NSF CyberTraining Winter Workshop

Introduction to Proper Orthogonal Decomposition

concepts, formulation and applications

Part II: POD Model Construction & Examples

Ming-Cheng Cheng

Department of Electrical & Computer Engineering
Clarkson University, Potsdam, NY 13699

Supported by NSF OAC-2118079

POD Simulation Methodology involves 2 projections

- **Mode Training:** maximizing mean square **data projection** onto each of the POD modes:

$$\left\langle \left(\int_{\Omega} Q(\vec{r},t)\eta(\vec{r})d\Omega \right)^2 \right\rangle \Big/ \int_{\Omega} \eta(\vec{r})^2 d\Omega : \ \left\langle \left(\text{Projection onto } \eta(\vec{r}) \right)^2 \right\rangle$$

$$➔ \ \int_{\Omega'} \langle Q(\vec{r},t) \otimes \ Q(\vec{r}\,',t) \rangle \eta(\,\vec{r}\,')d\Omega' = \lambda\eta(\vec{r})$$

This process ensures a minimum least square error with a smallest number of modes **if the training is properly done (i.e., if the data quality is sufficient)**.

- **"Galerkin" projection of the heat conduction equation** onto the the *i*th POD mode, $\eta_i(\vec{r})$

$$\int_{\Omega} \left(\frac{\partial \rho C T(\vec{r},t)}{\partial t} - \nabla \cdot k \nabla T(\vec{r},t) = P_d(\vec{r},t) \right) \eta_i(\vec{r}) d\Omega$$

$$➔ \ \sum_{j=1}^{M} c_{i,j} \frac{da_j}{dt} + \sum_{j=1}^{M} g_{i,j} a_j = P_{pod,i} \ ➔ \ Q(\vec{r},t) = \sum_{j=1}^{M} a_j(t)\,\eta_j(\vec{r})$$

The Galerkin projection closes the model and implement physical principles into the model.

Example, heat conduction in semiconductor Chips

- **Galerkin projection** (transformation) onto the the *i*th POD mode, $\eta_i(\vec{r})$

$$\int_\Omega \left(\eta_i \frac{\partial \rho C T}{\partial t} + \nabla \eta_i \cdot k \nabla T \right) d\Omega = \int_\Omega \eta_i P_d \, d\Omega + \int_S \eta_i k \nabla T \cdot d\vec{S}$$

- Using $T(\vec{r}, t) = \sum_{j=1}^{M} a_j(t) \eta_j(\vec{r})$ ➔ *M* mode POD model, a set of *M*-dimensional ODEs

$$\sum_{j=1}^{M} c_{i,j} \frac{d}{dt} + \sum_{j=1}^{M} g_{i,j} a_j = P_{pod,i},$$

$$c_{i,j} = \int_\Omega \rho C \eta_i \eta_j \, d\Omega, \quad g_{i,j} = \int_\Omega k \nabla \eta_i \cdot \nabla \eta_j \, d\Omega, \qquad P_{pod,i} = \int_\Omega \eta_i P_d(\vec{x}, t) \, d\Omega - \int_\Gamma \eta_i (-k \nabla T) \cdot d\vec{S}$$

These POD model parameters are thus pre-tabulated for solving the ODEs for $\vec{a} = \left[a_1, a_2, \ \ldots a_j, \ldots a_M \right]^T$

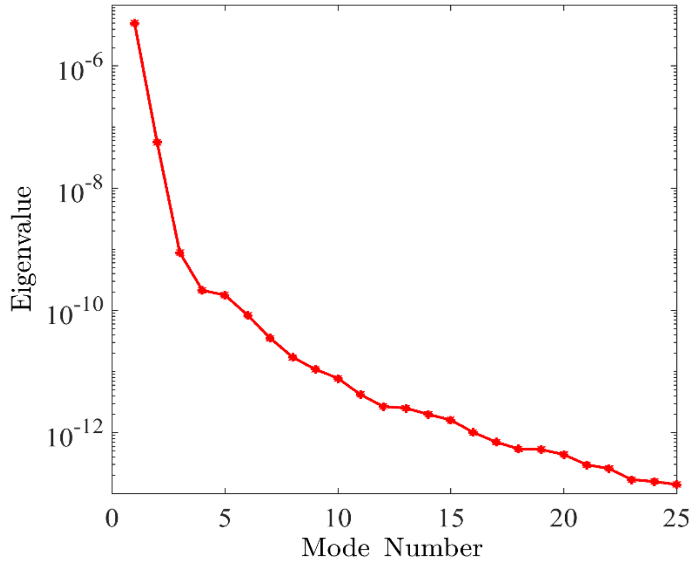
Procedure for Constructing the POD Simulation Model

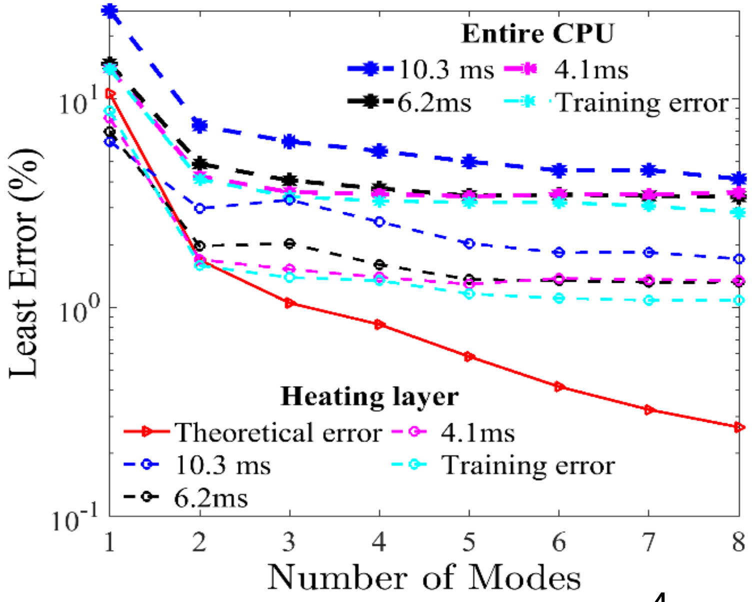
1. Data Collection from direct numerical simulation (DNS) ➔ $Q(\vec{r}, t)$
2. Solving the 2-point correlation eigenvalue problem for λ_j & $\eta_j(\vec{r})$.

   Observe the Eigenvalue spectrum to determine the number of modes, *M,* where λ_j represents the mean squared information captured by η_j Or estimate the least square error based on

$$Err_{LS,M} = \sqrt{\sum_{i=M+1}^{N_s} \lambda_i \Big/ \sum_{i=1}^{N_s} \lambda_i}$$

3. Project the governing equation onto the POD Space (accounting for physical principles) ➔ a set of *M* ODEs for a_j
4. Evaluate the model parameters (coefficients of the ODEs



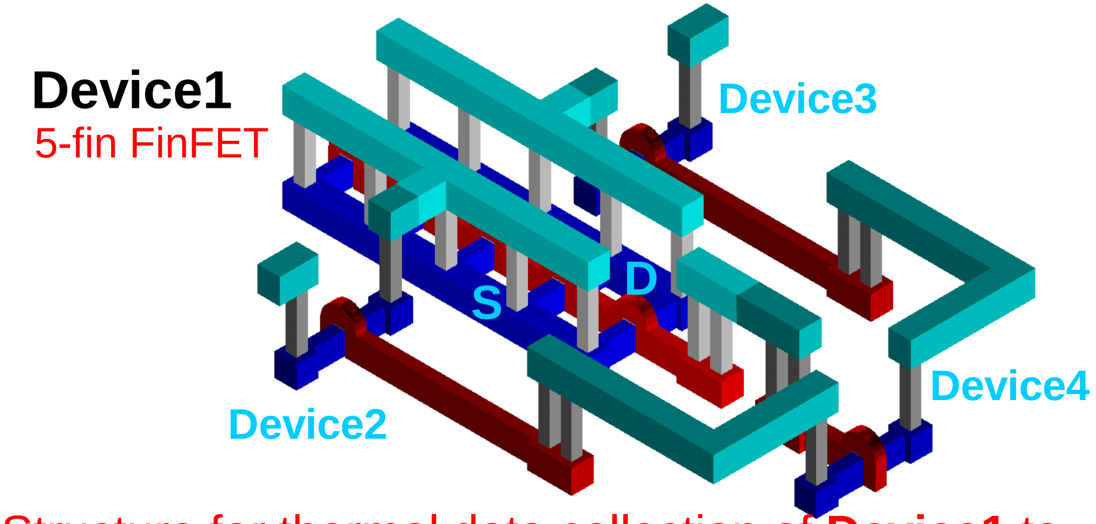


Implementation of POD in physics simulation

- Solve the ODEs to obtain a_j
- Post processing: the solution $Q(\vec{r}, t) = \sum_{j=1}^{M} a_j(t)\, \eta_j(\vec{r})$

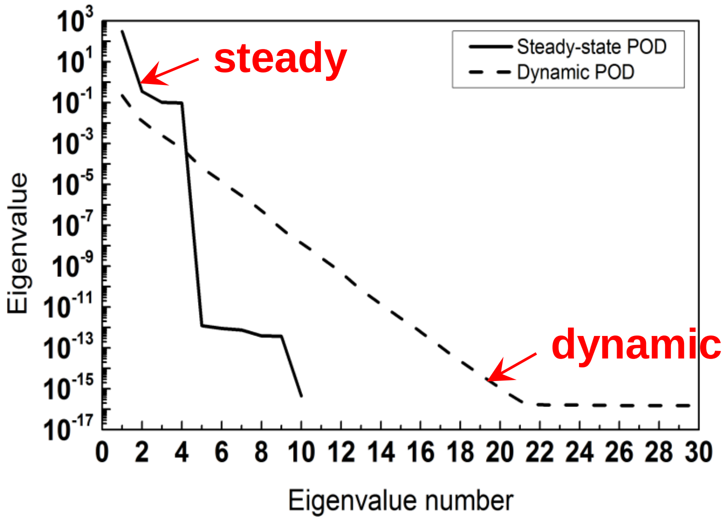
Application to FinFETs: steady-state

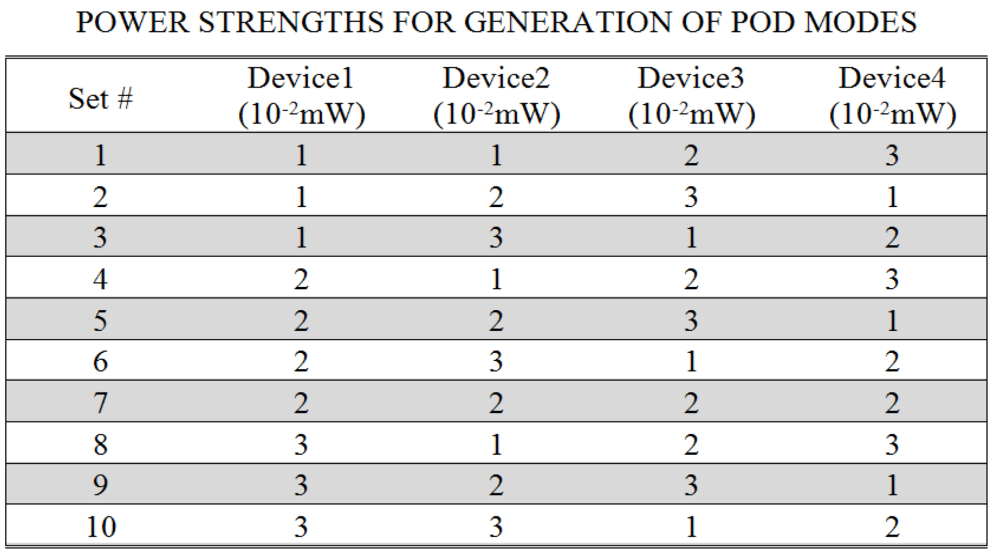
- **Thermal data collection** for extraction of POD modes and eigenvalues:



Structure for thermal data collection of **Device1** to account for the influences from the neighboring devices

Training settings for POD modes to account for realistic BC's **for Device1**



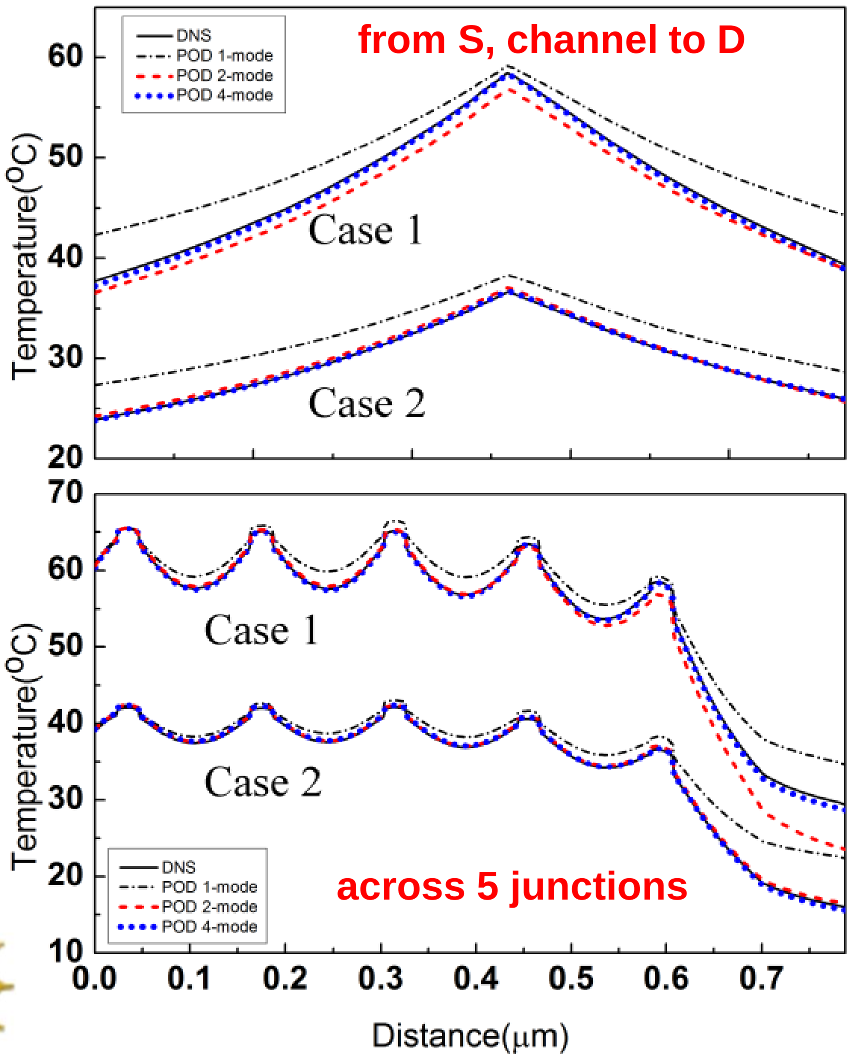


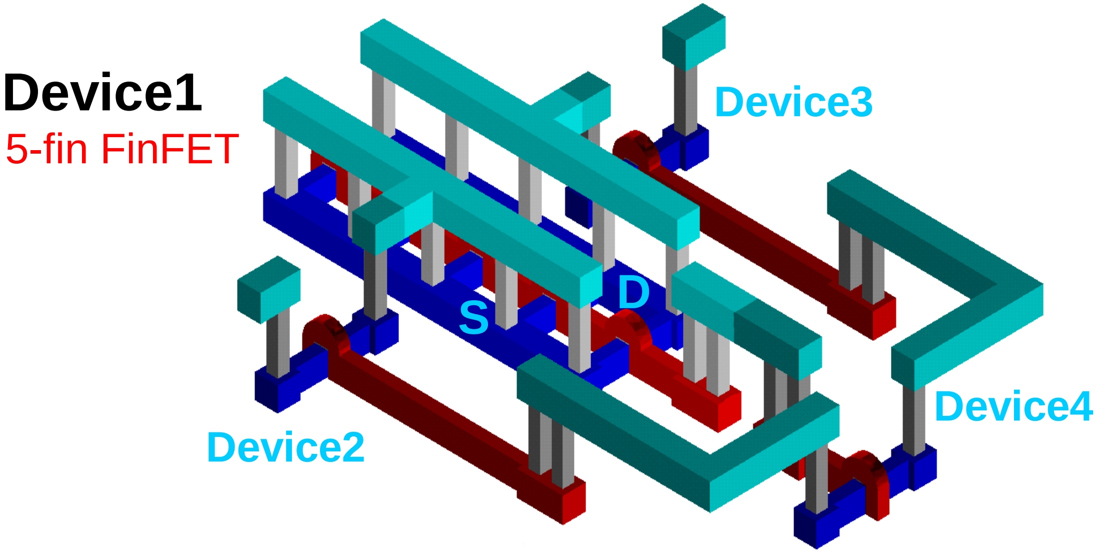
POWER STRENGTHS FOR GENERATION OF POD MODES

Set #	Device1 (10^{-2}mW)	Device2 (10^{-2}mW)	Device3 (10^{-2}mW)	Device4 (10^{-2}mW)
1	1	1	2	3
2	1	2	3	1
3	1	3	1	2
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	2	2	2	2
8	3	1	2	3
9	3	2	3	1
10	3	3	1	2

Application to FinFETs: steady-state

Demonstration





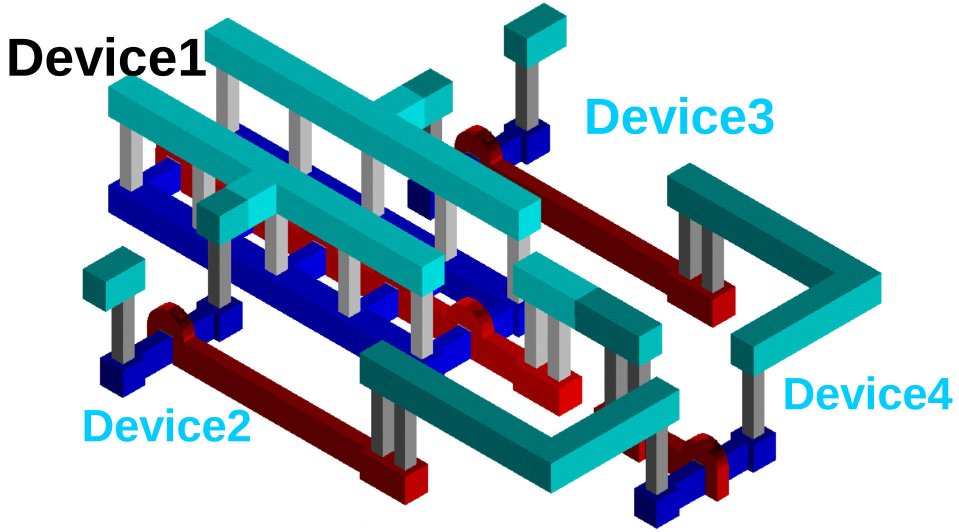
Power Strengths in Demonstration

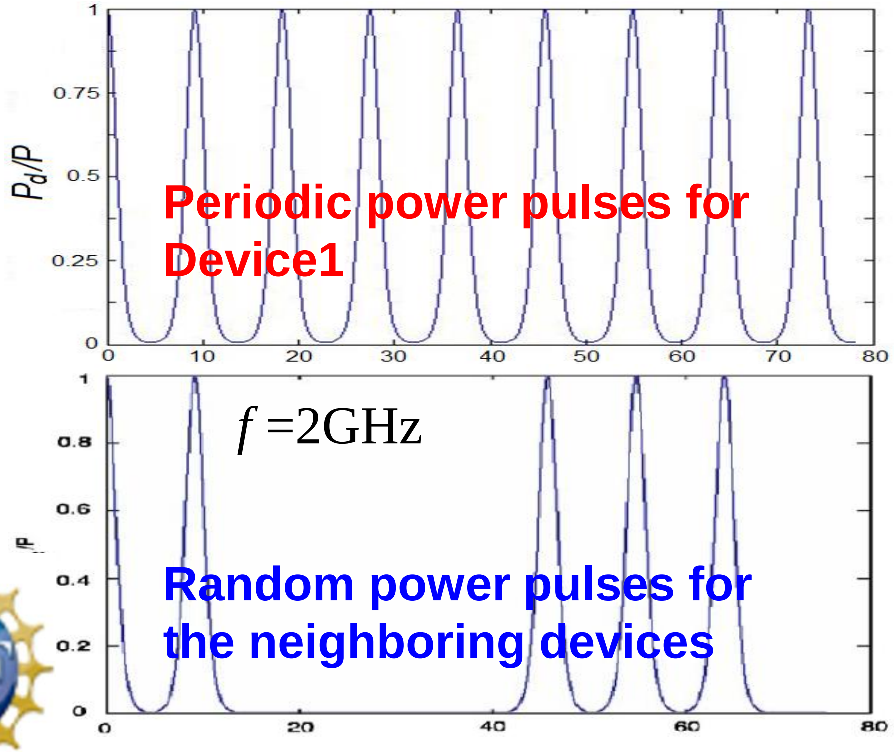
Cases	Device1 (10^{-2}mW)	Device2 (10^{-2}mW)	Device3 (10^{-2}mW)	Device4 (10^{-2}mW)
Case 1	2.5	0.8	1.2	1.6
Case 2	1.5	1.2	0.8	0.4

Application to FinFETs: **dynamic**

Data collection for extraction of POD modes and eigenvalues:

- A **periodic train** of power pulses applied to junctions of Device1
- Synchronized-**random** power pulses applied the neighboring devices

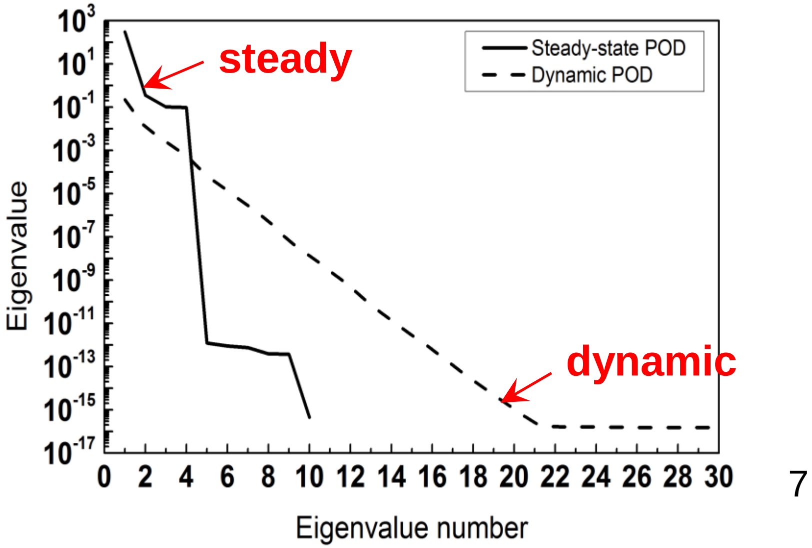




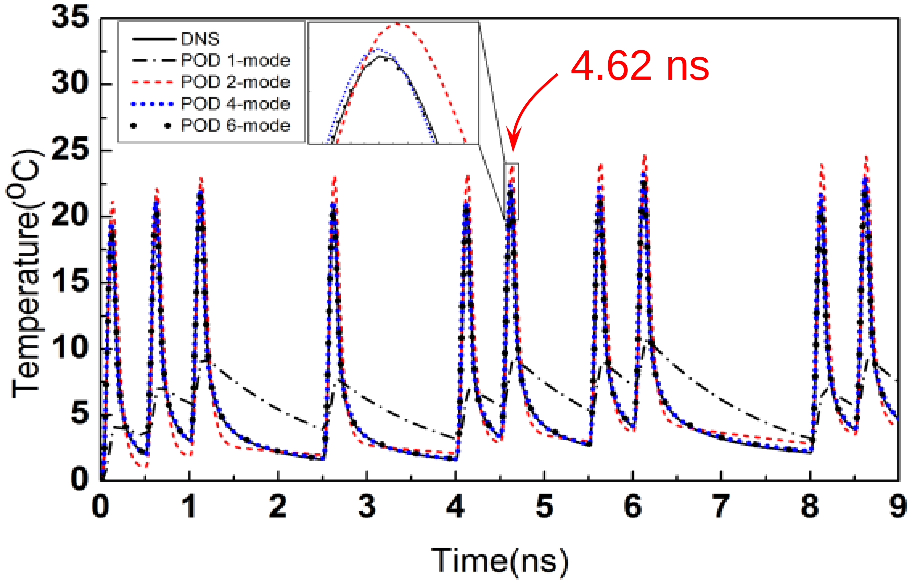
Periodic power pulses for Device1

f =2GHz

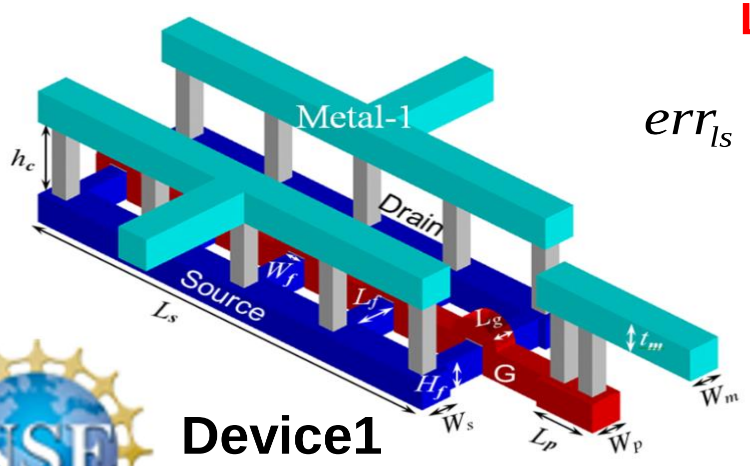
Random power pulses for the neighboring devices



Application to FinFETs: dynamic



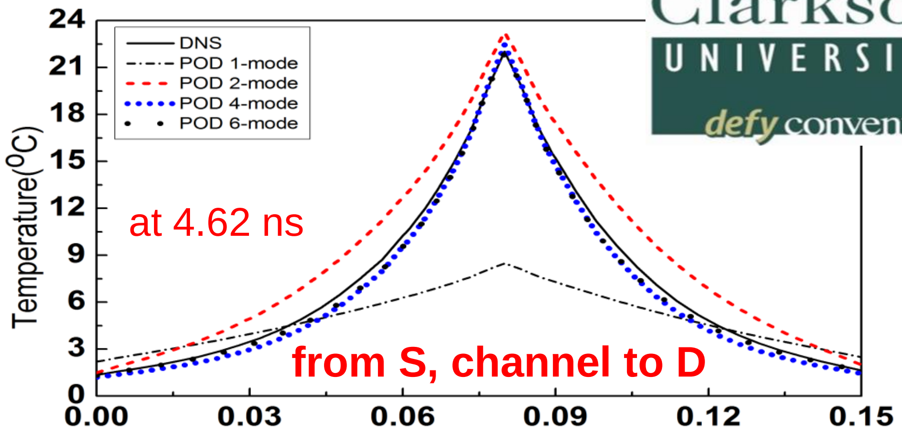
Random power pulses applied to all device junctions

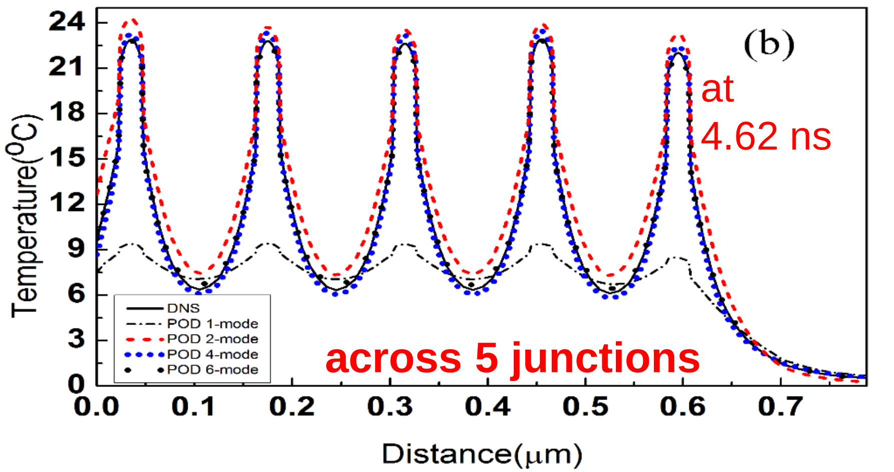


Device1

Least square error

$$err_{ls} = \sqrt{\int_{\Omega} \sum_{i=1}^{N_s} e_i^2 d\Omega \Big/ \left(N_s \Omega\right)}$$

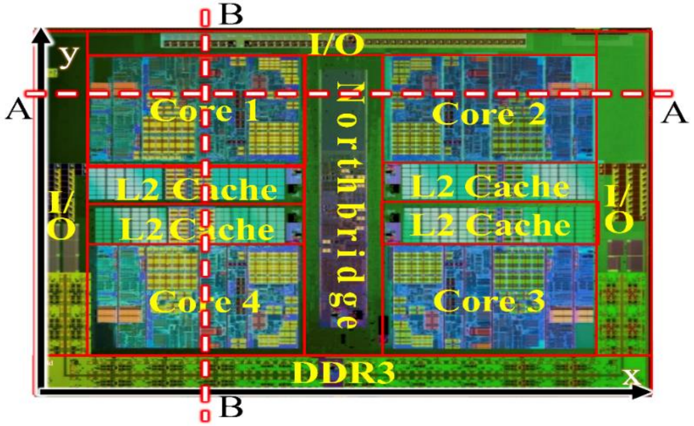




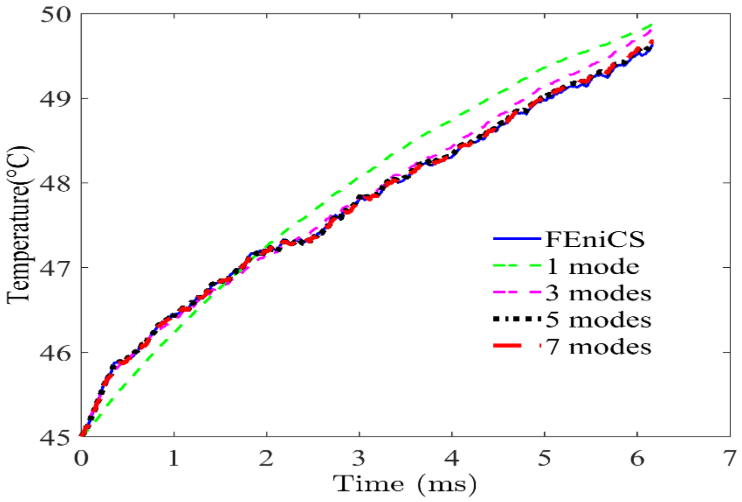
No. of POD modes	4 modes	6 modes
err_{ls}	0.063 °C	0.042 °C

A reduction in DoF by 5 orders of magnitude is achieved

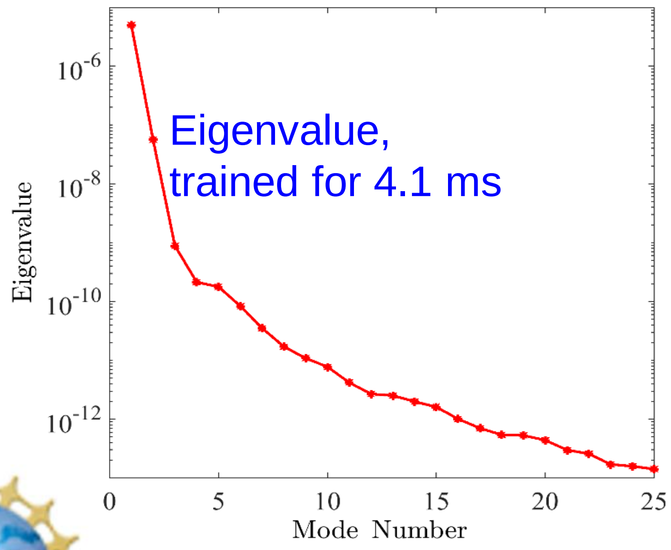
Application to a CPU, AMD ATHLON II X4 610e



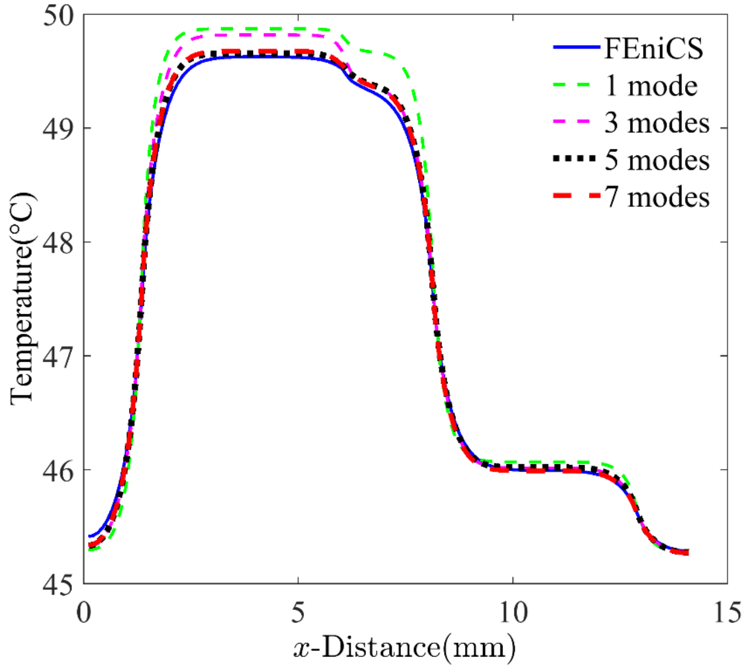
Dynamic power maps

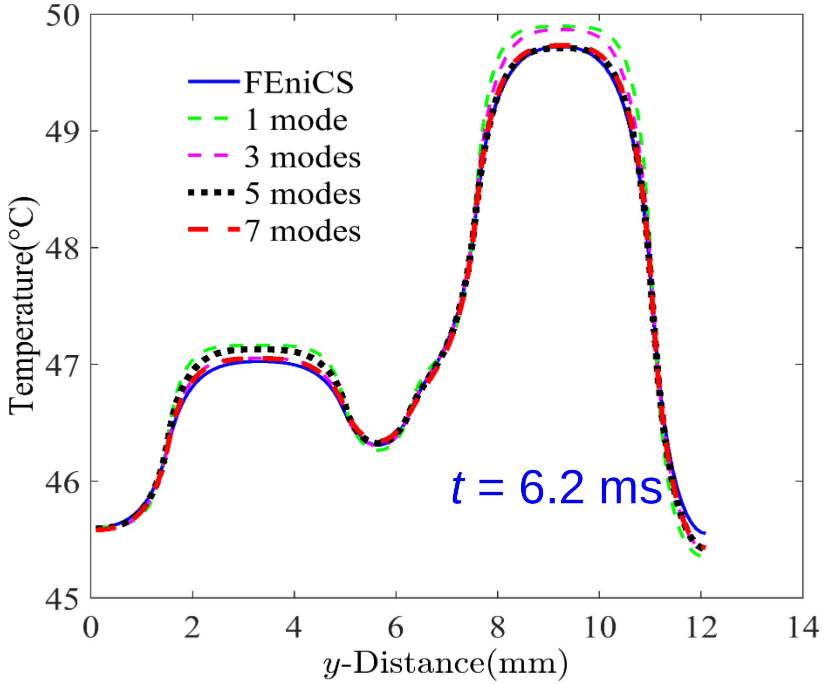


Very accurate prediction of the dynamic temperature distribution even beyond the training time



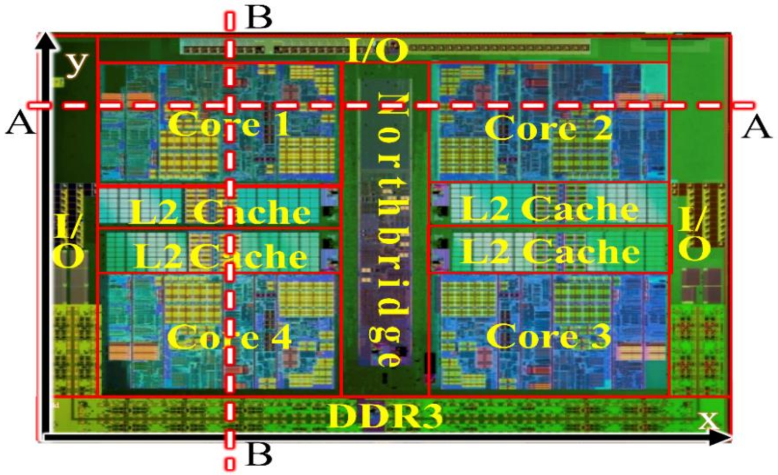
Eigenvalue, trained for 4.1 ms

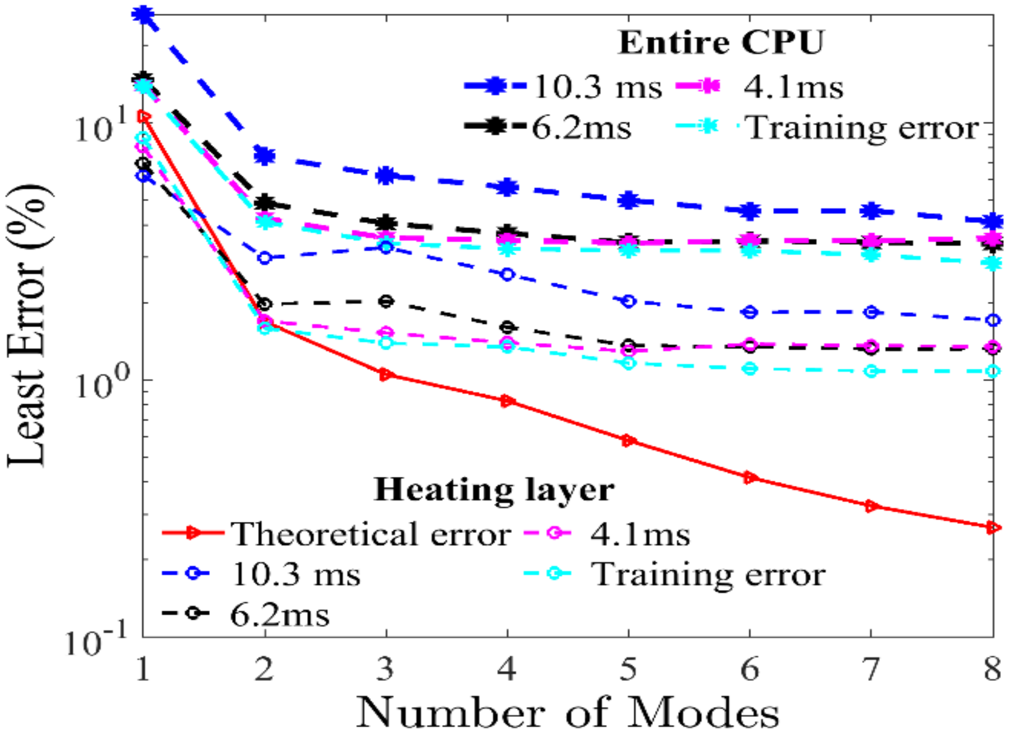




t = 6.2 ms

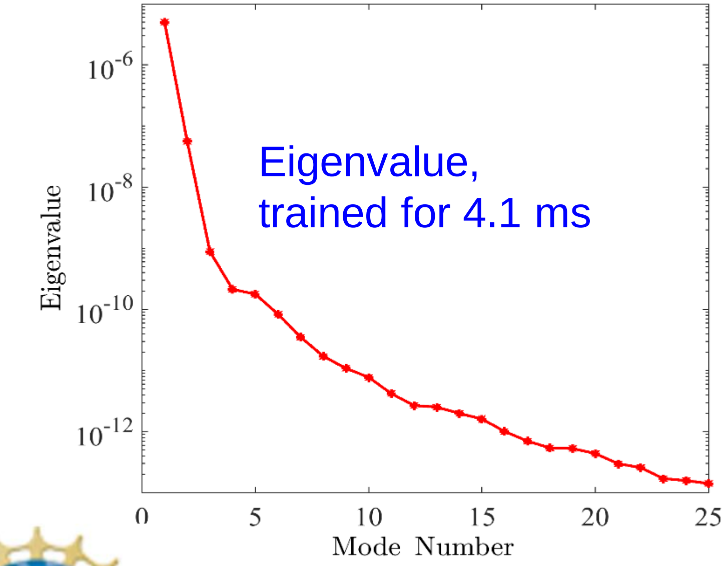
Application to a CPU, AMD ATHLON II X4 610e : Dynamic





Accurate prediction of the dynamic temperature distribution even beyond the training time

Post1: heting layer
Post2: entire CPU



Eigenvalue, trained for 4.1 ms

Simulation time (ms)	Time consumption (s)										
	FEniCS (FEM)		Number of POD Modes								
			1			3			5		
	Gmres	Mumps	ODE	Post1	Post2	ODE	Post1	Post2	ODE	Post1	Post2
4.1	6.78e3	1.35e5	0.11	0.07	1.10	0.11	0.15	2.08	0.11	0.27	3.16
6.2	1.02e4	2.00e5	0.15	0.11	1.28	0.16	0.23	2.68	0.16	0.40	4.78
10.3	1.69e4	3.33e5	0.26	0.18	2.14	0.28	0.38	4.51	0.30	0.69	8.05

Entire chip: a reduction in computational time over 3,500 times
Heating layer: a reduction in computational time near 26,000 times

More information on the POD Thermal Simulation Methodology

- In most applications relevant to thermal issues, thermal information only needed in high temperature region
  ➔ Post processing only need to perform at certain grid points
  ➔ Computational time is at least order shorter. $Q(\vec{r},t) = \sum_{j=1}^{M} a_j(t)\,\eta_j(\vec{r})$

- For very large domain structures, such as GPUs with hundreds or thousands of cores, the approaches can be modified to improve the training efficiency.
  - Multi-block POD: POD Blocks + Domain decomposition + Discontinuous Galekin
    - Fast Thermal Simulation of **FinFET Circuits** Based on a Multi-Block Reduced-Order Model, IEEE Trans. CAD ICs & Systems, 2016. DOI: 10.1109/TCAD.2015.2501305
    - A methodology for thermal simulation of **interconnects** enabled by model reduction with material property variation, J. Computational Sci..2022. doi.org/10.1016/j.jocs.2022.101665
    - Chip-level Thermal Simulation for a **Multicore Processor** Using a Multi-Block Model Enabled by Proper Orthogonal Decomposition, ITherm 2022. Doi: 10.1109/iTherm54085.2022.9899503
  - Ensemble POD: Individual POD + Domain Truncation + Superposition
    - Predicting Accurate Hot Spots in a **More Than Ten-Thousand-Core GPU** with a Million Time Speedup over FEM Enabled by a Physics-based Learning Algorithm”, ITherm 2024, May 28-May 31, 2024.

Example: Quantum Eigenvalue Problems for Nanostructure

Basic Concepts of Electron Wave functions (WFs) in nanostructure

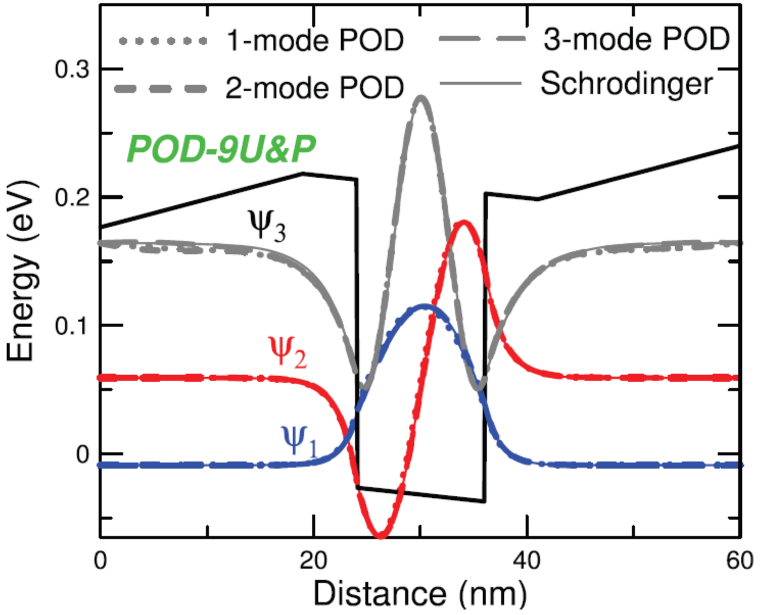
Schrödinger equation: $\nabla \cdot \left[-\frac{\hbar^2}{2m^*}\nabla + U(\vec{r}) \right] \psi(\vec{r}) = E\ \psi(\vec{r})$ Or $H\psi = E\psi$

Kinetic energy potential energy total energy

where H is the Hamiltonian operator and E is the total energy for an electron

Quantum eigenvalue problem ➔

- **Eigenvalues:** discreate energies of an electron in a small-scale confinement
- The larger the spatial confinement is, the closer the discrete energies are ➔ continuous energy ➔ classical
- **Eigenfunctions:** electron WFs ψ_i in different energy states (or eigenstates).
- $|\psi_i(\vec{r})|^2$ represents the probability density of the electron in the *i*-th quantum state at $\vec{r}$.



Example: Quantum Eigenvalue Problems for Nanostructure

Schrödinger Equation: $$\nabla\cdot\left[-\frac{\hbar^2}{2m^*}\nabla\psi(\vec{r})\right]+U(\vec{r})\psi(\vec{r})=E\psi(\vec{r})$$

- Galerkin projection of the Schrödinger equation onto the *i*th POD mode, $\eta_i(\vec{r})$

$$\int_\Omega\left(\nabla\cdot\left[-\frac{\hbar^2}{2m^*}\nabla\psi(\vec{r})\right]+U(\vec{r})\psi(\vec{r})=E\psi(\vec{r})\right)\underline{\eta_i(\vec{r})}d\Omega$$

Using the following identities:

$$\nabla\cdot\left(\eta_i\frac{\hbar^2}{2m^*}\nabla\psi\right)=\nabla\eta_i\cdot\frac{\hbar^2}{2m^*}\nabla\psi+\underline{\eta_i}\nabla\cdot\frac{\hbar^2}{2m^*}\nabla\psi$$

$$\text{Gauss's Law:}\ \underbrace{\int_\Omega\nabla\cdot\left(\eta_i\frac{\hbar^2}{2m^*}\nabla\psi\right)d\Omega}_{\text{Vol}}=\overbrace{\int_S\eta_i\frac{\hbar^2}{2m^*}\nabla\psi\cdot d\vec{S}}^{\text{Surf}}$$

- The projection leads to the weak form of the Schrödinger Equation

$$\int_\Omega\nabla\eta_i\cdot\frac{\hbar^2}{2m^*}\nabla\psi d\Omega+\int_\Omega\eta_i U\psi d\Omega-\int_s\eta_i\frac{\hbar^2}{2m^*}\nabla\psi\cdot d\vec{S}=E\int_\Omega\eta_i\psi d\Omega$$

Example: Quantum Eigenvalue Problems for Nanostructure

- Galerkin projection (transformation) onto the the *i*th POD mode, $\eta_i(\vec{r})$

$$\int_\Omega \nabla\eta_i \cdot \frac{\hbar^2}{2m^*}\nabla\underline{\psi} d\Omega + \int_\Omega \eta_i U \underline{\psi} d\Omega - \int_s \eta_i \frac{\hbar^2}{2m^*}\nabla\underline{\psi} \cdot d\vec{S} = E\int_\Omega \eta_i \underline{\psi} d\Omega$$

- Using $\psi(\vec{r}) = \sum_{j=1}^{M} a_j\eta_j(\vec{r})$ ➔ an *M*-dimensional eigenvalue problem in the POD space

$$\boldsymbol{H}_\eta \vec{a} = E\,\vec{a}, \qquad \text{where } \vec{a} = [a_1\ \ a_2\ \ \ldots\ a_M]^T$$

where Hamiltonian in the POD eigenspace: $\boldsymbol{H}_\eta = \boldsymbol{T}_\eta + \boldsymbol{U}_\eta + \boldsymbol{B}_\eta$

$$T_{\eta\ i,j} = \int_\Omega \nabla\eta_i(\vec{r}) \cdot \frac{\hbar^2}{2m^*}\nabla\eta_j(\vec{r})d\Omega, \qquad U_{\eta\ i,j} = \int_\Omega \eta_i(\vec{r})U(\vec{r})\psi(\vec{r})d\Omega\,, \qquad \boldsymbol{B}_{\eta\ i,j} = \int_s \eta_i(\vec{r})\frac{-\hbar^2}{2m^*}\nabla\psi(\vec{r}) \cdot d\vec{S}$$

Interior kinetic energy matrix

Potential energy matrix

Boundary kinetic energy matrix

These POD model parameters can be pre-tabulated for solving the ODEs for

$$\vec{a} = \left[a_1, a_2,\ \ldots a_j, \ldots a_M\right]^T$$

Example: Quantum Eigenvalue Problems for Nanostructure

$$\text{Schrödinger equation: } \nabla \cdot \left[-\frac{\hbar^2}{2m^*} \nabla + U(\vec{r}) \right] \psi(\vec{r}) = E \, \psi(\vec{r}) \quad OR \quad H\psi = E\psi$$

Different approaches to train the POD modes ➔

- **Individual-state POD model:** generates one set pf POD modes for each selected individual state

  N_{QS} selected quantum states ➔ N_{QS} sets of POD modes; ; *i.e.*, N_{QS} quantum POD models

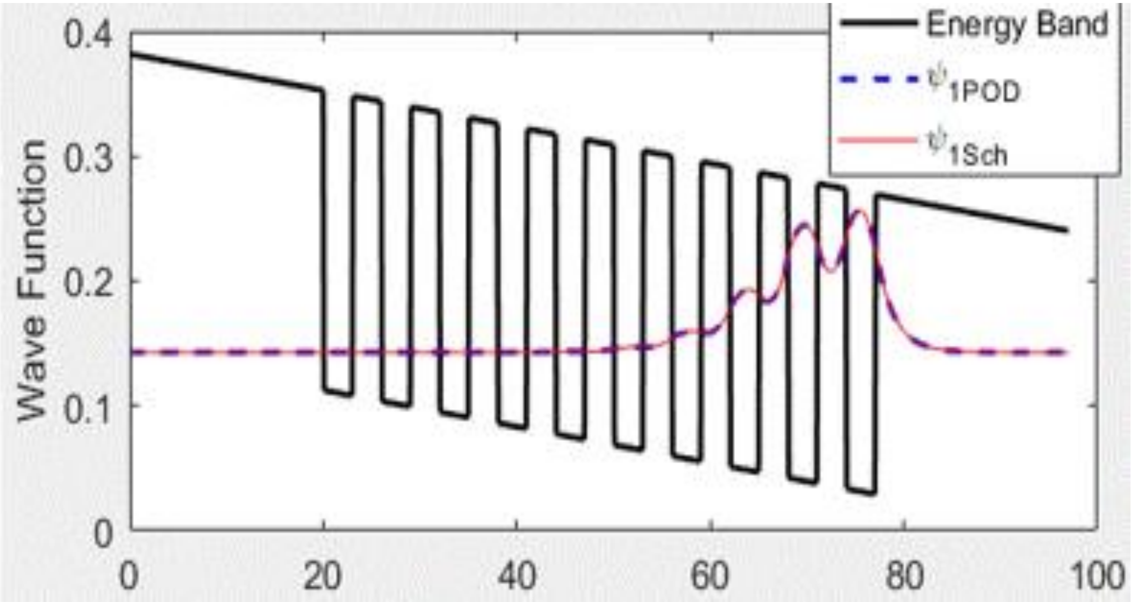
- **Global POD model:** generates only one set of POD modes for all the selected states

  After solving the *i*th state eigenenergy E_i, and eigenvector $\vec{a}_i$ from the Hamiltonian equation in the POD space, $\boldsymbol{H}_\eta \vec{a} = E \, \vec{a}$, the electron WF of the *i*th quantum state is calculated from

$$\psi_i(\vec{r}) = \vec{a}_i^T \cdot \vec{\eta} = \sum_{j=1}^{M} a_{j,i} \, \eta_j(\vec{r})$$

Multi Quantum Well Structure

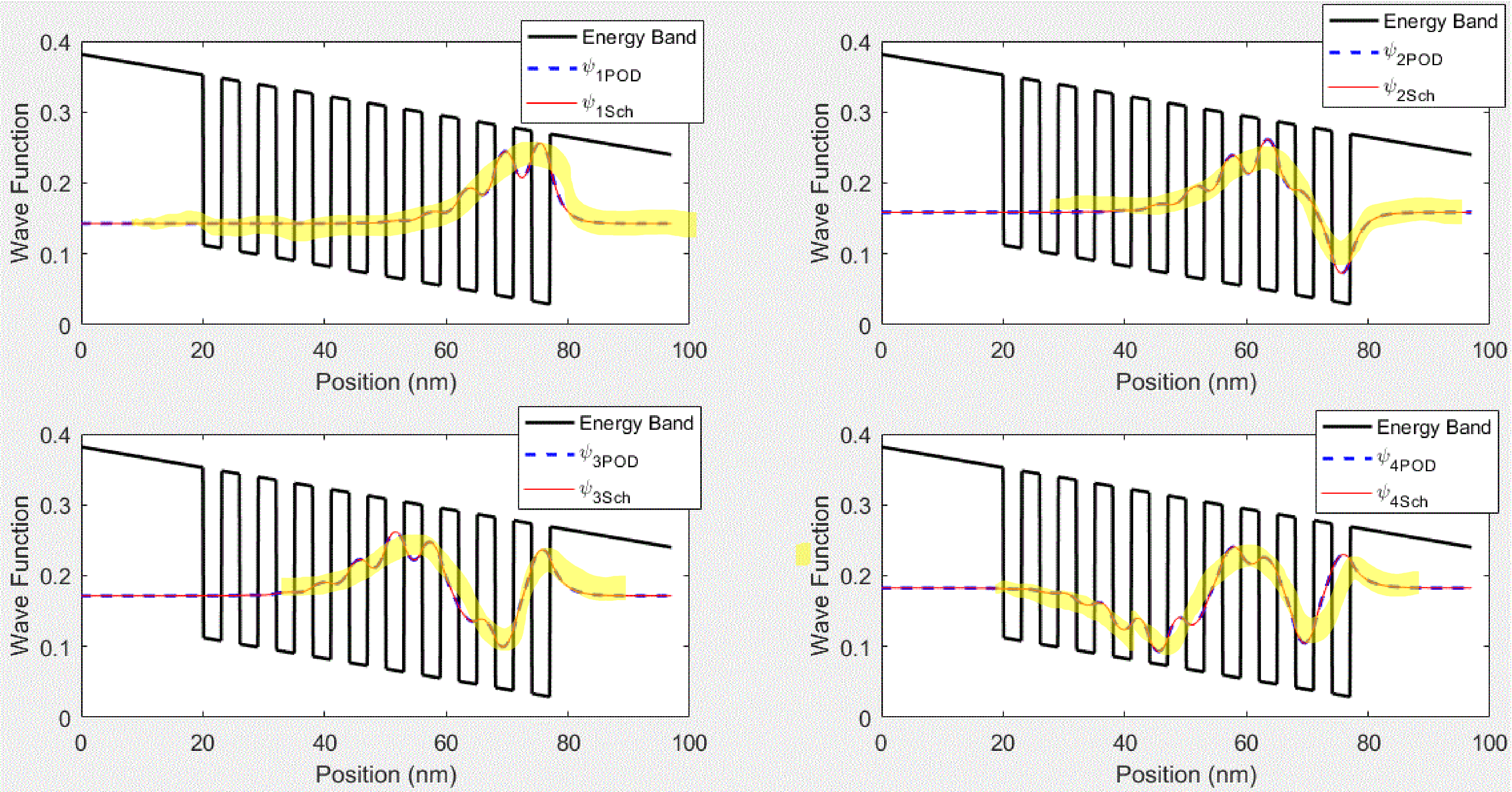
Wave Function (WF) Data Collection with potential (or electric field) variation



Apply various electric fields (indicated by the slope of the conduction band) to collect N_s sets of WF data, where $N_s = N_F \times N_{QS}$

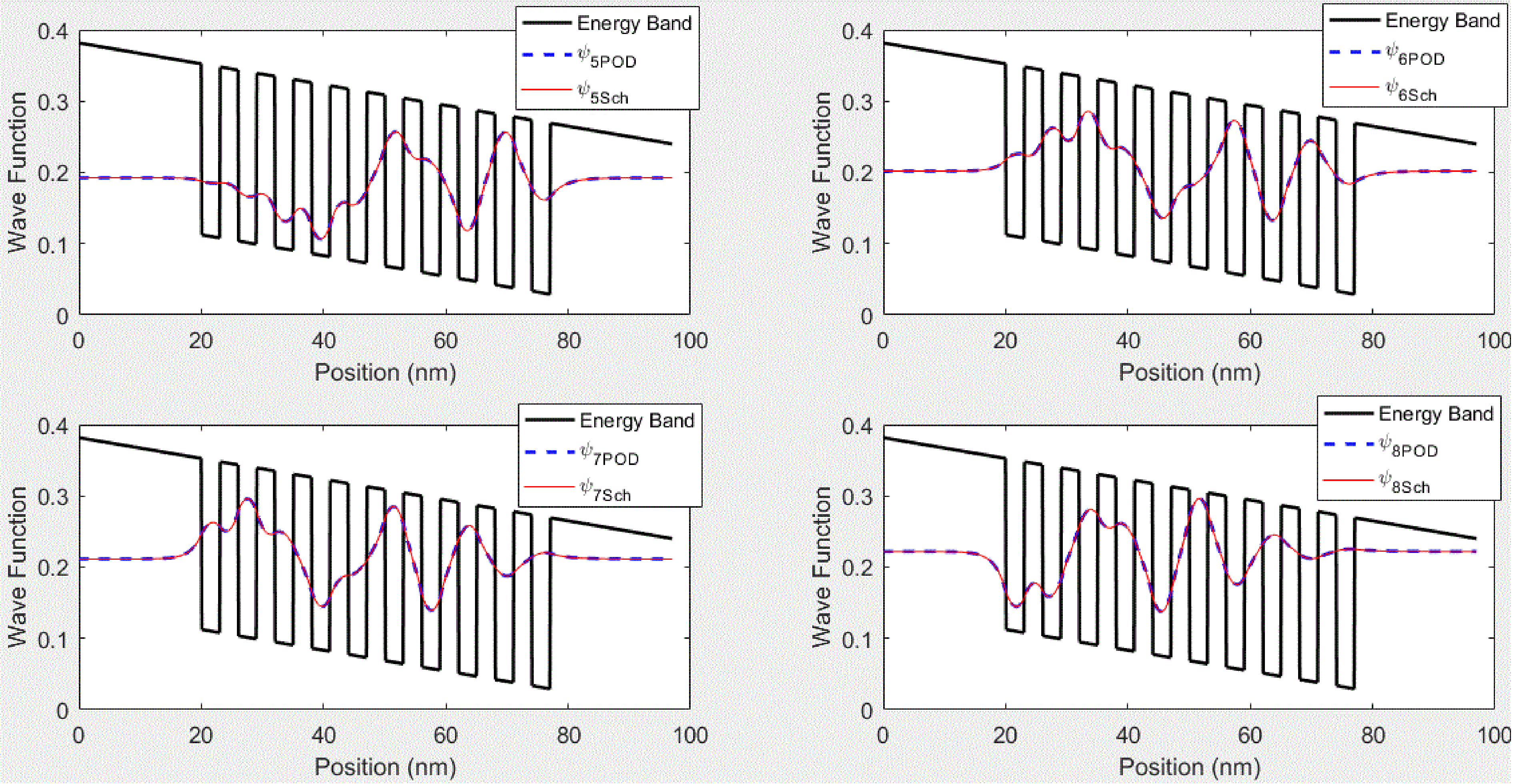
N_F is the number of applied electric fields

Multi Quantum Wells using 10 modes in POD model



Quantum State	1	2	3	4
LS error	0.015%	0.028%	0.037%	0.042%

Multi Quantum Wells using 10 modes in POD model



Quantum State	5	6	7	8
LS error	0.042%	0.041%	0.064%	0.018%