tag{13}$$

其中 $\boldsymbol{T}_c$ 是粗图周期解，而 $\boldsymbol{T}_{\text{sp}}$ 和$\boldsymbol{T}_{\text{sm}}$ 分别是平滑前后的细图相位样本。**粗图**。粗图以固定比例将每个细胞块合并为一个粗单元。这样会使系统规模缩小为原来的比例的三次方，而局部细节的损失可以接受，因为第III-B 节对先验的约束仅在形式上，而不涉及准确率。

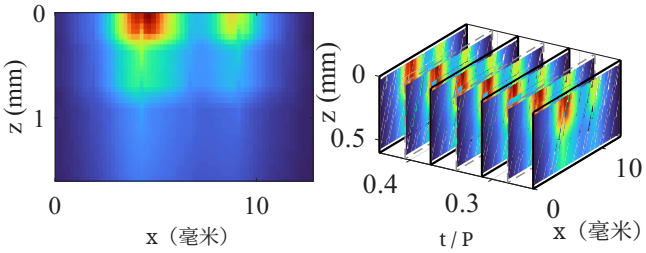
在此图上，我们直接在频域中求解周期稳态。由于 $\boldsymbol{q}$ 是 P-周期的，将 $\boldsymbol{q}$和周期稳态 $\boldsymbol{T}_c$ 展开为傅里叶级数，其中 $\omega=2\pi/P$ 会将方程（4）中的时间导数转化为乘以$\mathrm{j}\ell\omega$。算子是线性且时不变的，因此各谐波彼此解耦，每个谐波分别求解

$$(\boldsymbol{A}_c+\mathrm{j}\ell\omega\,\boldsymbol{C}_c)\,\boldsymbol{T}_\ell=\boldsymbol{q}_\ell,\qquad \ell=0,\dots,L, \tag{14}$$

其中，$\boldsymbol{A}_c$ 和 $\boldsymbol{C}_c$ 是粗图矩阵，而 $\boldsymbol{T}_\ell$ 和$\boldsymbol{q}_\ell$ 是温度和功率的第 ℓ个傅里叶系数。每个 ℓ 都是一次粗粒度规模的独立复数稀疏求解，将谐波求和还原后得到 $\boldsymbol{T}_c(t)$。这种周期性通过构造强制实现，而不是通过时间步进达到



(a) $\boldsymbol{T}_c$：粗图。 (b) $\boldsymbol{T}_{\mathrm{sp}}$：采样后 $\boldsymbol{T}_c$。



(c) $\boldsymbol{T}_{\mathrm{sm}}$：平滑后 $\boldsymbol{T}_{\mathrm{sp}}$。 (d) $\boldsymbol{T}_p$：沿时间切片。图 5 芯片堆叠切片上的先验构建。

因此无需初始条件，也无需模拟预热周期： $L+1$ 求解会立即落在周期稳态上。高次谐波也可以低成本地舍弃，因为$\|(\boldsymbol{A}_c+\mathrm{j}\ell\omega\boldsymbol{C}_c)^{-1}\|$ 会随 ℓ衰减，因此热响应呈低通特性，只需较小的 L即可。

细图细化。粗略求解 $\boldsymbol{T}_c$ 能够捕捉全局温度包络，但会平均掉热点附近的空间细节以及周期响应中剧烈的相位变化，而完整的细网格周期瞬态求解代价过高。我们利用层叠结构的一个结构特性：垂直耦合强于横向耦合，因此，仅对垂直方向进行隐式处理的平滑器既经济又高效；由于它将各列解耦，因此能够自然地映射到 GPU 上。

我们将周期划分为 M 个长度为 $\Delta t = P/M$ 的时间间隔，并将 $\boldsymbol{T}_c$ 在每个相位投影到细图上，

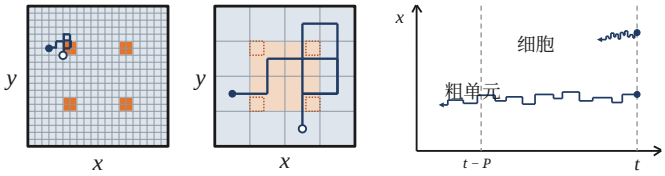
$$\boldsymbol{T}_{\mathrm{sp}}(m) = \boldsymbol{\Pi}\,\boldsymbol{T}_c(m\Delta t) \qquad m = 0,\ \ldots\ldots,\ M-1, \tag{15}$$

其中， $\boldsymbol{\Pi}$ 通过三线性插值将粗网格单元中心映射到细网格单元中心。从此处开始，参数 m表示相位索引，而非时间，因此 $\boldsymbol{T}_{\mathrm{sp}}(m)$表示相位 $m\Delta t$处的细图温度向量，索引按 M取模。接着，我们将细图矩阵 $\boldsymbol{A} = \boldsymbol{V} + \boldsymbol{H}$拆分，其中 $\boldsymbol{V}$包含对角项和垂直耦合， $\boldsymbol{H}$包含横向耦合；从 $\boldsymbol{T}_{\mathrm{sp}}(m)$开始，执行固定次数的周期性后向欧拉迭代

$$\left(\boldsymbol{V} + \frac{\boldsymbol{C}}{\Delta t}\right)\boldsymbol{T}_{\mathrm{sm}}(m) \leftarrow \boldsymbol{q}(m) + \frac{\boldsymbol{C}}{\Delta t}\boldsymbol{T}_{\mathrm{sm}}(m-1) - \boldsymbol{H}\,\boldsymbol{T}_{\mathrm{sm}}(m), \tag{16}$$

其中， $\boldsymbol{q}(m)$表示对以 $m\Delta t$结束的时间间隔内的 $\boldsymbol{q}$ 取平均值。每次迭代都会根据前一次迭代更新所有相位，因此垂直耦合采用隐式处理，而横向耦合和前一相位则滞后处理。

由于 $\boldsymbol{V}$ 仅耦合堆叠在同一列中的单元格，因此每一列和每个相位彼此独立：一次迭代相当于批量执行 M 乘以列数个相互独立的三对角方程求解，并作为单个 GPU内核运行。因此，固定的用户指定扫描次数能够以确定性成本得到细网格周期温度场 $\boldsymbol{T}_{\mathrm{sm}}$ ，该温度场构成平滑的确定性分量，其方程残差将驱动后续的 RandomWalk。



(a) 通孔解析与特征对齐。 （b）在反向时间中的随机游走。

图 6 特征对齐的网格界定跳跃率；通孔作为材料处理。

比较图5(b) 与图5(c) 可以看出扫描所恢复的内容：热点变得更加清晰，层间界面和 TSV 柱重新出现，而这些在粗略求解中全部丢失。

插值。 $\boldsymbol{T}_p(t)$ 对 $\boldsymbol{T}_{\mathrm{sm}}(m)$ 在相邻相位之间进行线性插值（图 5(d)）。随后 $\boldsymbol{T}_p$ 和 $\boldsymbol{q}$ 在每个区间内均为常数，而残差在该区间内随时间呈线性变化。

因此，先验在时间上具有 P周期性且呈分段线性，正如第III-B节所要求的那样，残差积分无需进一步进行时间离散化或采样。其成本为 $L+1$ 次粗网格求解和少量列并行扫描，并且它吸收了主要分布，因此随机游走只需采样一个小得多的场。

D. TSV 感知的多尺度图构建

若用共形网格解析每个微米尺度 TSV，随机游走会慢得难以接受，因为跳跃率对单元厚度高度敏感。根据方程（5），厚度为 Δz 的单元，其热导按 $k/\Delta z$ 缩放，电容按 $c_v\Delta z$缩放，因此

$$\frac{\sum_j G_{ij}}{C_i} \sim \frac{k}{c_v\,\Delta z^2}. \tag{17}$$

共形网格会将局部跳跃率提高几个数量级。每一步只能推进极短的反向时间，而回溯一个周期则需要极其庞大的步数（图 6(a)和 6(b)）。通孔也不能省略：通孔阵列沿通孔方向的导热能力远高于横向导热能力，而这种方向性行为主导着堆叠封装中的热量散逸。图结构必须限制跳跃率，同时保留这种各向异性。

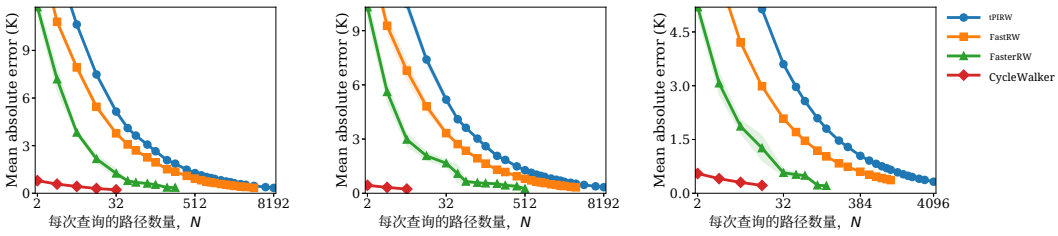
特征对齐离散化。我们使用非均匀笛卡尔网格来保留材料界面和热源边界，且不会细化到通孔的尺度。因此，单元尺寸由网格分辨率决定，而非由最小 TSV 决定，相应地，跳跃率也受到限制。

子网格 TSV 均质化。我们将通孔视为材料属性，而不是几何结构。设 χ为单元横截面中通孔所占的面积分数，k_h 和 k_c 分别为基体材料和铜填充物的热导率。有效横向热导率 κ和轴向热导率为

$$\kappa = \left(\frac{1-\chi}{k_h} + \frac{\chi}{k_c}\right)^{-1}, \qquad \eta = (1-\chi)\,k_h + \chi\,k_c, \tag{18}$$

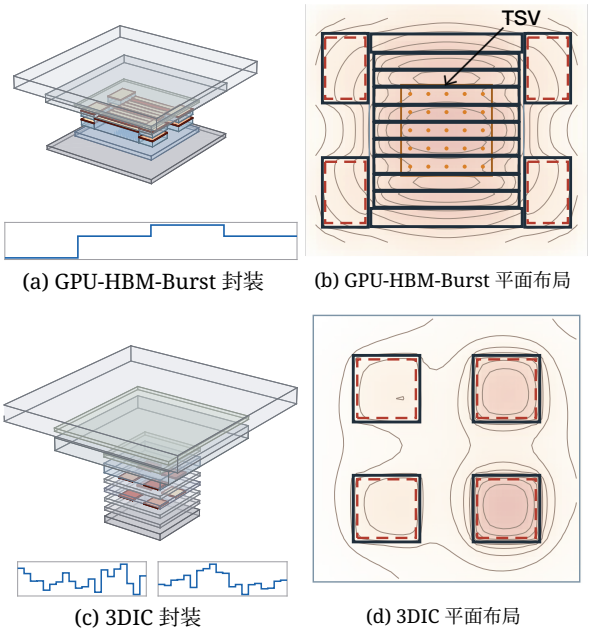
体积热容采用线性混合。两种材料均为各向同性，但该单元并非如此：热沿通孔方向的传导比横穿通孔更容易。

我们按常规方式计算热导 G ，根据 κ 和 η ，因此游走无需针对均质化单元制定特殊规则：图 6(a) 中阴影单元包含通孔，但无需将其单独解析。这种层状结构也正是第III-C 节垂直块平滑所利用的。



(a) 3DIC-FastCycle (b) 3DIC-NonAligned (c) GPU-HBM-Burst

图 7 三种场景下误差与路径预算的关系。



(a) GPU-HBM-Burst 封装 (b) GPU-HBM-Burst 平面布局

(c) 3DIC 封装 (d) 3DIC 平面布局

图 8 基准测试案例。(a)(c) 封装爆炸视图，各堆叠结构下方标示单周期总功耗。(b)(d) 下层有源平面的平面布局：单元轮廓位于 TSV 阵列上方，并根据该平面上的参考稳态温度进行着色。

IV. 实验

我们在三种周期性热分析场景中评估*CycleWalker*：首先，在相同的精度约束下与三种随机基线进行比较；然后，通过对其估计量进行累积消融来评估；最后，在 TSV 感知图构建上进行评估。

A. 实验设置

测试案例。图8 展示了这两种封装。3DIC- NonAligned 和 3DIC-FastCycle 是相同的十层 3D 堆叠结构（35,600 个单元格），分别采用 0。2s 和 0。08s 周期，前者的查询相位落在切换时刻之间。GPU- HBM-Burst 是一种 2.5D TSV 中介层封装（82,420 个单元格），包含一个 31×31 TSV 阵列和位于三个有源平面上的 19 个功率单元，由突发式工作负载驱动。功耗模型固定在 300K参考温度下，因此该问题是线性的。

实现。准确率以相同离散算子上的 512 次谐波确定性解为基准进行测量，相对残差低于 1.1×10^{-13}。四种方法共用一个 CUDA 实现，并使用同一台主机（Intel Core i5-12500H、NVIDIA RTX 3050 Ti 笔记本电脑 GPU）；参考结果离线生成，不计入任何计时。***CycleWalker***采用 8 次谐波粗略先验、全路径方程残差估计和无偏俄罗斯轮盘赌。

表 I 两个目标下的等精度比较。

ε (K)	方法	误差±标准差 (K)	N	时间 (s)	加速比
3DIC-NonAligned					
0.2	tPIRW	0.190 ± 0.008	24576	806.97	1.00
	FastRW	0.213 ± 0.006	8192	134.17	6.01
	FasterRW	0.193 ± 0.059	1260	99.62	8.10
	CycleWalker	0.215 ± 0.006	**9**	**17.09**	**47.22**
0.4	tPIRW	0.380 ± 0.002	6144	199.69	1.00
	FastRW	0.419 ± 0.011	2048	41.32	4.83
	FasterRW	0.355 ± 0.141	384	38.27	5.22
	CycleWalker	0.362 ± 0.015	**3**	**14.71**	**13.58**
3DIC-FastCycle					
0.2	tPIRW	0.191 ± 0.011	24576	746.63	1.00
	FastRW	0.200 ± 0.004	12288	178.74	4.18
	FasterRW	0.204 ± 0.018	1536	108.19	6.90
	CycleWalker	0.214 ± 0.008	**32**	**17.52**	**42.62**
0.4	tPIRW	0.381 ± 0.009	6144	182.56	1.00
	FastRW	0.397 ± 0.013	3072	54.42	3.35
	FasterRW	0.414 ± 0.075	236	28.38	6.43
	CycleWalker	0.413 ± 0.017	**8**	**17.80**	**10.26**
GPU-HBM-Burst					
0.2	tPIRW	0.186 ± 0.004	12288	3309.53	1.00
	FastRW	0.210 ± 0.004	3072	258.76	12.79
	FasterRW	0.210 ± 0.029	128	116.50	28.41
	CycleWalker	0.218 ± 0.003	**16**	**43.63**	**75.85**
0.4	tPIRW	0.405 ± 0.011	2560	704.64	1.00
	FastRW	0.386 ± 0.009	896	128.60	5.48
	FasterRW	0.438 ± 0.089	104	89.70	7.86
	CycleWalker	0.411 ± 0.009	**4**	**56.09**	**12.56**

FastRW 和 FasterRW 使用 0.5 的截断值。

基线方法。tPIRW [20]是一种不使用先验的周期性费曼–卡茨游走，因此其截断残差会随绝对温度而缩放，并要求截断值为 3×10^{-4}。**Fas-tRW** [24]加入了周期性终端先验、跨查询尾部复用和贝叶斯融合。**FasterRW** [25]进一步加入协方差估计和周期性克里金细化。

指标。准确率和成本分别以平均绝对误差（MAE）和加速比表示，

$$\mathrm{MAE}=\frac{1}{Q}\sum_{j=1}^{Q}\left|T_j-T_j^{\mathrm{ref}}\right|,\qquad \mathrm{Speedup}=\frac{T_{\mathrm{tPIRW}}}{T_{\mathrm{method}}}. \tag{19}$$

加速比采用端到端求解器时间，因为先验辅助方法会将成本转移至确定性阶段；图构建和参考解不计入其中。方差是查询的每路径方差，~~SE~~ $=\sqrt{\mathrm{Var}/N}$ 是其标准误差。

B. 等精度比较

针对每个名义目标 ε，在三个校准种子上选择预算，然后在五个独立的正式种子上评估所选配置。Err 报告

表 II 与确定性先验自身准确率相匹配。

方法	误差±标准差 (K)	N	时间 (s)	加速比
周期先验	0.0130	–	1266.86	1.00
CycleWalker	$\mathbf{0.0126 \pm 0.0003}$	**5120**	**457.61**	**2.77**

表 I Cumulative ablation of the three mechanisms.

Case	Cfg.	N	Var. (K^2)	|偏差| (K)	Len. (步数)	平均绝对误差时间 (K)	(s)
NonAligned	截断	2048	1377.7	<4.0e−2	3923	0.6214	17.7
	+残差	2048	0.576	3.4e−2	3923	0.0342	31.7
	+轮盘赌	2048	0.704	<9.1e−4	6585	0.0144	39.4
	+自适应	4096	0.705	<6.4e−4	6587	**0.0102**	55.6
FastCycle	截断	2048	1392.4	<4.3e−2	3923	0.6299	17.6
	+残差	2048	1.849	3.0e−2	3923	0.0349	27.7
	+轮盘赌	2048	2.291	<1.8e−3	6586	0.0260	38.4
	+自适应	8192	2.292	<8.8e−4	6586	**0.0136**	112.6

这五次正式运行的均值和样本标准差。由于预算选择和正式评估使用不同的种子，正式运行的平均绝对误差可能高于名义校准目标。

如表 I 所示，*CycleWalker* 在全部六个运行点上都是最快的方法。它比 tPIRW 快 $10.26\times$–$75.85\times$，而且仍比最强的基准方法 FasterRW 快 $1.59\times$–$6.18\times$。

目标越严格，观察到的优势越大。对于方差主导的蒙特卡洛估计量，随机误差的期望缩放关系为

$$\text{error} \propto N^{-1/2}, \tag{20}$$

其中 N 为路径数。在图 7中，tPIRW 和 FastRW 紧密遵循这一趋势，而 FasterRW 和 *CycleWalker* 则表现出依赖于方法和案例的有限预算行为。因此，收紧目标并不会使每种方法的预算都按相同倍数增加。

在选定的六个运行点上，当名义目标从 *CycleWalker* 的加速比在名义目标从 0.4K 收紧至 0.2 K时，CycleWalker 相对于 tPIRW 的加速比有所提高。方法之间的差异不能仅归因于每路径方差，因为各方法在先验、残差估计器、协方差处理、融合和克里金法方面也存在差异。第 IV-C 节分离考察了其中若干机制。

表 II 将 GPU-HBM-Burst 目标再收紧一个数量级，达到 0。0130K，这与仅使用确定性先验在 1267s内达到的平均绝对误差相当。在 $N = 5120$，*CycleWalker*在 0.0126 ± 0.0003K 个种子上经过 458 ± 12 s获得了 $2.77\times$ 的运行时间缩减，同时匹配了仅先验参考的平均绝对误差。

C. 机制分析

在三个种子上比较了 Fourcumulative 在 NonAligned 和 FastCycle 上的配置，每种配置都在前一种配置的基础上增加一个机制。**截断**会在截断点停止每条路径。**+残差** 增加全路径方程残差估计。**+轮盘赌**增加无偏的俄罗斯轮盘赌。**+自适应** 增加由试运行驱动的采样，并构成完整的 *CycleWalker*。所有配置均共享一个包含 8 个谐波的先验和 32 次平滑迭代；前三种配置固定为 $N = 2048$，截断值为 0。5，而**+自适应**会设定自身的预算。在表 III中，写作$< x$ 的偏差低于三个种子能够分辨的范围，与零一致。

残差估计。从**截断**到 **+残差**，见表 III，其中路径数、截断值和平均长度均

表 IV 固定精度目标下的先验敏感性。

方法	先验	N	Var. (K^2)	MAE (K)	Time (s)	加速比
tPIRW	none	12288	—	0.186	3309.5	1.00×
CycleWalker	H8+0	4096	222.6	0.186	372.4	8.9×
	H8+32	32	1.27	0.158	43.0	**77.0**×

表 V 显式解析 TSV 的成本。

	已解析	子网格（我们的方法）	比率
TSV 核心 ν_{p99} (s^{-1})	4.40e5	1.97e4	22.3×
TSV 核心 $1/\nu_{\text{p99}}$ (µs)	2.27	50.7	0.045×
全局 ν_{p50} (s^{-1})	8471	1468	5.77×
单元格	175,068	82,420	2.12×
平均绝对误差 (K)，$N = 4096$	0.481	0.451	–
每查询标准误 (K)，$N = 4096$	0.545	0.546	–
回看期数（周期）	45.84	45.70	1.00×
每期步数	1788.9	**356.4**	**5.02**×
GPU 时间 (s)	1.990	**0.379**	**5.25**×

保持相等时，每路径方差下降了 $753\times$–$2392\times$，平均绝对误差下降了 $18\times$。由于唯一不同的是采样量，我们将全部收益归因于此，并得出结论：先验无需准确，只要吸收主导分量即可。

无偏俄罗斯轮盘赌。从 **+残差** 到 **+轮盘赌**在表 III 中，保持 $N = 2048$不变时，偏差 3.4×10^{-2}K和 3.0×10^{-2}K 均降至种子值无法分辨的程度，路径变得 $1.68\times$ 更长，平均绝对误差降低了 $2.4\times$和 $1.3\times$。由此得出结论：轮盘赌消除了截断偏差。

自适应采样。从 +轮盘赌切换到 +自适应，在表 III 中，N 提高到 4096 和 8192，以达到 0.035 K 的标准误差目标，平均绝对误差降低了 $1.4\times$和 $1.9\times$。由于此处尚未手动设定 N，但仍满足 0.035 K 的标准误差目标，因此我们得出结论：自适应采样会自动调整N，用户只需指定目标。先验敏感性。表 IV 将目标固定为 tPIRW 在 GPU-HBM-Burst 上达到的 0.186K，并将 HL+s记为 L谐波和 s扫描。减少扫描次数会增大方差 $175\times$ 和预算 $128\times$，增加 $8.7\times$ 运行时间，但仍保留相较于 tPIRW 的 $8.9\times$余量。由此得出结论：先验不可或缺，但不必精确。

D. TSV 感知图构建

表V在相同的估计量和零先验条件下解析了 TSV，并以五个随机种子的中位数测量时间 $N = 64$，准确率为 $N = 4096$。TSV 核心中的跳跃率上升 $22.3\times$，全局跳跃率上升 $5.77\times$，因此每个周期需要 $5.02\times$ 更多步数和$5.25\times$ 更多 GPU 时间，而报告的平均绝对误差仅相差 0。03K。我们得出结论：子网格均质化以一半的单元格数量和约五分之一的 GPU 时间，达到了与解析图相当的准确率。

V. 结论

本文介绍了*CycleWalker*，这是一种随机游走求解器，仅在被查询的位置和时间点计算 2.5D/3D集成电路的周期稳态。将周期稳态表述为无限历史反向游走，消除了初值瞬态游走中的人为预热。仅对周期先验的方程残差进行采样，使每路径方差随先验误差而非温度变化。无偏的俄罗斯轮盘赌消除了截断偏差，而试运行驱动的预算则无需手动调参即可满足指定的误差目标。子网格

TSV均匀化使跳跃率在堆叠结构的多尺度结构中保持有界。在准确率相同的情况下，*CycleWalker* 相较于 tPIRW 可将端到端求解器时间缩短 $10.26\times$–$75.85\times$；在更严格的 0.2K 误差目标下，则可缩短 $42.62\times$–$75.85\times$ ，使其成为 2.5 D/3D集成电路周期性热分析的实用工具，尤其适用于仅查询少数位置和相位的情况。

参考文献

[1] Y. Ma、L. Delshadtehrani、C. Demirkiran 等，“TAP-2.5D：面向2.5D系统的热感知芯粒布局方法”，发表于 DATE 会议论文集，2021年，第1246–1251页。[2] Q. Wang、X. Li、T. Jia 等，“ATPlace2.5D：面向大规模2.5D集成电路的解析式热感知芯粒布局框架”，发表于 ICCAD 会议论文集，2024年，第39:1–39:9页。[3] F. Eris、A. Joshi、A. B. Kahng 等，“利用2.5D系统中的热感知芯粒组织以回收暗硅”，发表于 DATE 会议论文集，2018年，第1441–1446页。[4] J. Cong 和 Y. Zhang，“三维集成电路的热通孔规划”，发表于 ICCAD 会议论文集，2005年，第745–752页。[5] A. Sridhar、A. Vincenzi、M. Ruggiero 等，“3D-ICE：采用层间液冷的三维集成电路快速紧凑型瞬态热建模”，发表于 ICCAD 会议论文集，2010年，第463–470页。[6] P. Liu、H. Li、L. Jin、W. Wu、S. X.-D. Tan 和 J. Yang，“用于运行时温度跟踪与管理的快速热仿真”，IEEE TCAD，第25卷，第12期，第2882–2893页，2006年。[7] W. Kang、S. Chung、J. Y. Kim、Y. Lee、K. Lee、J. Lee、K. G. Shin 和 H. S. Chwa，“DNN-SAM：面向实时目标检测的DNN拆分合并执行”，发表于 RTAS 会议论文集，2022年，第160–172页。[8] S. Moulik、A. Sarkar 和 H. K. Kapoor，“TARTS：温度感知的实时截止期限分区公平调度器”，《Journal of Systems Architecture》，第112卷，第101847页，2021年。[9] P. Liu、Z. Qi、H. Li、L. Jin、W. Wu、S. X.-D. Tan 和 J. Yang，“面向架构级动态热管理的快速热仿真”，发表于 ICCAD 会议论文集，2005年，第639–644页。[10] Q. Wang、T. Zhu、Y. Lin、R. Wang 和 R. Huang，“ATSim3D：面向考虑非线性漏电与电导率的异构三维集成电路的高精度热仿真器”，发表于 ISEDA 会议论文集，2024年。[11] T.-Y. Wang 和 C. C.-P. Chen，“3-D Thermal-ADI：线性时间芯片级瞬态热仿真器”，IEEE TCAD，第21卷，第12期，第1434–1445页，2002年。[12] R. Telichevesky、K. S. Kundert 和 J. K. White，“基于无矩阵克雷洛夫子空间方法的高效稳态分析”，发表于 DAC 会议论文集，1995年，第480–484页。[13] S. R. Sarangi、G. Ananthanarayanan 和 M. Balakrishnan，“LightSim：漏电感知的超快速温度仿真器”，发表于 ASPDAC 会议论文集，2014年，第855–860页。[14] V. A. Chhabria、V. Ahuja、A. Prabhu 等，“使用卷积编码器-解码器网络进行热分析与IR压降分析”，发表于 ASPDAC 会议论文集，2021年，第690–696页。[15] D. Zhang、D. Niu、Z. Jin 等，“面向2.5D集成电路快速热预测的新型频率-空间域感知网络”，发表于 DATE 会议论文集，2025年，第1–7页。[16] Z. Liu、Y. Li、J. Hu 等，“DeepOHeat：基于算子学习的三维集成电路设计超快速热仿真”，发表于 DAC 会议论文集，2023年，第1–6页。[17] Y.-F. Yang、W.-S. Wang、Y.-J. Lee 等，“面向三维集成电路的高效基于机器学习的瞬态热预测”，发表于 ASPDAC 会议论文集，2025年，第628–634页。[18] T. Zhu、Q. Wang、Y. Lin 等，“FaStTherm：考虑非线性效应的快速稳定全芯片瞬态热预测器”，发表于 ICCAD 会议论文集，2024年，第144:1–144:9页。[19] Y. Liang、W. Yu 和 H. Qian，“一种用于集成电路三维热分析的混合随机游走算法”，发表于 ASPDAC 会议论文集，2014年，第849–854页。[20] Z. Dong、Y. Meng、C. Yan、Z. Bi、K. Zhu、S.-G. Wang、D. Zhou 和 X. Zeng，“采用混合边界条件进行瞬态热分析的高精度随机求解器”，IEEE TCPMT，第16卷，第2期，第351–364页，2026年。[21] L. Yang、C. Ding、C. Yan 等，“采用傅里叶热传递罗宾边界条件进行稳态热分析的高精度随机求解器”，发表于 ICCAD 会议论文集，2022年，第1–9页。[22] Z. Dong、L. Yang、C. Ding、C. Yan、Z. Bi、S.-G. Wang、D. Zhou 和 X. Zeng，“pPIRW：一种用于带罗宾边界条件的稳态热仿真的高效高精度预计算路径积分随机游走求解器”，IEEE TCAD，第44卷，第4期，第1489–1502页，2025年。[23] Y. Luo、Y. Wang、Z. Dong、C. Yan、Z. Bi、K. Zhu、S.-G. Wang 和 X. Zeng，“带罗宾边界条件的稳态热分析高精度随机求解器的 GPU 加速”，发表于 ISEDA 会议论文集，2025年，第287–292页。[24] Z. Wang、T. Hou、C. Wang、Z. Zhuang、T.-Y. Ho、F. Farnia 和 B. Yu，“FastRW：一种用于稳态热分析的高效随机游走方法”，发表于 DATE 会议论文集，2026年。[25] Z. Wang、T. Hou、C. Wang、Z. Zhuang、L. Jin、T.-Y. Ho、F. Farnia 和 B. Yu，“FasterRW：一种高效的先验辅助随机游走求解器”，2026年，预印本。 [在线]。可访问：https://shiningsord.github.io/

[26] Y. Zhan、B. Goplen 和 S. S. Sapatnekar，《纳米级集成电路的电热分析与优化技术》，载于*ASPDAC* 会议论文集，2006年，第219–222页。[27] A. Sridhar、M. M. Sabry 和 D. Atienza，《液冷集成电路的半解析热建模框架》，*IEEE TCAD*，第33卷，第8期，第1145–1158页，2014年。[28] Z. Yuan、P. Shukla、S. Chetoui等，《PACT：面向新兴集成与冷却技术的可扩展并行热仿真器》，*IEEE TCAD*，第41卷，第4期，第1048–1061页，2022年。[29] S. Lan、M. Tang、L. Chen等，《用于集成电路与封装高效热仿真和设计的热阻网络导数（TREND）模型》，载于*DAC* 会议论文集，2024年，第113:1–113:6页。[30] Z. Dong、Y. Luo、Y. Meng等，《LCTMwalk：通过改进紧凑热模型电路网络上的随机游走，对液冷2.5D/3D集成电路进行GPU加速瞬态热仿真》，载于*ICCAD* 会议论文集，2025年，第1–9页。