Assignment 10

314652021

Q1

Consider a forward SDE:

$$dx_t = f(x_t, t)\, dt + g(x_t, t)\, dW_t$$

Show that the corresponding PF ODE is written as:

$$dx_t = \left[f(x_t, t) - \frac{1}{2}\frac{\partial}{\partial x} g^2(x_t, t) - \frac{g^2(x_t, t)}{2}\frac{\partial}{\partial x}\log p(x_t, t) \right] dt$$

Proof: Consider a forward SDE:

$$dx_t = f(x_t, t)\, dt + g(x_t, t)\, dW_t$$

With the corresponding Fokker–Planck equation:

$$\frac{\partial}{\partial t} p(x, t) = -\frac{\partial}{\partial x}(f(x, t)p(x, t)) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(g^2(x, t)p(x, t))$$

We can rewrite the Fokker–Planck equation as follows:

$$\begin{aligned} \frac{\partial}{\partial t} p &= -\frac{\partial}{\partial x}(fp) + \frac{1}{2}\frac{\partial^2}{\partial x^2}(g^2 p) \\ &= \frac{\partial}{\partial x}\Bigg(- fp + \frac{1}{2}\Big(p\frac{\partial}{\partial x} g^2 + g^2 p \frac{\partial}{\partial x}\log p \Big) \Bigg) \end{aligned}$$

Then rewrite the FPE as a continuity equation:

$$\frac{\partial}{\partial t} p = -\frac{\partial}{\partial x}(\tilde{f} p)$$

Replace the term $\partial_t p$ with $-\partial_x(\tilde{f}p)$ and integrate both sides:

$$-\tilde{f}p = -fp + \frac{p}{2}\frac{\partial}{\partial x} g^2 + \frac{g^2 p}{2}\frac{\partial}{\partial x}\log p$$

Dividing both sides by $-p$, we obtain:

$$\tilde{f} = f - \frac{1}{2}\frac{\partial}{\partial x} g^2 - \frac{g^2}{2}\frac{\partial}{\partial x}\log p$$

Therefore, the PF-ODE is as follows:

$$dx_t = \left[f(x_t, t) - \frac{1}{2}\frac{\partial}{\partial x} g^2(x_t, t) - \frac{g^2(x_t, t)}{2}\frac{\partial}{\partial x}\log p(x_t, t) \right] dt$$

Q2

1. A future AI capability

I think fully automated chip physical behavior simulation and design optimization will come true, and the primary challenge can be partitioned into the following:

- **Scientific / Modeling Challenges:** Multi-scale coupling complexity, difficult embedding of physical constraints.
- **Technical / Computational Challenges:** High-fidelity simulation cost, hardware limitations.

The most important challenge is embedding physical constraints when modeling the behavior of the chip. If we can overcome this, we could simulate the chip and predict which regions might overheat. Since excessively high local temperatures cause thermal throttling, we could adjust the layout or power distribution to ensure that the chip operates at high performance without increasing clock time.

With this improvement, hardware limitations could be alleviated, as the main constraints in IC design are becoming more pronounced due to the slowing of Moore' s law. As a result, we could obtain more efficient processors, reducing the time cost of simulations.

2. Involved machine learning types

To implement it:

- **Supervised learning:** Learn the mapping between chip design parameters and local physical behaviors such as temperature, current density, or quantum states.
- **Reinforcement learning:** Automatically optimize chip design strategies, such as component layout or power distribution.
- **Unsupervised learning:** Extract latent patterns in chip structure and physical behaviors.
- **Physics-informed learning:** Embedding PDEs or physical constraints directly into neural networks (Hard-PINN).

must be involved.

3. The first step of Modelization

Consider energy conservation:

$$\sum_i C_i \frac{d\tilde{T}_i}{dt} = \sum_i P_i + \text{heat flux}$$

Notation:

- $\tilde{T}_i$: predicted temperature of node i
- C_i: heat capacity of node i
- $\frac{d\tilde{T}_i}{dt}$: temperature change over time
- P_i: power dissipated by the i-th node
- heat flux: heat flux passing through the boundary

Hard-PINN can be written as:

$$u(x) = g(x) + u_\theta(x) \cdot f(x)$$

If the chip surface or module geometry is mapped onto a sphere (θ, ϕ), the boundary conditions can be represented using spherical harmonics $Y_l^m(\theta, \phi)$:

$$g(\theta, \phi) = \sum_{l=0}^{L} \sum_{m=-l}^{l} c_l^m Y_l^m(\theta, \phi)$$

where

$$c_l^m = \int_{S^2} T_b(\theta, \phi) \overline{Y_l^m(\theta, \phi)} d\Omega$$

Adaptive loss balancing:

$$\mathcal{L}_{\text{PDE}} = \sum_{i=1}^{N} w_i |r_i|^2$$

Conservation-aware term:

$$\mathcal{L}_{\text{cons}} = \left(\oint_{\partial\Omega} q_\theta \cdot dS - Q_{\text{total}} \right)^2$$

Total loss function:

$$\mathcal{L}_\theta = \sum_{i=1}^{N} w_i |r_i|^2 + \lambda \left(\oint_{\partial\Omega} q_\theta \cdot dS - Q_{\text{total}} \right)^2$$

Unanswered Questions

Is there an efficient way to decide the weight w_i of image interpolation with PF-ODE, where $\sum_i w_i = 1$? Since we do not have prior knowledge of the distribution of the initial state x_0^i, deciding the weights to preserve the characteristics of the image is costly.

References

- https://ncatlab.org/nlab/show/stereographic+projection
- https://arxiv.org/html/2403.00599v1