Machine Learning
Final Project

LO Wei-Hsiang (羅瑋翔)
314652021
xianglo9121.sc14@nycu.edu.tw

December 4, 2025

I. Future AI Capability

To overcome the bottlenecks posed by Moore's Law and Amdahl's Law, this project aims to simulate chip heat distribution. Through high-fidelity simulation, we can optimize layouts to reduce thermal throttling. This prevents single-processor underclocking (addressing Amdahl's Law) and allows for higher processor density (addressing Moore's Law). Furthermore, Reinforcement Learning is employed to discover optimal layouts. These improvements enable significantly higher clock rates, increasing throughput by encapsulating more processors per die.

Beyond merely simulating heat, the future AI will act as an autonomous physicist, creating a real-time Digital Twin of the chip. Unlike current TCAD tools that take hours to simulate, this AI will provide instantaneous multi-physics feedback, shifting the paradigm from 'verification after design' to 'inverse design', where the AI proactively optimizes the layout for thermal limits before a human engineer even finalizes the circuit.

II. Essential Ingredients

- **Data:** We require multi-scale physical field data, spanning from the microscopic to the macroscopic level. This includes tensor data encompassing gradients and boundary conditions. The dataset combines high-fidelity simulation data and sparse real-world sensor data. Since measuring every point on a chip is impossible, we use numerical methods like FEM or FDM to generate high-fidelity training data. Additionally, sparse sensor data from real chips is used to calibrate predictions, ensuring the results reflect real-world conditions (Sim-to-Real)
- **Tools:** A **differentiable physics engine** is needed to solve this problem. AI should incorporate physical laws rather than rely solely on data fitting. This requires PINNs and neural operators (e.g., Fourier neural operators), enabling the AI directly solve PDEs. And since chips

are inherently non-Euclidean graph structures, Geometric Deep Learning (e.g., Graph Neural Networks) will be an essential ingredient to encode the complex 3D topology and connectivity of interconnects, going beyond simple 2D grid representations.

- **Learning Setup:** hybrid learning paradigm
  - **Self-Supervised Learning:** Enable the AI to learn fundamental conservation laws (e.g., conservation of energy and charge) without the need for labeled data.
  - **Reinforcement Learning:** An AI agent explores the design space and searching for the optimal layouts based on feedback from the physics engine to increase the clock rate of a processor.

Additionally, uncertainty quantification is critical. The system must recognize out-of-distribution scenarios (e.g., a novel 3D packaging architecture) and automatically trigger high-fidelity simulations to retrain itself (Active Learning), ensuring reliability.

III. Involved Machine Learning Types

Implementing this vision requires a deep integration of physics-informed-neural-network, reinforcement learning, and supervised-learning.

Pure supervised-learning struggles to handle unseen physical scenarios and relies heavily on expensive simulation data (requiring vast amounts of data). We require PINNs to act as fast, physics-compliant environment/simulator, and RL to act as a designer for decision optimization.

The input data is chip geometry layouts, material parameters (σ, k), and operating conditions (voltage V). The target is physical fields that satisfy PDE constraints (electric potential φ and temperature T). The RL reward function is composed of performance gains (clock rate) and penalties (peak temperature).

The combination is not just about using PINN as a fast simulator, but exploiting its differentiability. Unlike traditional solvers, a PINN allows gradients to flow from the target metric (e.g., min temperature) back to the design parameters (layout coordinates). This enables Gradient-based Inverse Design, which is significantly more efficient than standard trial-and-error Reinforcement Learning. Additionally, Meta-Learning could be employed to allow the model to rapidly adapt to new chip architectures with minimal data.

IV. Solvable Model Problem

1. **Problem Formulation:**

   This study aims to solve the steady-state electro-thermal multi-physics coupling problem on a 2D chip. The goal is to train a neural network to simultaneously predict the electric potential distribution and temperature distribution given a specific chip architecture and operating voltage.

   The problem is defined by two systems of coupled PDEs on the domain $[-1, 1] \times [-1, 1]$:

   - **Electrical (Charge conservation):** Governed by the Laplace equation.

   $$\nabla \cdot (\sigma \nabla \varphi) = 0$$

   - **Thermal (Energy conservation):** Governed by the Poisson equation. The heat source includes both Logic gate switching power and joule heating.

   $$k\nabla^2 T + Q_{\text{logic}} + \underbrace{\frac{1}{\sigma} \|J\|^2 = 0}_{\text{Coupling Term}}$$

   - **Boundary conditions:**
     - voltage: $\varphi|_{\text{top}} = 1.0V$, $\varphi|_{\text{bottom}} = 0.0V$
     - Temperature: Satisfying the Dirichlet condition $T|_{\partial\Omega} = 0$

2. **Methodology:**

   The implementation of the model proposes a coupled-electro-thermal Hard-PINN architecture.

   - **Hard-PINN:** This model adopts a hard constraint technique

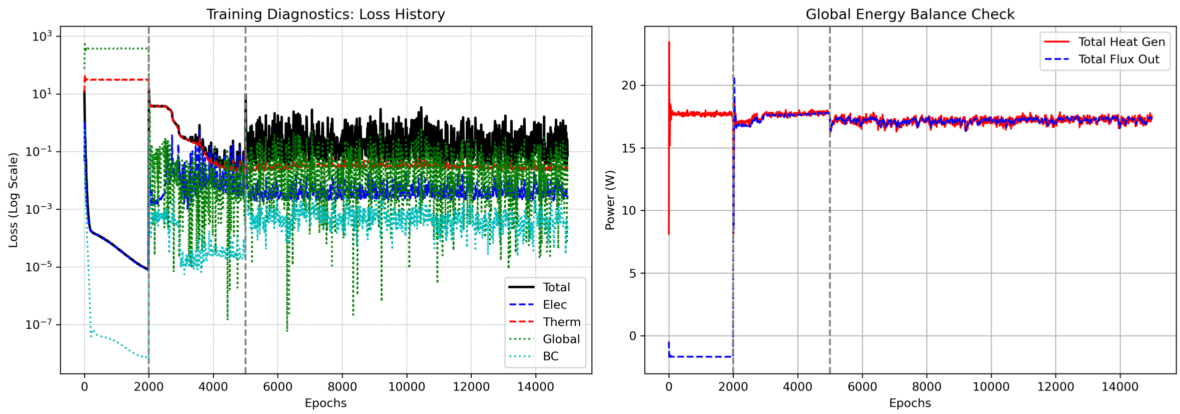
   $$\tilde{T}(x, y) = \mathbb{N}(x, y) \times (1 + x)(1 - x)(1 + y)(1 - y) + 0$$

     to guarantee the temperature is exactly 0 at the boundaries.
   - **Global energy conservation:** Forces the model to adhere the first law of thermodynamics-the total heat generated inside the chip must equal to the total heat flux dissipating through the boundaries.
   - **Curriculum learning strategy:** To address instability in multi-physics training, a three-phase training strategy is designed:
     a. Phase 1: Freeze the thermal equation; focus on training potential φ to stabilize current distribution. (Enabling the model to learn the electric potential)
     b. Phase 2: Introduce the thermal equation to learn local temperature gradient shape. (Since electric potential is almost as same as the true value, and the heat source is steady, the model can learn the temperature)

c. Phase 3: Enable the global integral constraint to correct the overall energy balance. (Correct the model, forcing it to satisfy energy conservation)

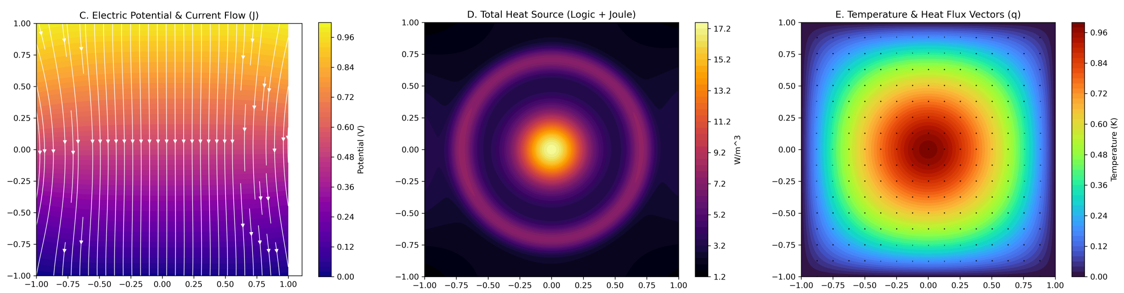
3. **Implementation and Result:**

- **Implementation:** In the file `final.ipynb`.
- **Result:**



A. **Loss History:** In 0-2000 epochs, electrical and BC loss decrease. Following in 2000-5000 epochs, thermal loss decreases. Finally, the global loss decreases and tend to be stabilizes in 5000-15000 epochs.

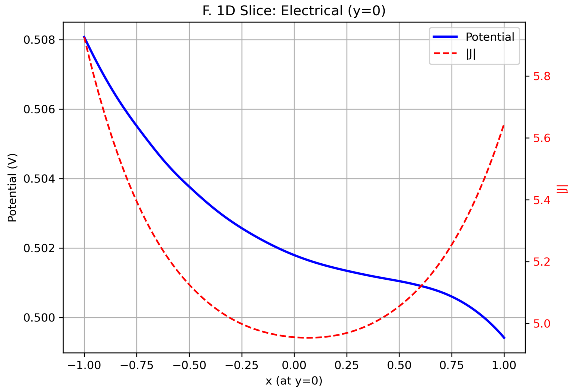
B. **Global Energy Balance Check:** It shows us that the model satisfies the first law of thermodynamics after training.

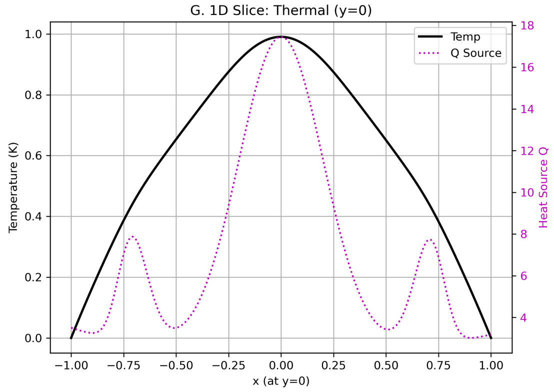


C. **Electric Potential:** The model satisfies the BCs we assumed (top=1V bottom=0.0V), while the streamlines flow from top to bottom.

D. **Total Heat Source:** The figure perfectly matches our assumption. The bright part in the center is the core (CPU) while the outer bright circle is the cache according to `q_dyn_1` and `q_dyn_2` in the code.

E. **Temperature and Heat Flux:** The temperature in the center is the highest (red), and it decays when approaching the boundaries (yellow→green→light blue→dark blue). And the arrows (heat flux) point from center toward the boundaries.





F. **Electrical:** The potential hovering around $0.5V$ is reasonable because $y = 0$ corresponds to the exact center of the chip.

   In the central region, the current density remains relatively uniform.

G. **Thermal:**

   A. The central peak of purple dashed line corresponds to Core A and the two smaller side peaks of purple dashed line represent Core B.

   B. The peak of the black line corresponds to the strongest heat source, Core A.

- **Error Analysis:** On an unseen test grid, the mean residual of the thermal PDE converged to a negligible level, with boundary temperature error remaining at exactly 0.

4. **Discussion:**

Via implementing this model,

- I discovered that simultaneously training electric and thermal fields is extremely difficult, as the model tends to diverge. Curriculum Learning is proved to be an effective strategy for solving such stiff PDE systems
- Future AI simulators must enforce constraints at both microscopic differential and macroscopic integral scales to solve the numerical drift issues and raise the accuracy.

Key challenges revealed,

- We observed that simultaneous training of electric and thermal fields leads to divergence. This is due to the disparate scales and convergence rates of different physical fields. Minor noise in the electric potential is amplified through the non-linear joule heating term ($|\nabla\varphi|^2$), destabilizing the thermal field.
- Relying solely on local PDE residuals allows the model to learn correct gradient shapes but often fails to satisfy global energy conservation. This demonstrates that PINNs cannot solely rely on differentiation. Future AI must incorporate Global Constraint Layers. However,

enforcing these constraints on complex 3D packaging structures poses a significant computational challenge, as calculating precise integrals over irregular 3D boundaries is expensive.