Deep Learning model for real-time SPARC vertical stability observers

A.Kumar[1], C. Clauser[1], T. Golfinopoulos[1], F. Carpenese[2], O.Nelson[3], D.T. Garnier[1], D. Boyer[3], D. Battaglia[4], A. Saperstein[1], C. Rea[1]

1. Plasma Science and Fusion Center | MIT, USA
2. Neural Concept, Switzerland
3. Columbia University, USA
4. Commonwealth Fusion System, USA

3rd Technical Meeting on Plasma Disruptions and their Mitigation, 2024

ITER Headquarters

Contents

❑ Motivation behind this work

- **Real-Time Compatibility** of Vertical Stability Observer (VSO) metrics for SPARC Off-Normal Warning systems.
- **Efficiency:** Ensure fast and reliable analysis for SPARC Plasma Control, leveraging computationally inexpensive surrogates over traditional physics-based models.
- **Optimize ARC Power Plant Design**: Vertical stability is key - fast, accurate models are essential for system-wide optimization.

❑ Scope of this work

- Brief review of Physics–based Vertical Stability models (basically linear and Non-linear PDEs).
- **Surrogate Model Development:** To generate real-time VSO metrics are:
  - Non-linear VDE growth rate (γ_Z)
  - Maximum Controllable Displacement (ΔZ_{max})

Current Physics-Based Models for Vertical Stability Observers

❑ VDE growth rates: Comparing Linear vs. Non-Linear Approaches

Rigid Body Model (Linear)

- Plasma treated as a rigid body (RZIp).
- Plasma circuit equation is formulated assuming a fixed shape of the current distribution.
- Motion is constrained to purely vertical.

Non-Rigid Body Model (Non-Linear)

- Plasma treated as non-rigid (deformable) problem.
- Dynamics of conductor current evolution.
- Coupled to the resistive plasma current decay on subsequent states of free-boundary equilibria.
- Flexible to parametrized state-space.

More details Francesco PhD thesis , 2021

Current Physics-Based Models for Vertical Stability Observers

SPARC

❑ VDE growth rates: Comparing Linear vs. Non-Linear Approaches

Rigid Body Model (Linear)

- Plasma treated as a rigid body (RZIp).
- Plasma circuit equation is formulated assuming a fixed shape of the current distribution.
- Motion is constrained to purely vertical.

$$M_{\mathrm{ss}}\dot{I}_{\mathrm{s}} + R_{\mathrm{ss}}I_{\mathrm{s}} + \frac{\partial \psi_{\mathrm{sp}}}{\partial z}\dot{z} = V_{\mathrm{s}} \qquad 0 = \frac{\partial F_z}{\partial z}\dot{z} + \frac{\partial F_z}{\partial I_s}\dot{I}_s$$

$$M_{\mathrm{ss}}\dot{I}_{\mathrm{s}} + R_{\mathrm{ss}}I_{\mathrm{s}} + \frac{\partial \psi_{\mathrm{sp}}}{\partial z}\frac{\partial z}{\partial I_{\mathrm{s}}}\dot{I}_{\mathrm{s}} = \left(M_{\mathrm{ss}} + \frac{\partial \psi_{\mathrm{sp}}}{\partial z}\frac{\partial z}{\partial I_{\mathrm{s}}}\right)\dot{I}_{\mathrm{s}} + R_{\mathrm{ss}}I_{\mathrm{s}}$$

$$= L_{*\mathrm{s}}\dot{I}_{\mathrm{s}} + R_{\mathrm{ss}}I_{\mathrm{s}} = V_{\mathrm{s}},$$

Non-Rigid Body Model (Non-Linear)

- Plasma treated as non-rigid (deformable) problem.
- Dynamics of conductor current evolution.
- Coupled to the resistive plasma current decay on subsequent states of free-boundary equilibria.
- Flexible to parametrized state-space.

$$M_{ss}\dot{I}_s + R_{ss}I_s + X_{ff}\dot{I}_s = V_s \qquad \Delta^*\psi = -\mu_0 R^2 p' - FF'$$

$$X_{ff} \equiv \frac{\partial\, \psi_{(p)}}{\partial I_s}$$

- Change in plasma current and a profile shape parameters to the circuit equation

$$\frac{\partial\, \psi_{(p)}}{\partial \boldsymbol{x}}\dot{\boldsymbol{x}} \equiv Y\dot{I}_s$$

More details Francesco PhD thesis , 2021

- Solved by MATLAB-based MEQ code suite.

Current Physics-Based Models for Vertical Stability Observers

SPARC

❑ VDE growth rates: Comparing Linear vs. Non-Linear Approaches

Rigid Body Model (Linear)

- Plasma treated as a rigid body (RZIp).
- Plasma circuit equation is formulated assuming a fixed shape of the current distribution.
- Motion is constrained to purely vertical.

$$M_{ss}\dot{I}_s + R_{ss}I_s + \frac{\partial \psi_{sp}}{\partial z}\dot{z} = V_s \qquad 0 = \frac{\partial F_z}{\partial z}\dot{z} + \frac{\partial F_z}{\partial I_s}\dot{I}_s$$

$$M_{ss}\dot{I}_s + R_{ss}I_s + \frac{\partial \psi_{sp}}{\partial z}\frac{\partial z}{\partial I_s}\dot{I}_s = \left(M_{ss} + \frac{\partial \psi_{sp}}{\partial z}\frac{\partial z}{\partial I_s}\right)\dot{I}_s + R_{ss}I_s$$

$$= L_{*s}\dot{I}_s + R_{ss}I_s = V_s,$$

$$A \equiv -L_{*s}R_{ss}$$

- Eigenvalues of this state matrix, A are the VDE growth rate.

Non-Rigid Body Model (Non-Linear)

- Plasma treated as non-rigid (deformable) problem.
- Dynamics of conductor current evolution.
- Coupled to the resistive plasma current decay on subsequent states of free-boundary equilibria.
- Flexible to parametrized state-space.

$$M_{ss}\dot{I}_s + R_{ss}I_s + X_{ff}\dot{I}_s = V_s \qquad \Delta^*\psi = -\mu_0 R^2 p' - FF'$$

$$X_{ff} \equiv \frac{\partial\, \psi_{(p)}}{\partial I_s}$$

- Change in plasma current and a profile shape parameters to the circuit equation

$$\frac{\partial\, \psi_{(p)}}{\partial \boldsymbol{x}}\dot{\boldsymbol{x}} \equiv Y\dot{I}_s$$

More details Francesco PhD thesis , 2021

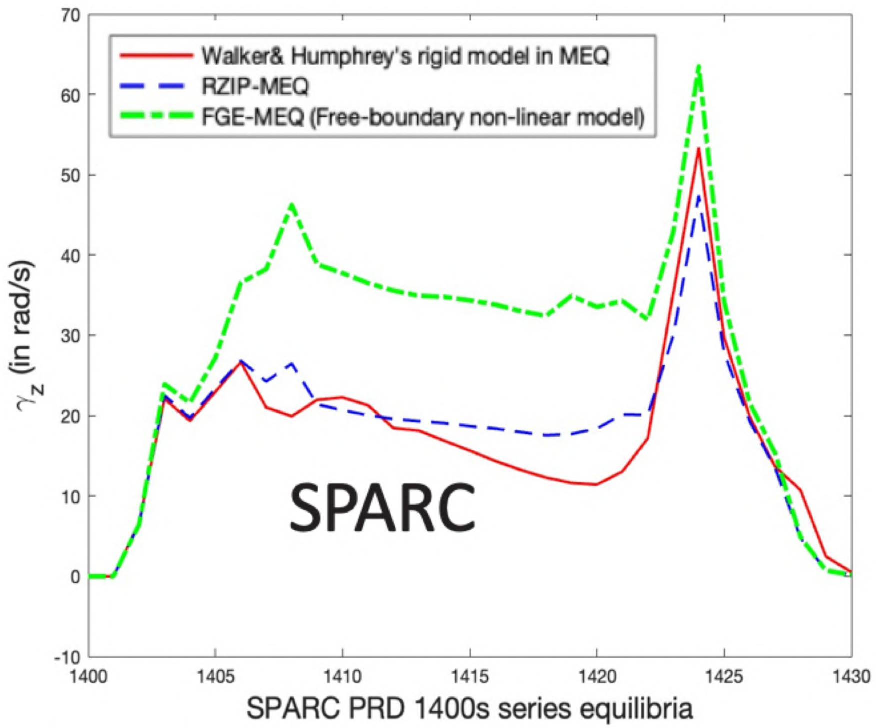
- Solved by MATLAB-based MEQ code suite.

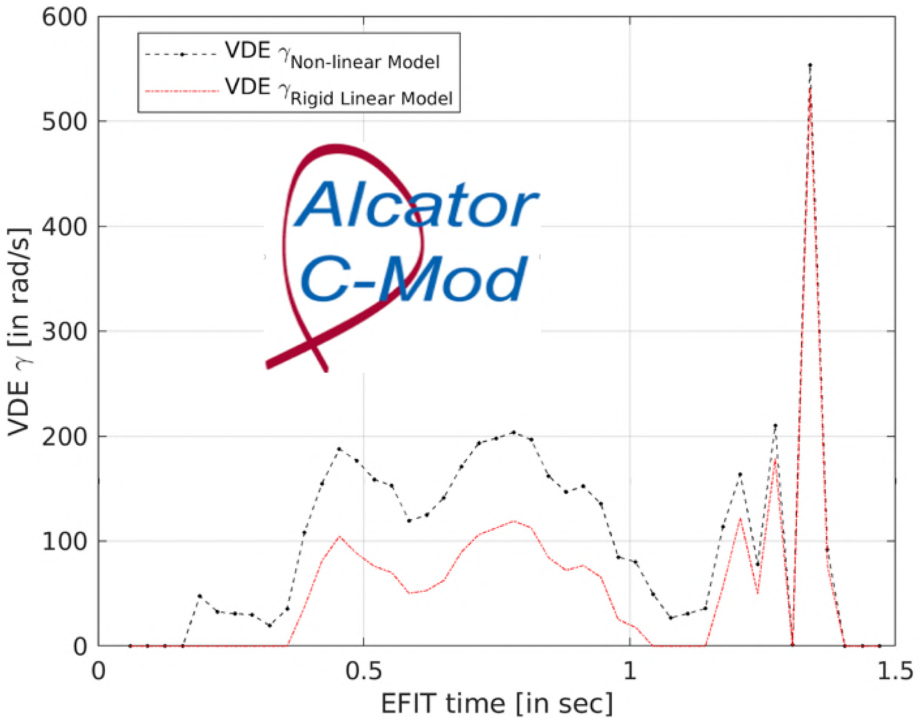
Comparing Rigid vs. Non-Rigid Models for VDE Growth Rates in SPARC and C-Mod

*MEQ-FGEL – physics-based code which solves the non-linear VDE γ_z

Key Findings: Non-linear models predict better VDE growth rates.

Differences observed mainly during the flat-top period.





Comparing Rigid vs. Non-Rigid Models for VDE Growth Rates in SPARC and C-Mod

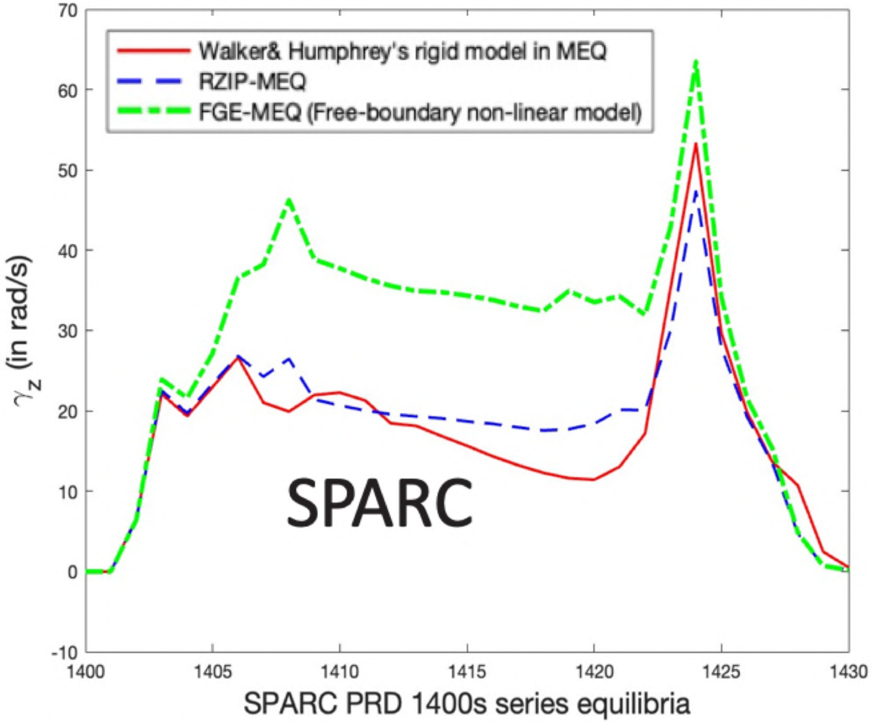
*MEQ-FGEL – physics-based code which solves the non-linear VDE γ_z

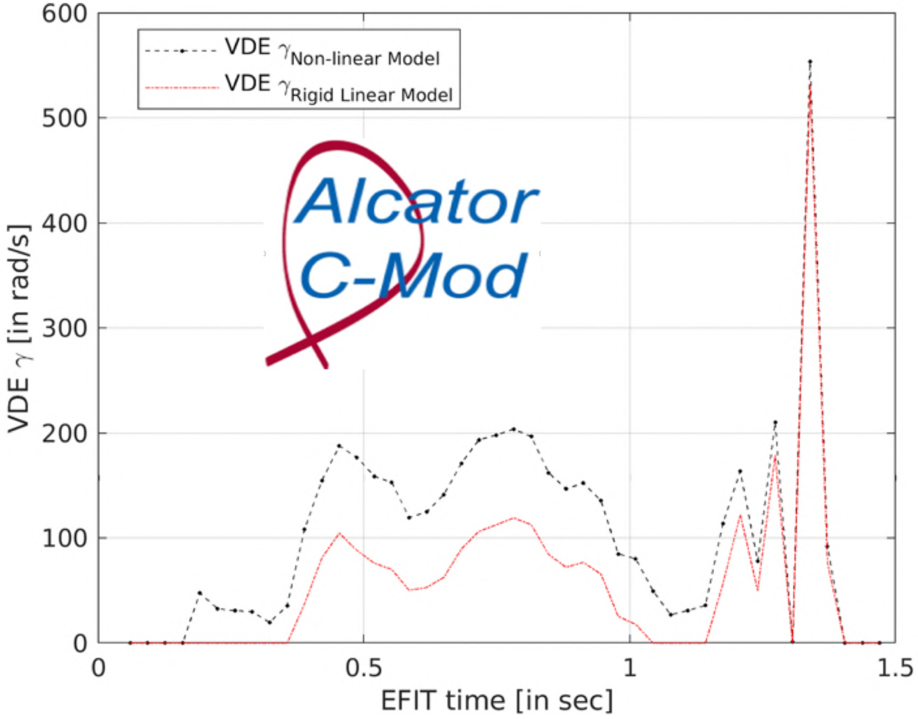
Key Findings: Non-linear models predict better VDE growth rates.

Differences observed mainly during the flat-top period.

Model Comparison: Non-linear models differ by a factor of ~2 from linear models.

Non-rigid models are not suitable for real-time processing.





Comparing Rigid vs. Non-Rigid Models for VDE Growth Rates in SPARC and C-Mod

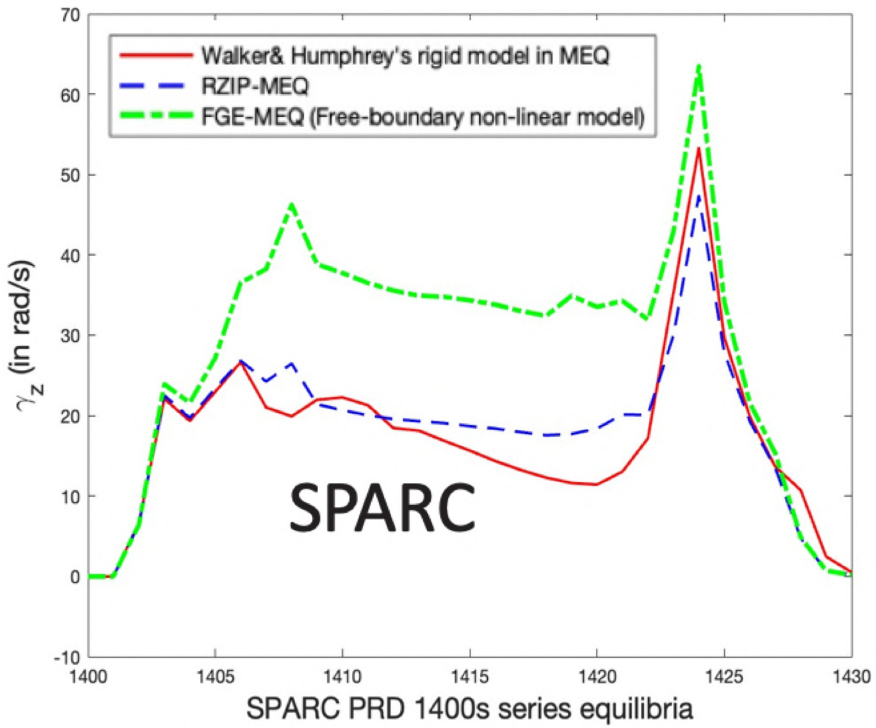
*MEQ-FGEL – physics-based code which solves the non-linear VDE γ_z

Key Findings: Non-linear models predict better VDE growth rates.

Differences observed mainly during the flat-top period.

Model Comparison: Non-linear models differ by a factor of ~2 from linear models.

Non-rigid models are not suitable for real-time processing.



Challenges of solving Non-Linear PDE model made not compatible for real-time processing.

Comparing Rigid vs. Non-Rigid Models for VDE Growth Rates in SPARC and C-Mod

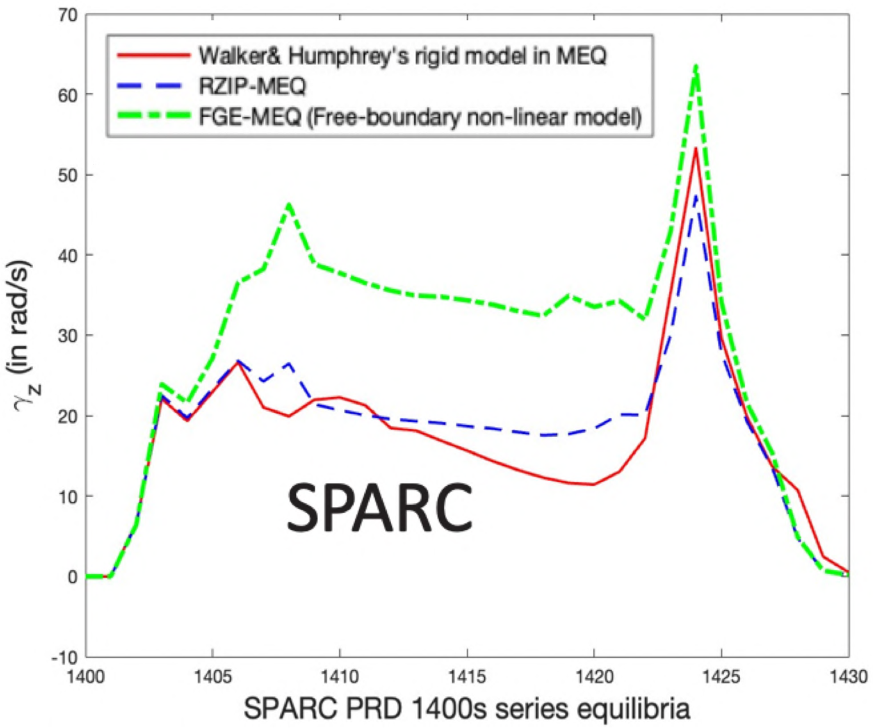
*MEQ-FGEL – physics-based code which solves the non-linear VDE γ_z

Key Findings: Non-linear models predict better VDE growth rates.

Differences observed mainly during the flat-top period.

Model Comparison: Non-linear models differ by a factor of ~2 from linear models.

Non-rigid models are not suitable for real-time processing.



Challenges of solving Non-Linear PDE model made not compatible for real-time processing.

Current Physics-Based Models for Vertical Stability Observers

SPARC

❑ ΔZ_{max} : Maximum controllable displacements

Key Components:

- Simulation Principle:
  - ΔZ_{max} is determined via time-dependent simulations that constrain coil current or voltage saturation.
  - Simulation considers non-linear/linear plasma behavior and evolving equilibrium states, allowing for more accurate predictions under varying conditions.

Current Physics-Based Models for Vertical Stability Observers

SPARC

- $\Delta \boldsymbol{Z}_{\boldsymbol{max}}$: Maximum controllable displacements

Key Components:

- Simulation Principle:
  - ΔZ_{max} is determined via time-dependent simulations that constrain coil current or voltage saturation.
  - Simulation considers non-linear/linear plasma behavior and evolving equilibrium states, allowing for more accurate predictions under varying conditions.
- Plasma Control System (PCS):

  Accounts for PCS latency, power supply limits, and perturbed equilibrium plasma dynamics, response time considerations.

Current Physics-Based Models for Vertical Stability Observers

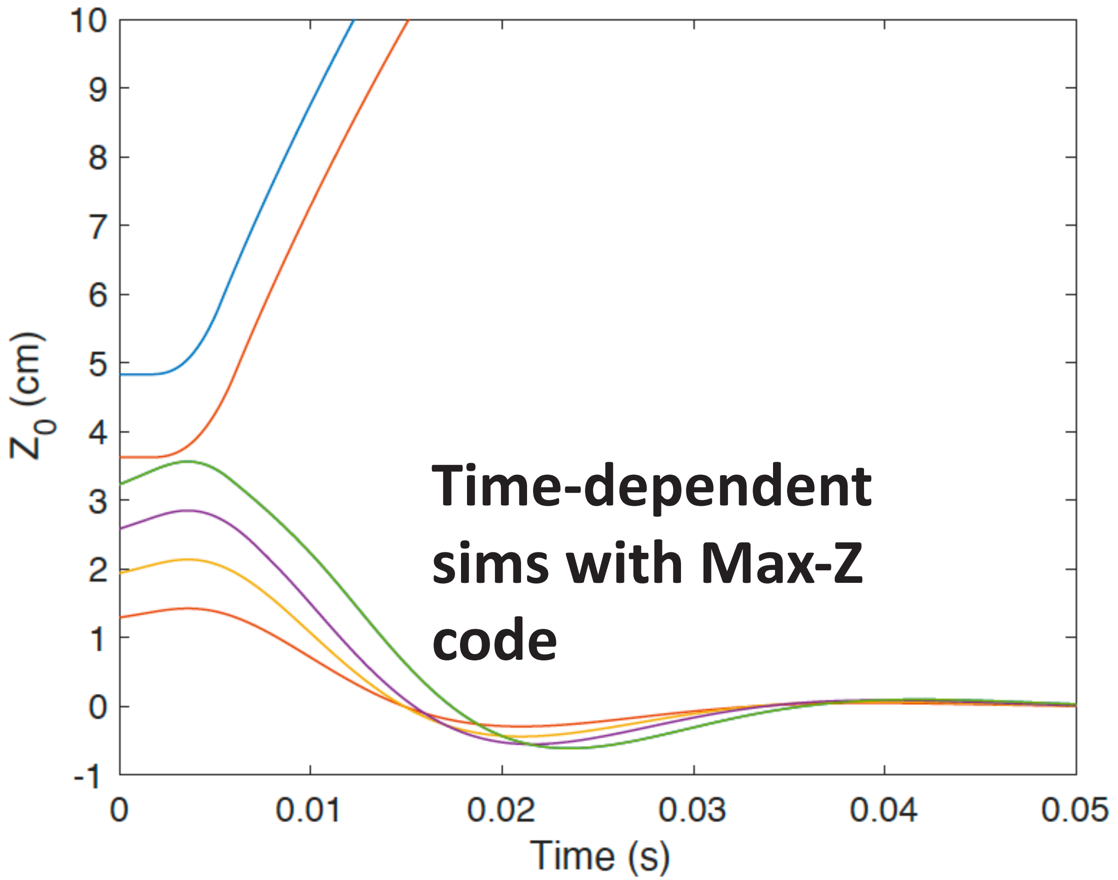
SPARC

❑ ΔZ_{max} : Maximum controllable displacements

Key Components:

- Simulation Principle:
  - ΔZ_{max} is determined via time-dependent simulations that constrain coil current or voltage saturation.
  - Simulation considers non-linear/linear plasma behavior and evolving equilibrium states, allowing for more accurate predictions under varying conditions.
- Plasma Control System (PCS):

  Accounts for PCS latency, power supply limits, and perturbed equilibrium plasma dynamics, response time considerations.



Current Physics-Based Models for Vertical Stability Observers

SPARC

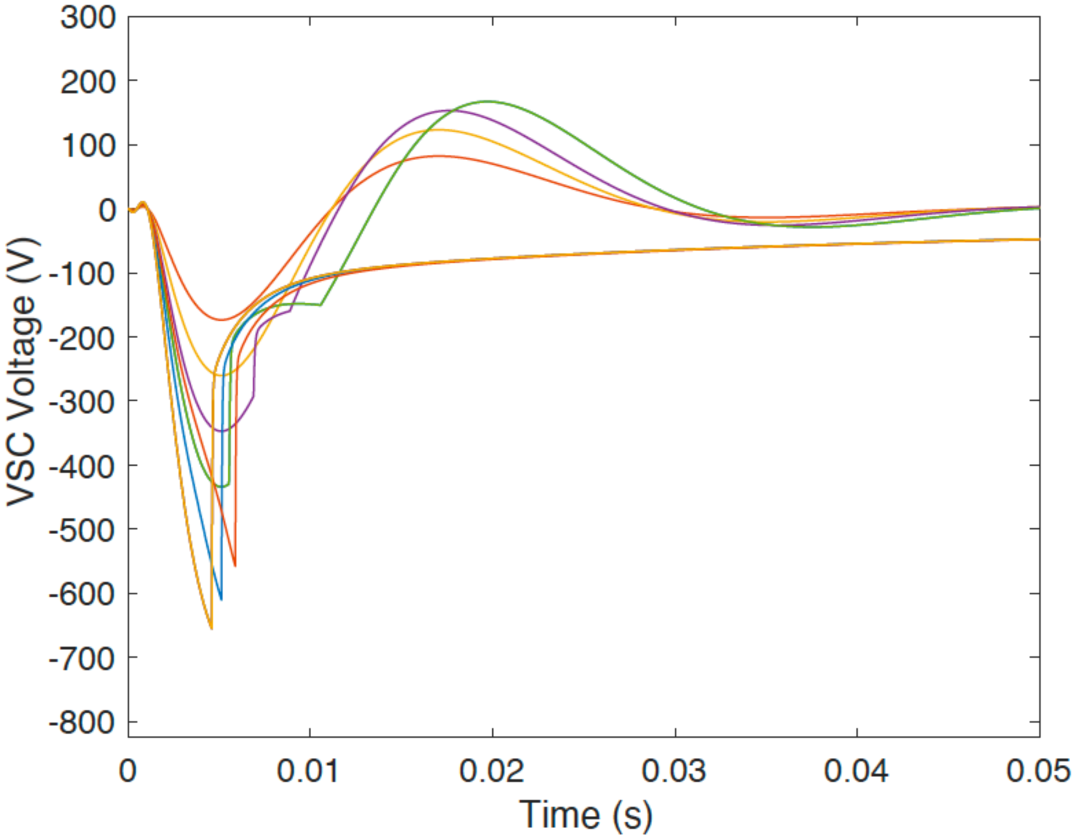
❑ ΔZ_{max} : Maximum controllable displacements

Key Components:

- Simulation Principle:
  - ΔZ_{max} is determined via time-dependent simulations that constrain coil current or voltage saturation.
  - Simulation considers non-linear/linear plasma behavior and evolving equilibrium states, allowing for more accurate predictions under varying conditions.

- Plasma Control System (PCS):

  Accounts for PCS latency, power supply limits, and perturbed equilibrium plasma dynamics, response time considerations.



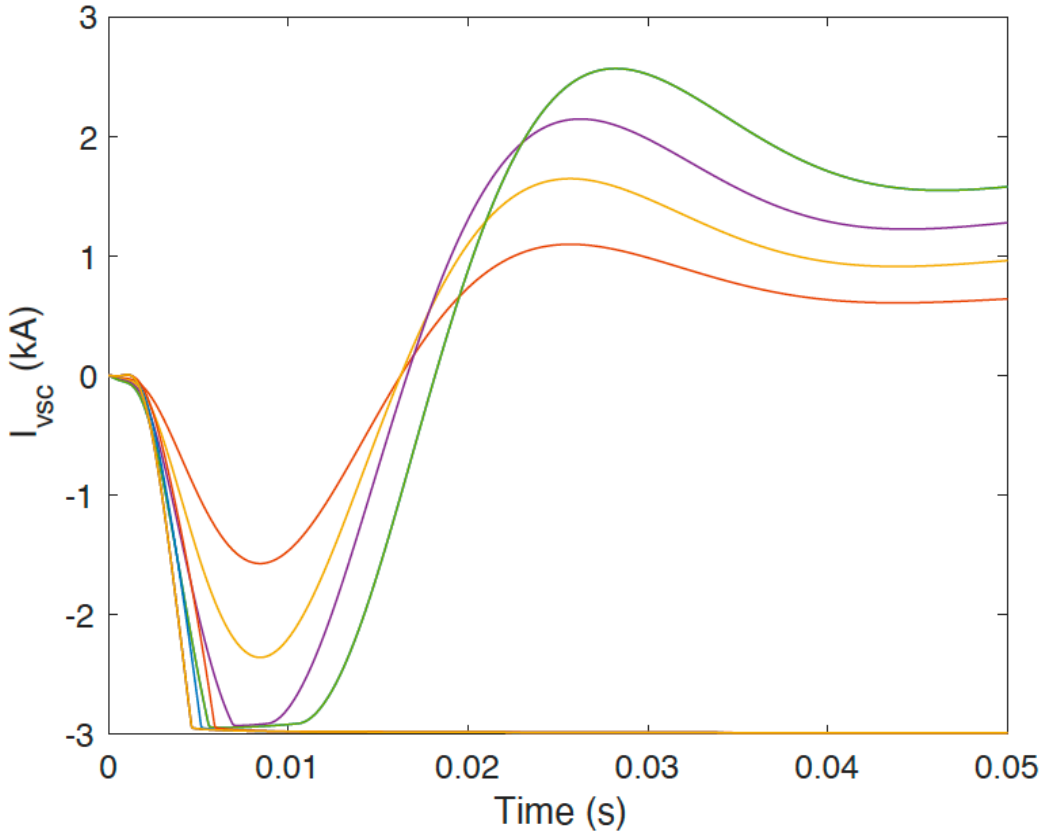
Current Physics-Based Models for Vertical Stability Observers

❑ ΔZ_{max} : Maximum controllable displacements

Key Components:

- Simulation Principle:
  - ΔZ_{max} is determined via time-dependent simulations that constrain coil current or voltage saturation.
  - Simulation considers non-linear/linear plasma behavior and evolving equilibrium states, allowing for more accurate predictions under varying conditions.
- Plasma Control System (PCS):

  Accounts for PCS latency, power supply limits, and perturbed equilibrium plasma dynamics, response time considerations.



Current Physics-Based Models for Vertical Stability Observers

SPARC

❑ ΔZ_{max} : Maximum controllable displacements

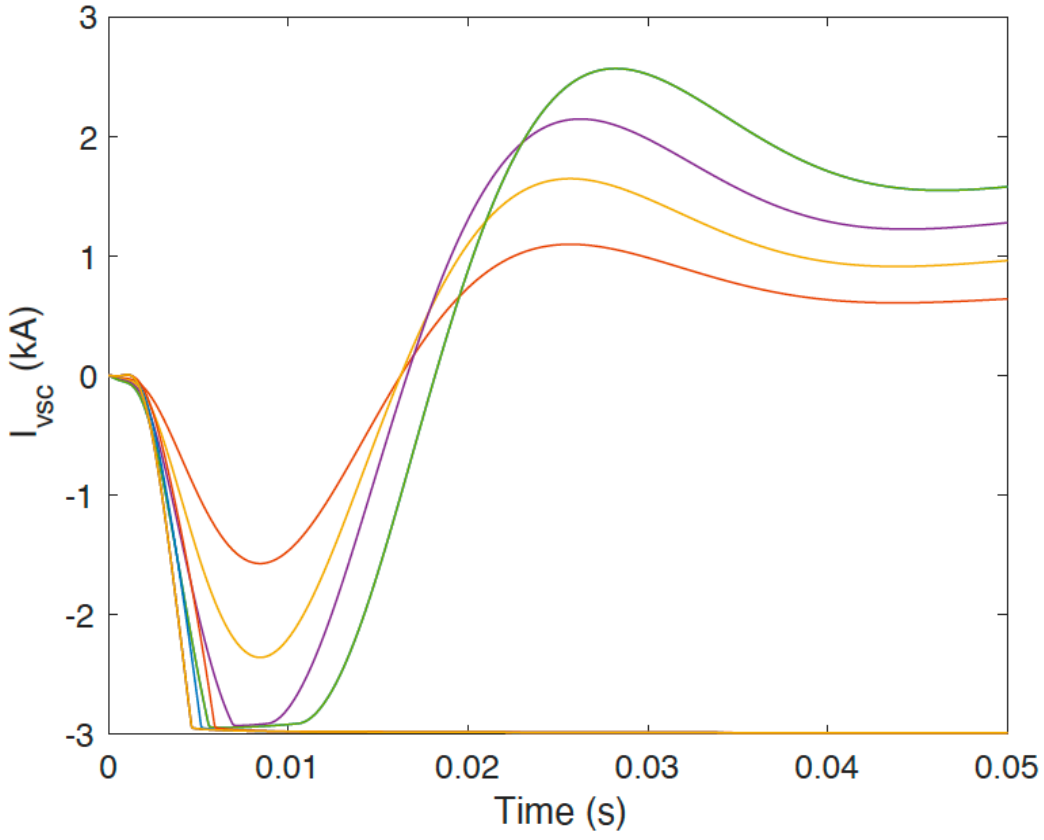
Key Components:

- Simulation Principle:
  - ΔZ_{max} is determined via time-dependent simulations that constrain coil current or voltage saturation.
  - Simulation considers non-linear/linear plasma behavior and evolving equilibrium states, allowing for more accurate predictions under varying conditions.
- Plasma Control System (PCS):

  Accounts for PCS latency, power supply limits, and perturbed equilibrium plasma dynamics, response time considerations.

Also exist, Humphrey's analytical ΔZ_{max} threshold :

$$\Delta Z_{max} \sim -\frac{\partial z}{\partial I_e}\overrightarrow{\delta I_e} L_{*e}^{-1} b_{ez} \frac{I_{max}R_c}{\gamma_z} e^{-\gamma_z T_{ps}}$$



Current Physics-Based Models for Vertical Stability Observers

SPARC

- ΔZ_{max} : Maximum controllable displacements

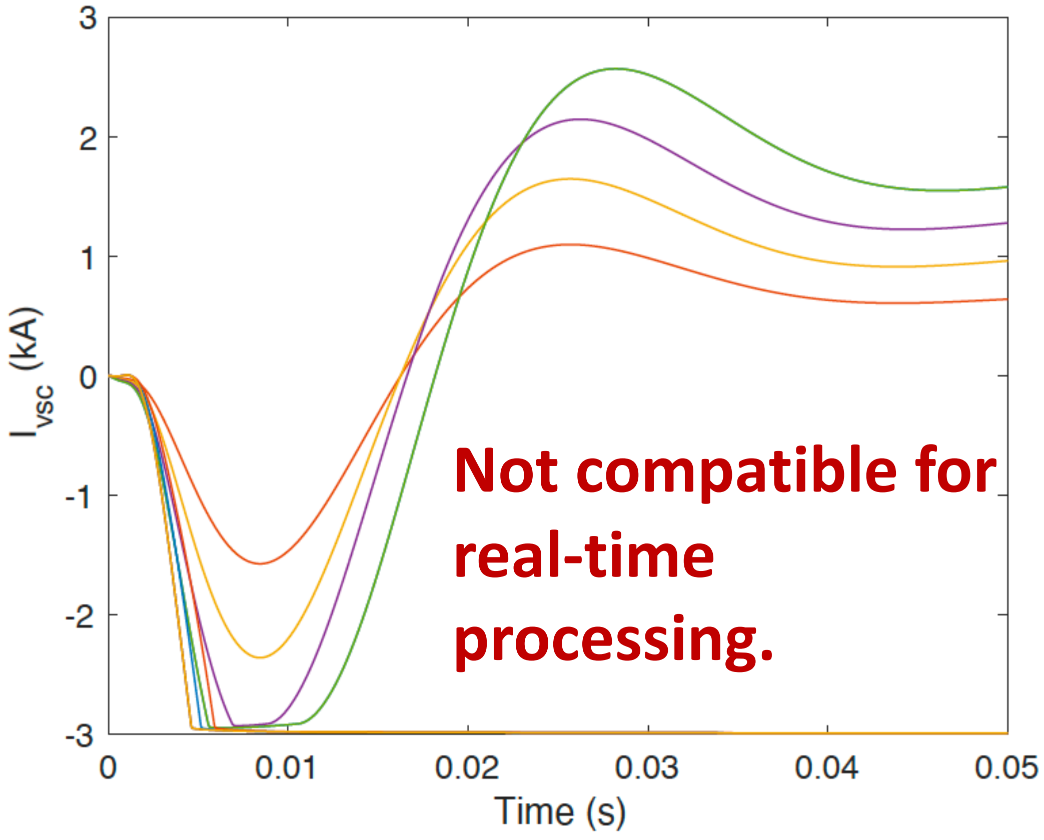
Key Components:

- Simulation Principle:
  - ΔZ_{max} is determined via time-dependent simulations that constrain coil current or voltage saturation.
  - Simulation considers non-linear/linear plasma behavior and evolving equilibrium states, allowing for more accurate predictions under varying conditions.
- Plasma Control System (PCS):

  Accounts for PCS latency, power supply limits, and perturbed equilibrium plasma dynamics, response time considerations.

Also exist, Humphrey's analytical ΔZ_{max} threshold :

$$\Delta Z_{max} \sim -\frac{\partial z}{\partial I_e}\overrightarrow{\delta I_e} L_{*e}^{-1} b_{ez} \frac{I_{max} R_c}{\gamma_z} e^{-\gamma_z T_{ps}}$$



Current Physics-Based Models for Vertical Stability Observers

SPARC

❑ ΔZ_{max} : Maximum controllable displacements

Key Components:

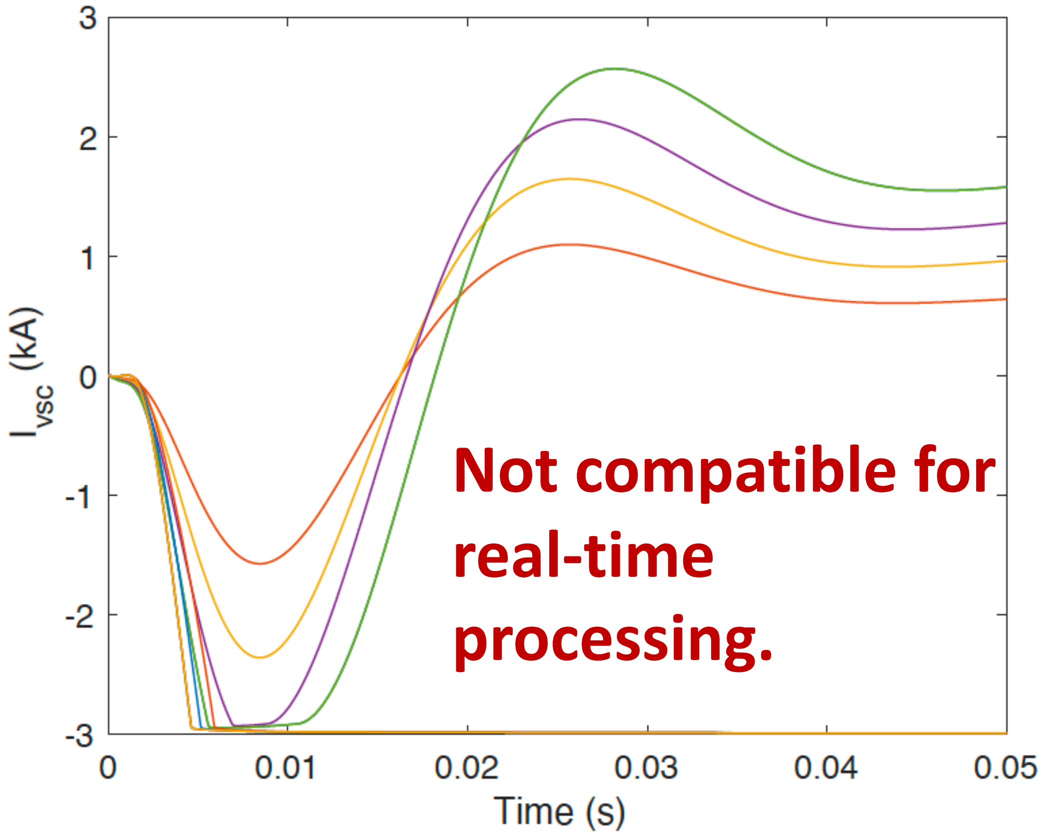
- Simulation Principle:
    - ΔZ_{max} is determined via time-dependent simulations that constrain coil current or voltage saturation.
    - Simulation considers non-linear/linear plasma behavior and evolving equilibrium states, allowing for more accurate predictions under varying conditions.

- Plasma Control System (PCS):

  Accounts for PCS latency, power supply limits, and perturbed equilibrium plasma dynamics, response time considerations.

Also exist, Humphrey's analytical ΔZ_{max} threshold :

$$\Delta Z_{max} \sim -\frac{\partial z}{\partial I_e}\overrightarrow{\delta I_e} L_{*e}^{-1} b_{ez} \frac{I_{max} R_c}{\gamma_z} e^{-\gamma_z T_{ps}}$$



Not compatible for real-time processing.

Surrogate model for γ_z & ΔZ_{max} with Deep-learning framework
Transolvers

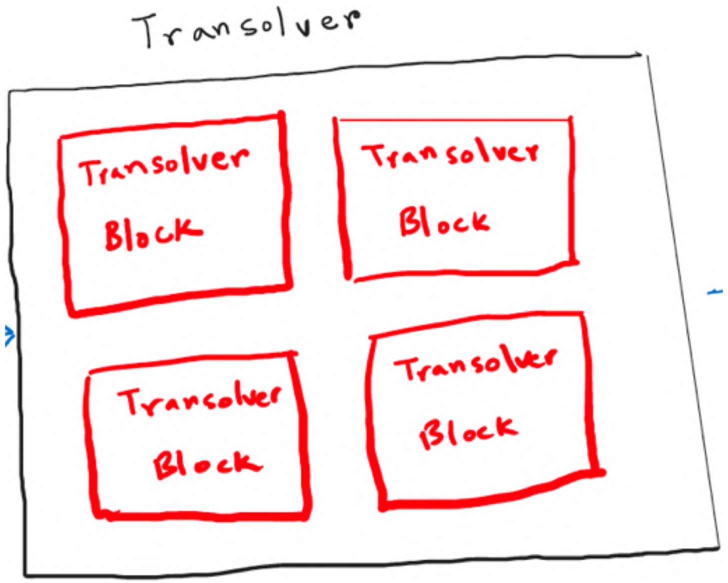
Leveraging Transformer-based Neural Net Architectures

Learning operators with Physics-Attention mechanism

Transolver: AI model designed to solve PDEs with complex geometries

Haixu Wu, Huakun Luo et. al https://arxiv.org/abs/2402.02366

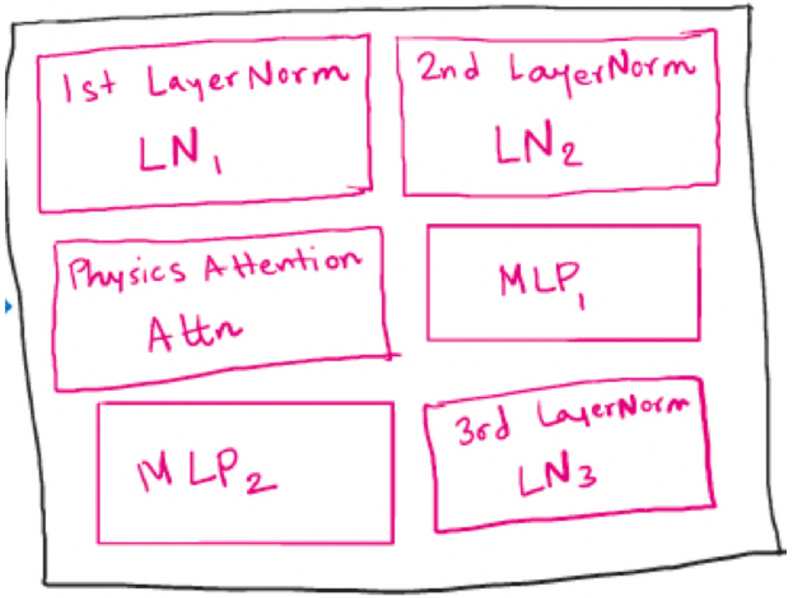
ICML, 2024 Spotlight



Transolver: AI model designed to solve PDEs with complex geometries

Haixu Wu, Huakun Luo et. al https://arxiv.org/abs/2402.02366

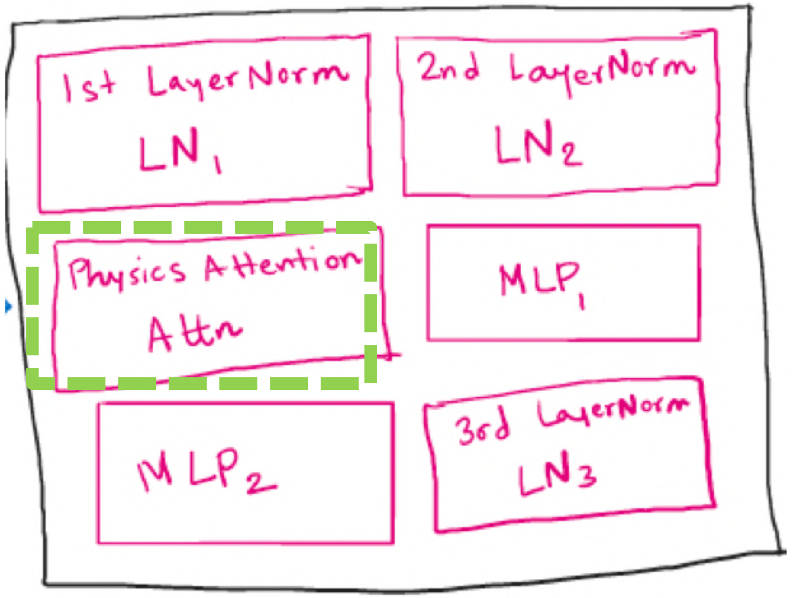
ICML, 2024 Spotlight



Transolver: AI model designed to solve PDEs with complex geometries

Haixu Wu, Huakun Luo et. al https://arxiv.org/abs/2402.02366

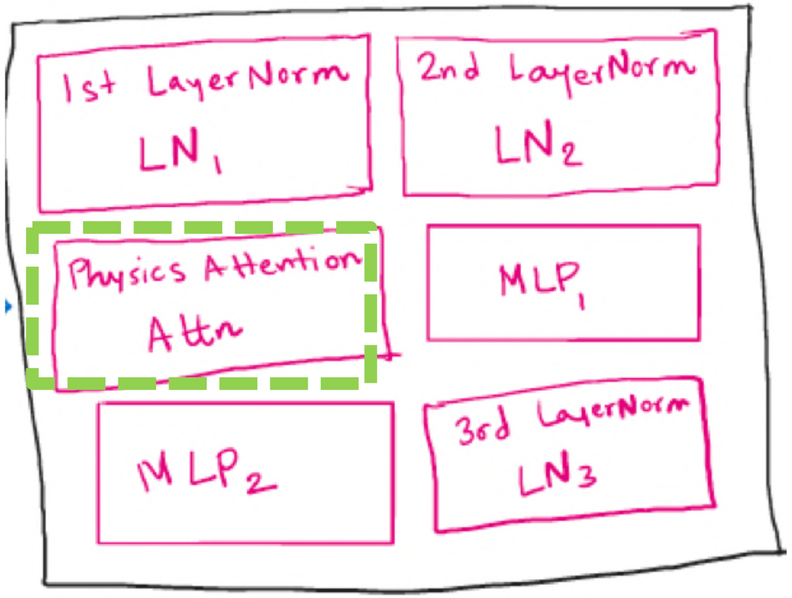
ICML, 2024 Spotlight

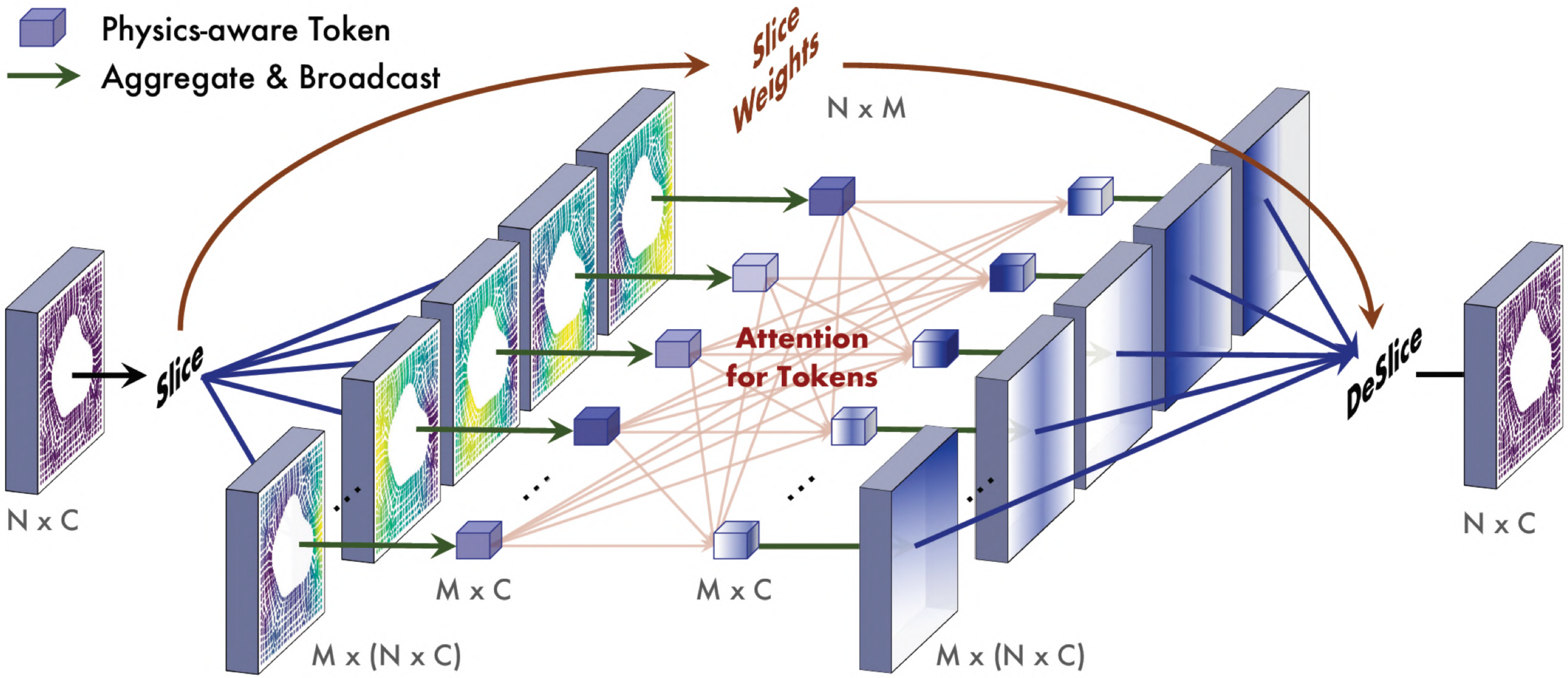


Transolver: AI model designed to solve PDEs with complex geometries

Haixu Wu, Huakun Luo et. al https://arxiv.org/abs/2402.02366

ICML, 2024 Spotlight

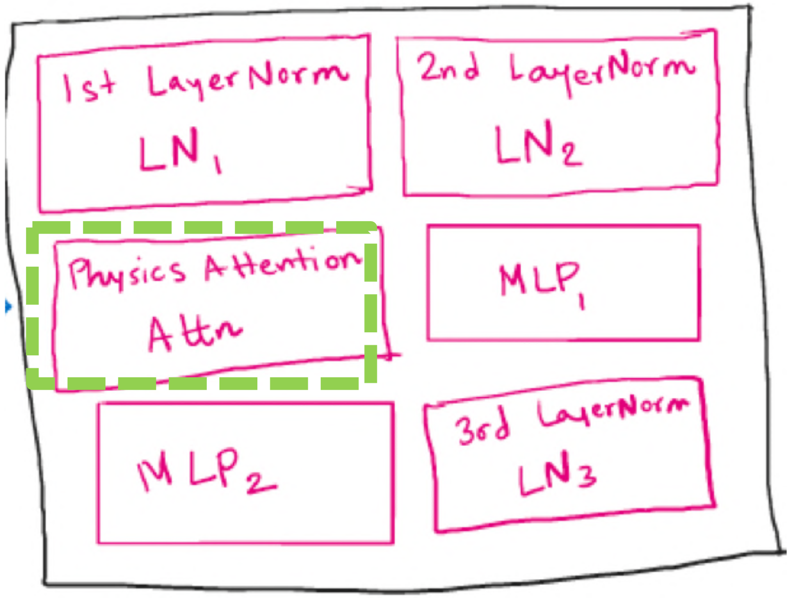


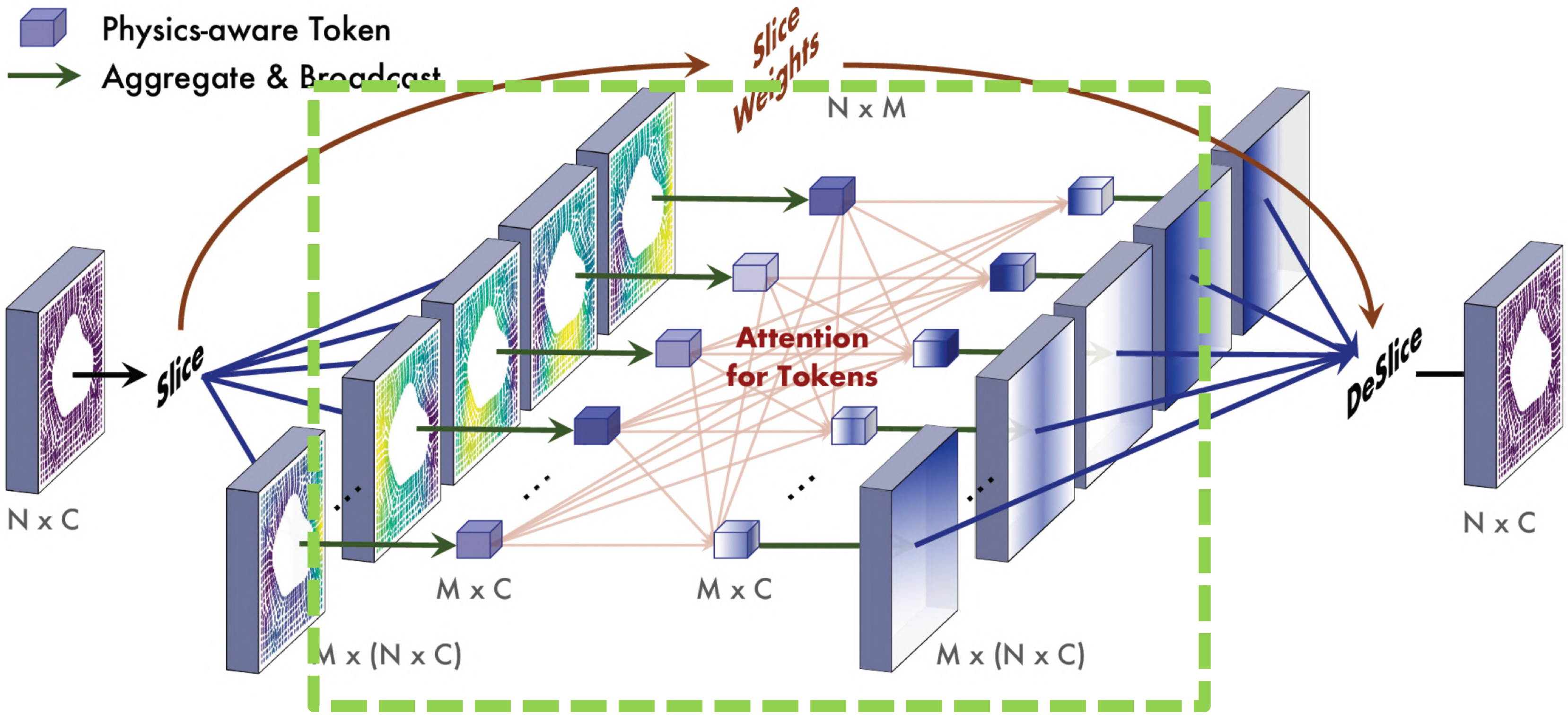


Transolver: AI model designed to solve PDEs with complex geometries

Haixu Wu, Huakun Luo et. al https://arxiv.org/abs/2402.02366

ICML, 2024 Spotlight

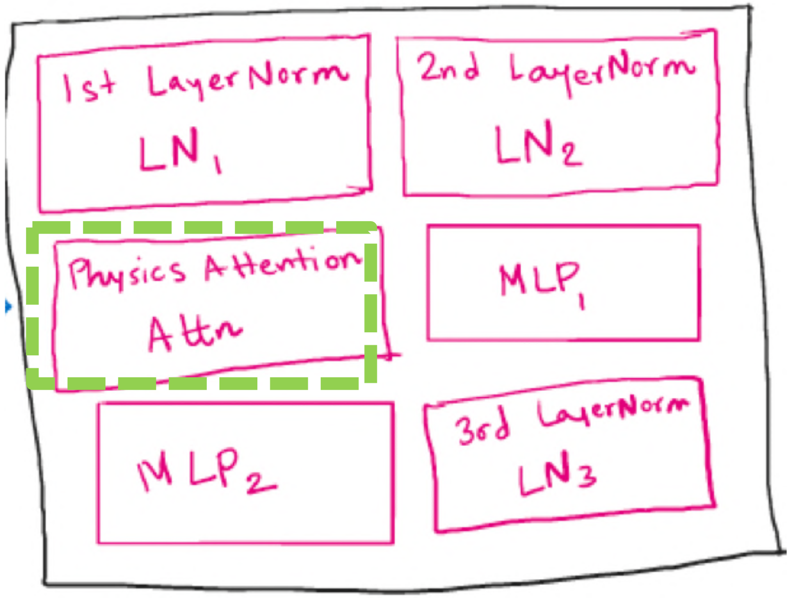


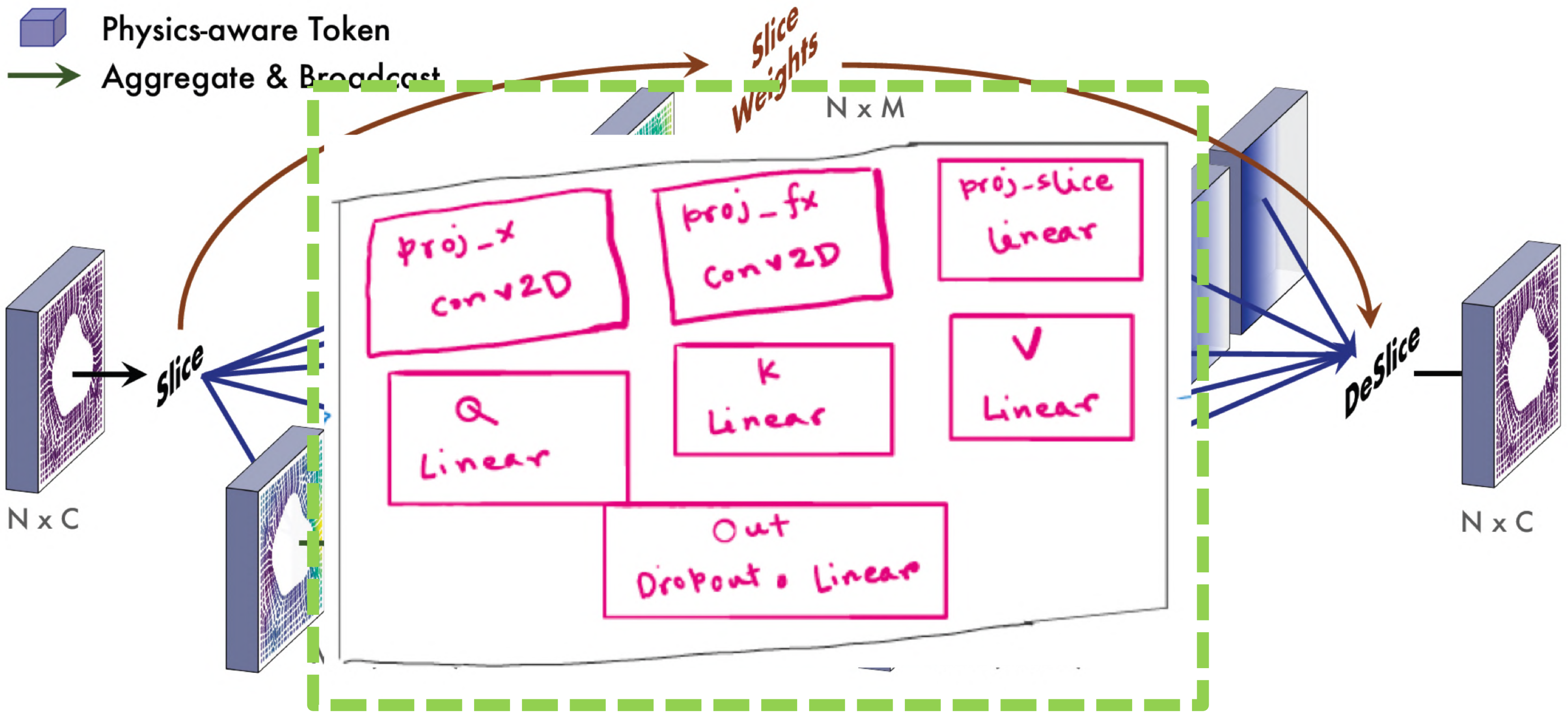


Transolver: AI model designed to solve PDEs with complex geometries

Haixu Wu, Huakun Luo et. al https://arxiv.org/abs/2402.02366

ICML, 2024 Spotlight

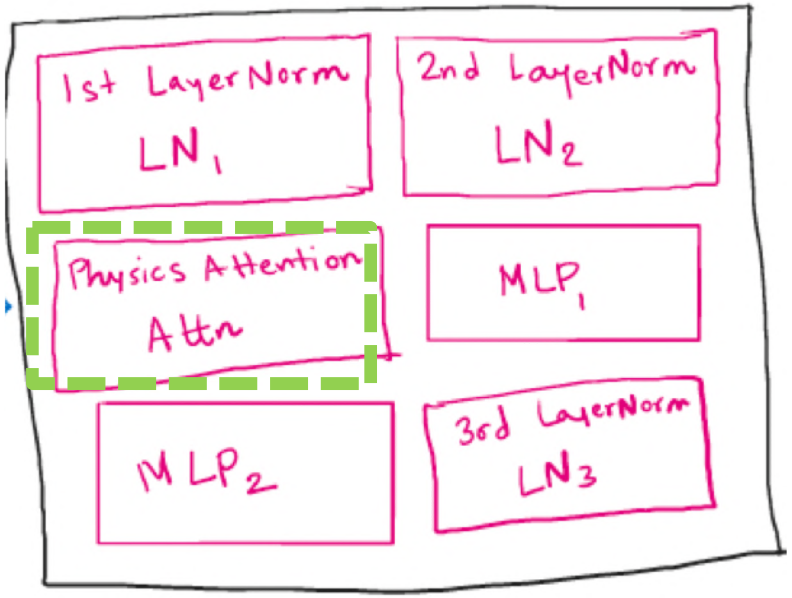


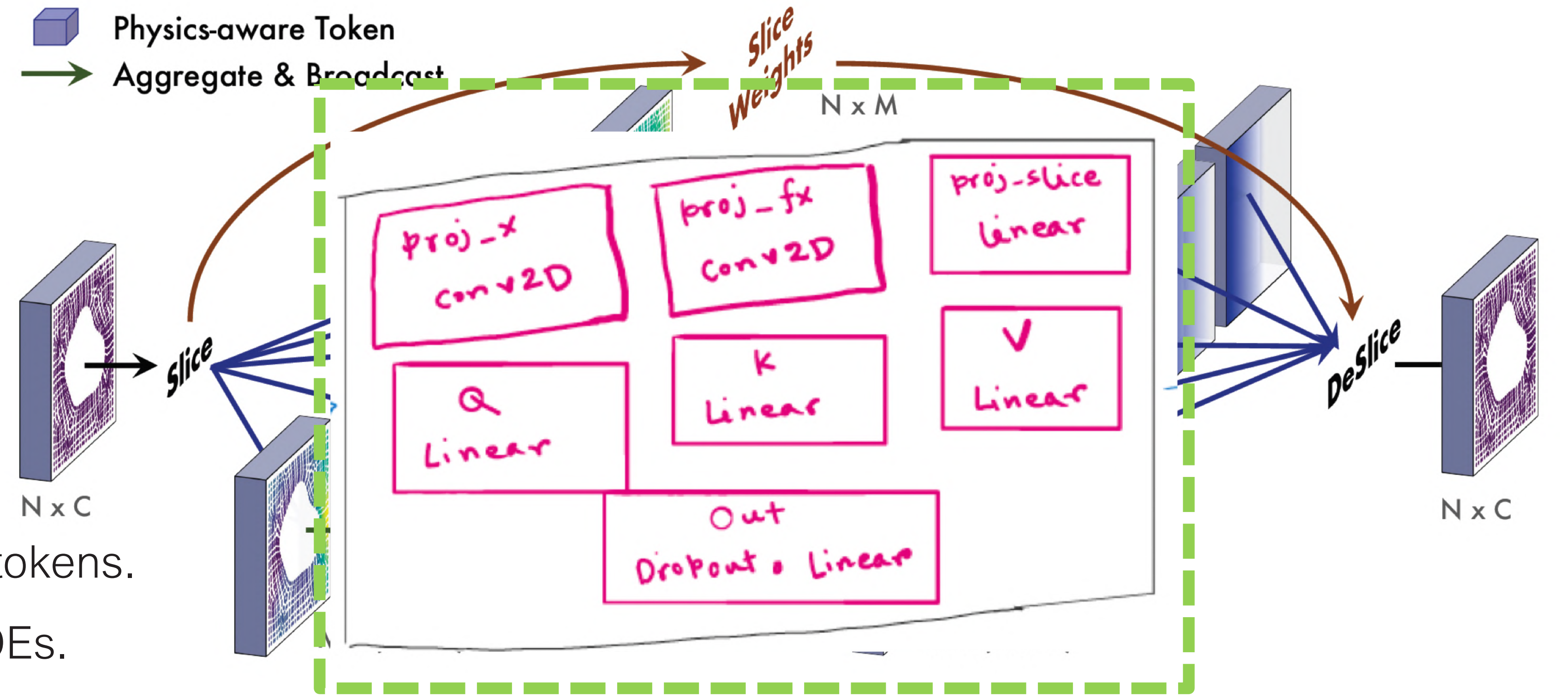


Transolver: AI model designed to solve PDEs with complex geometries

Haixu Wu, Huakun Luo et. al https://arxiv.org/abs/2402.02366

ICML, 2024 Spotlight



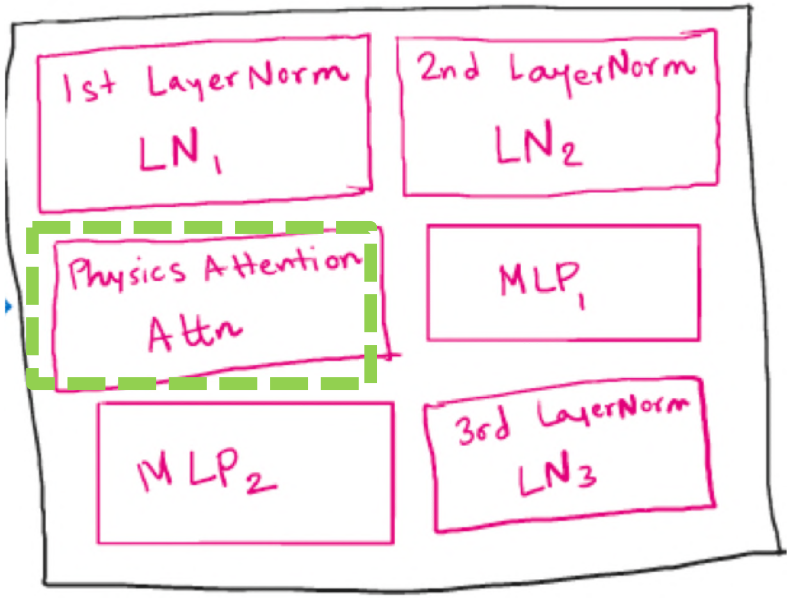


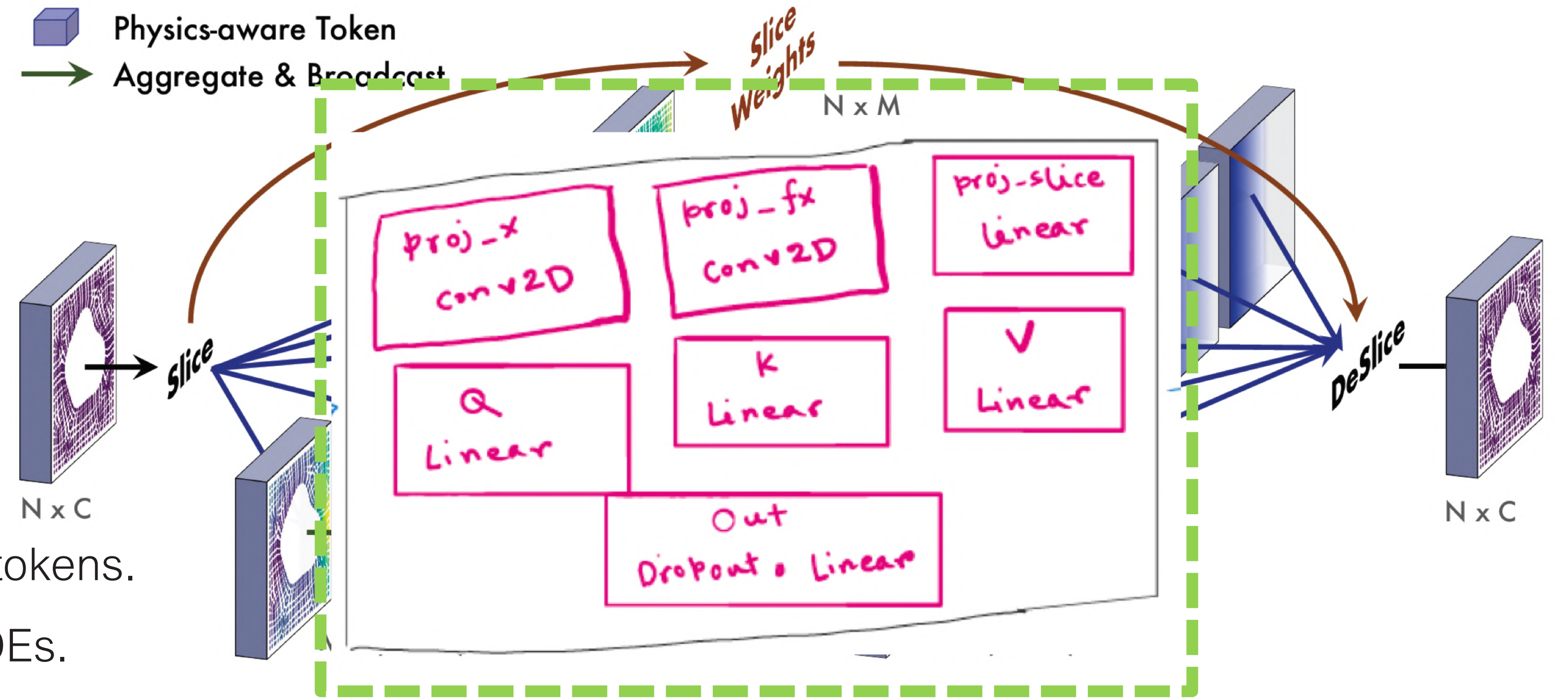
- Attention among physics-aware tokens.
- Approximate Integral to solve PDEs.

Transolver: AI model designed to solve PDEs with complex geometries

Haixu Wu, Huakun Luo et. al https://arxiv.org/abs/2402.02366

ICML, 2024 Spotlight





- Attention among physics-aware tokens.
- Approximate Integral to solve PDEs.

Transolvers out performs other PDE-NN solvers.

Building a Transolver Surrogate: Leveraging C-Mod and SPARC Databases for Comprehensive Training

Building a Transolver Surrogate: Leveraging C-Mod and SPARC Databases for Comprehensive Training

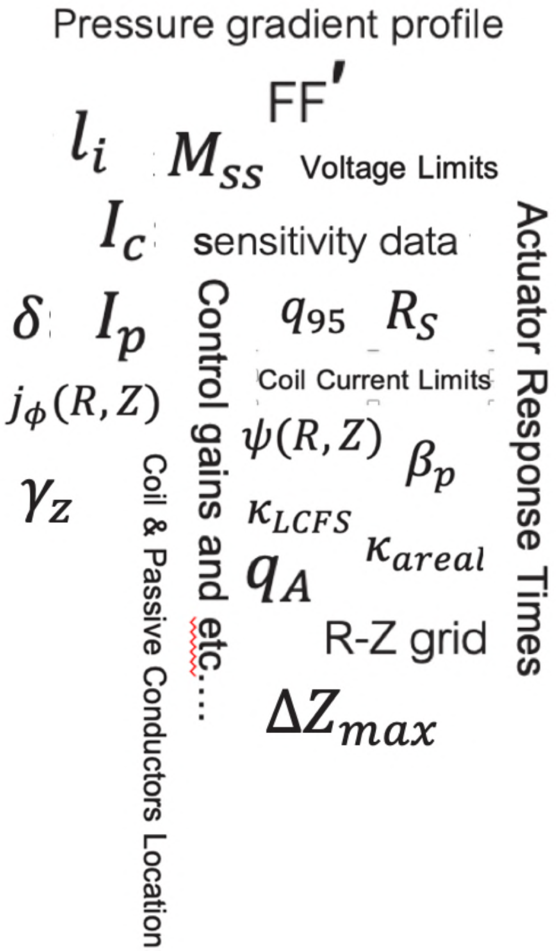
- C-Mod disruptive database 2012-2016 (also includes stable discharges)
- Includes > 10,000 discharge scenarios with diverse plasma conditions.
- Key for capturing the dynamics of disruptions and equilibrium shifts in various conditions.
- SPARC Primary Reference Discharge (1400s, 1500s, 1600s, 1700s series)
    - With L-mode, H-mode scenarios.
    - With L-H transition scenarios.

Building a Transolver Surrogate: Leveraging C-Mod and SPARC Databases for Comprehensive Training

- C-Mod disruptive database 2012-2016 (also includes stable discharges)
- Includes > 10,000 discharge scenarios with diverse plasma conditions.
- Key for capturing the dynamics of disruptions and equilibrium shifts in various conditions.
- SPARC Primary Reference Discharge (1400s, 1500s, 1600s, 1700s series)
  - With L-mode, H-mode scenarios.
  - With L-H transition scenarios.

Building a Transolver Surrogate: Leveraging C-Mod and SPARC Databases for Comprehensive Training

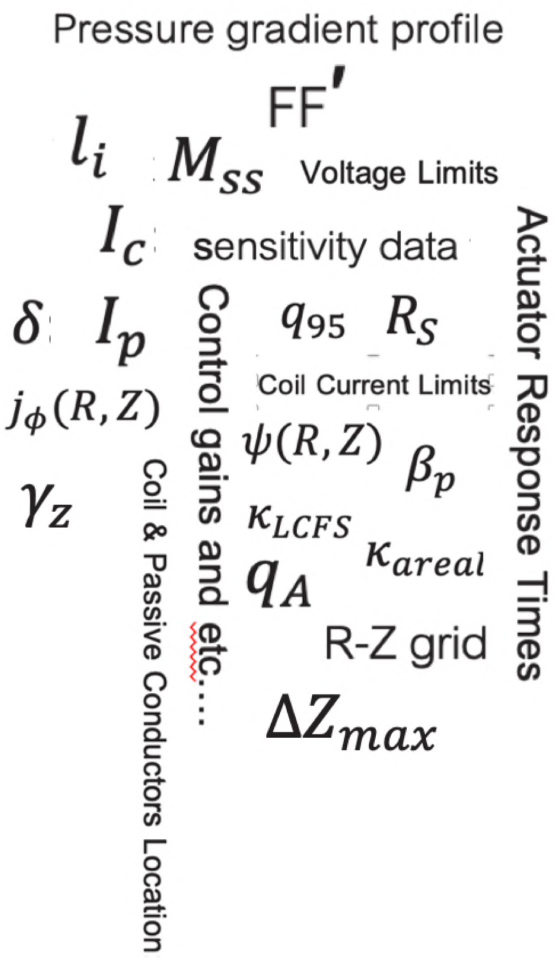
- C-Mod disruptive database 2012-2016 (also includes stable discharges)
- Includes > 10,000 discharge scenarios with diverse plasma conditions.
- Key for capturing the dynamics of disruptions and equilibrium shifts in various conditions.
- SPARC Primary Reference Discharge (1400s, 1500s, 1600s, 1700s series)
  - With L-mode, H-mode scenarios.
  - With L-H transition scenarios.



Building a Transolver Surrogate: Leveraging C-Mod and SPARC Databases for Comprehensive Training

- C-Mod disruptive database 2012-2016 (also includes stable discharges)
- Includes > 10,000 discharge scenarios with diverse plasma conditions.
- Key for capturing the dynamics of disruptions and equilibrium shifts in various conditions.
- SPARC Primary Reference Discharge (1400s, 1500s, 1600s, 1700s series)
  - With L-mode, H-mode scenarios.
  - With L-H transition scenarios.

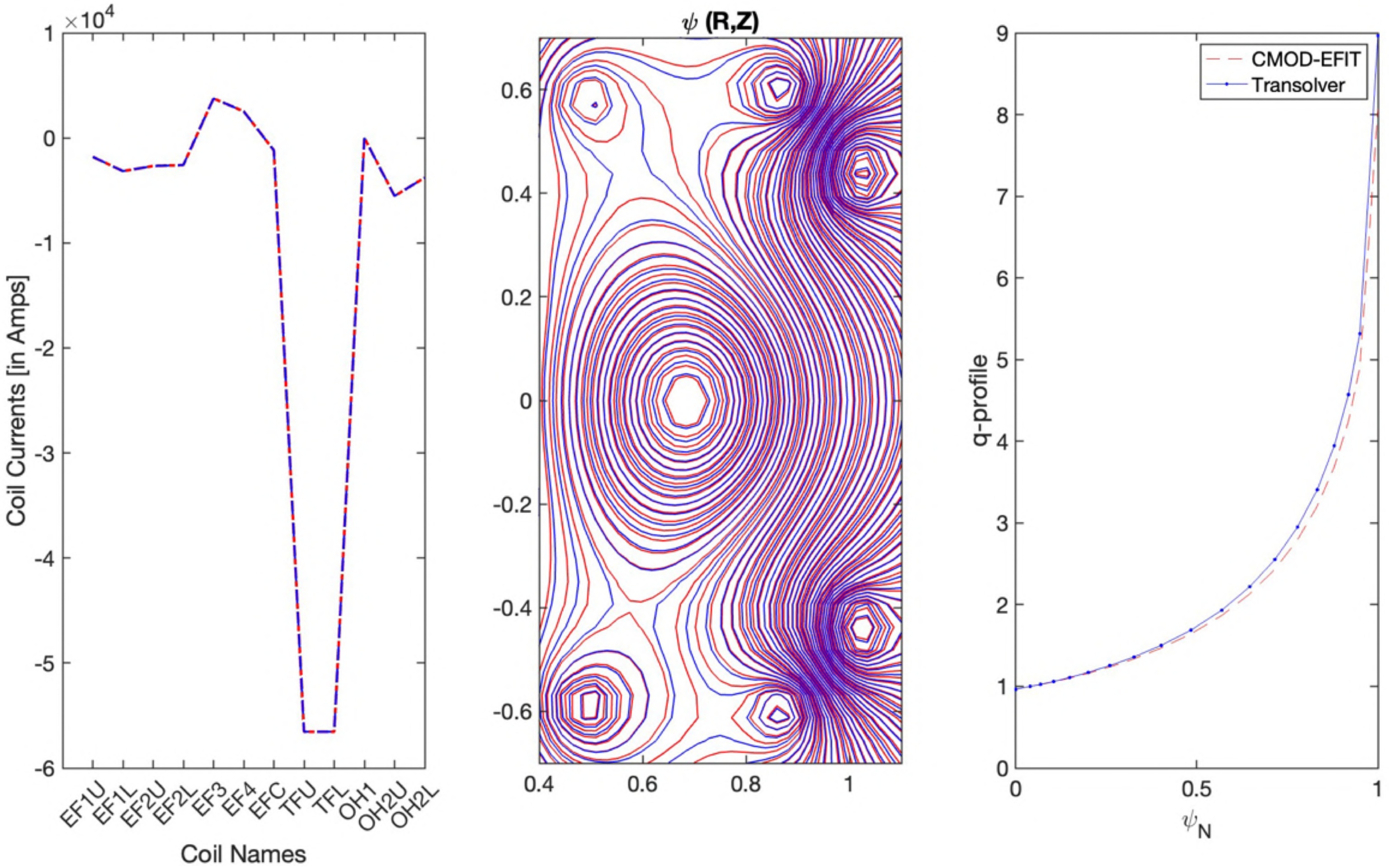
Alcator C-Mod



Surrogate model is being trained on 6 A100 NVIDIA GPUs to accelerate model development.

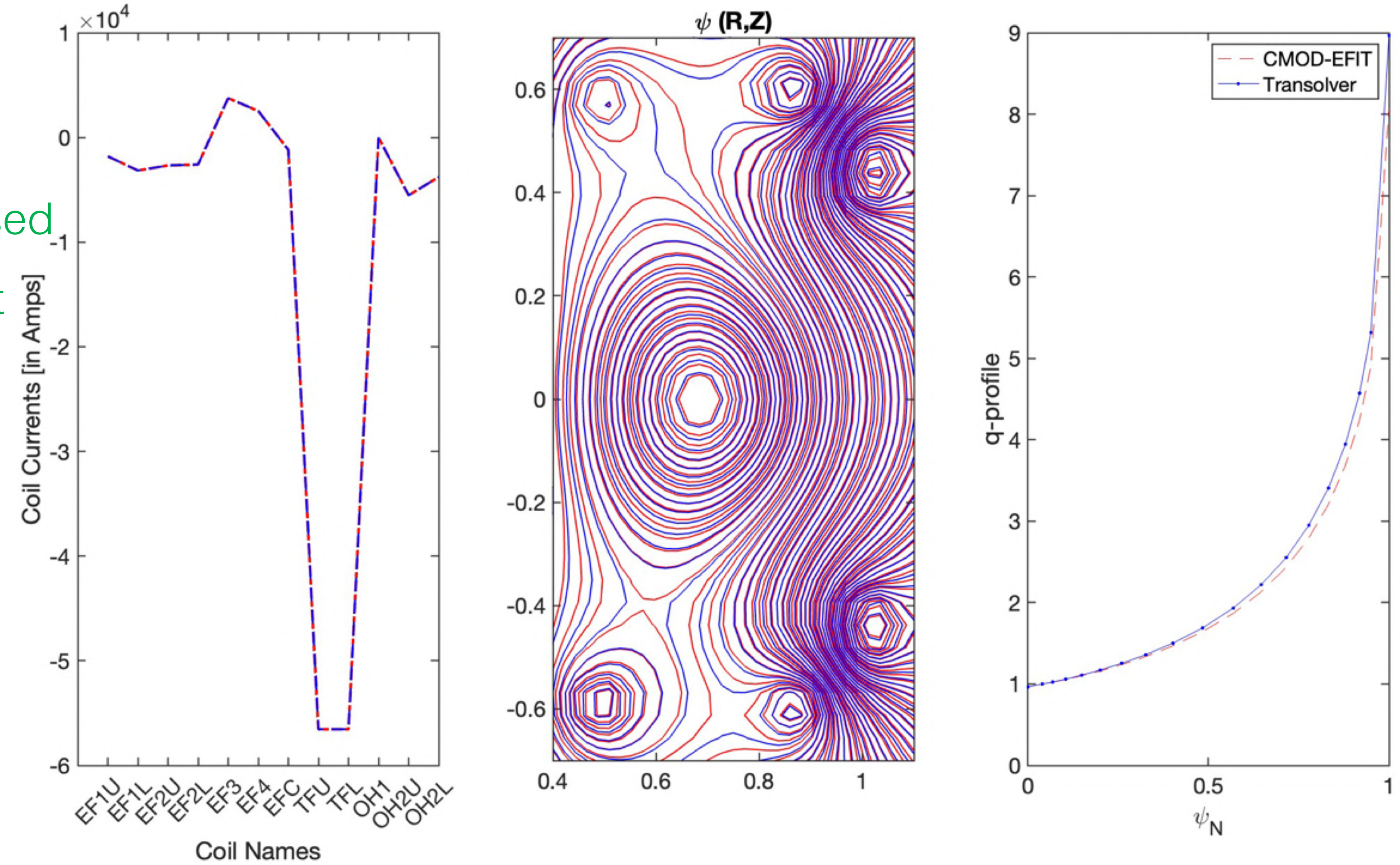
Transolver Surrogate: Achieving Near-Perfect Agreement with C-Mod EFIT

- Transolver surrogate also acts as an fast grad-shafranov equilibrium solver.



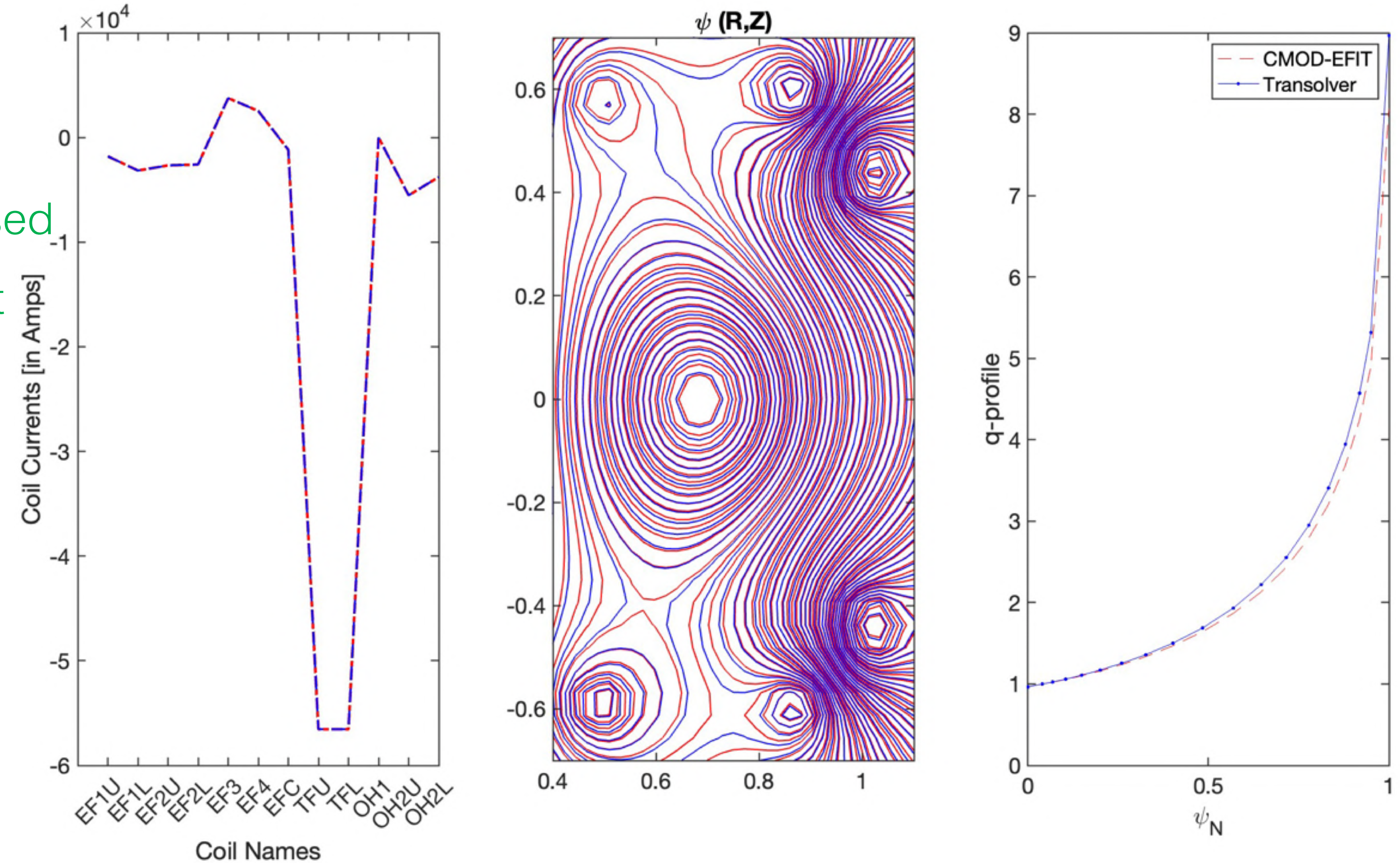
Transolver Surrogate: Achieving Near-Perfect Agreement with C-Mod EFIT

- Transolver surrogate also acts as an fast grad-shafranov equilibrium solver.
- Predictive capabilities of Transolver based surrogate shows an 94-95 % agreement with C-Mod EFIT.



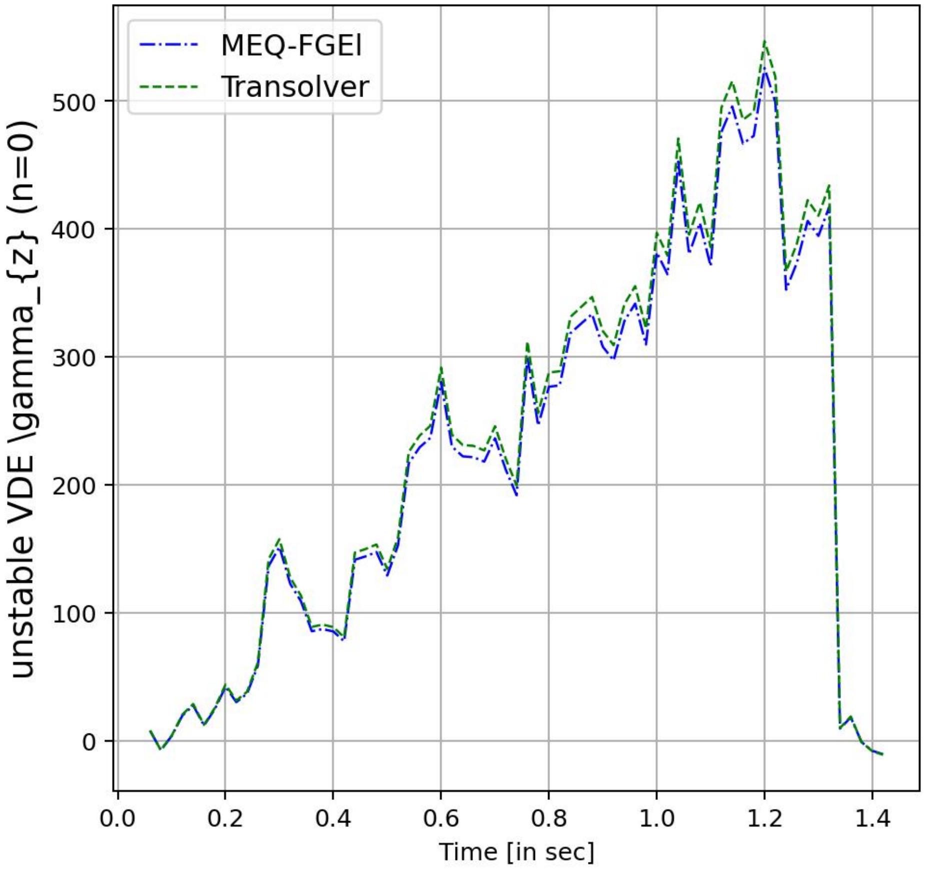
Transolver Surrogate: Achieving Near-Perfect Agreement with C-Mod EFIT

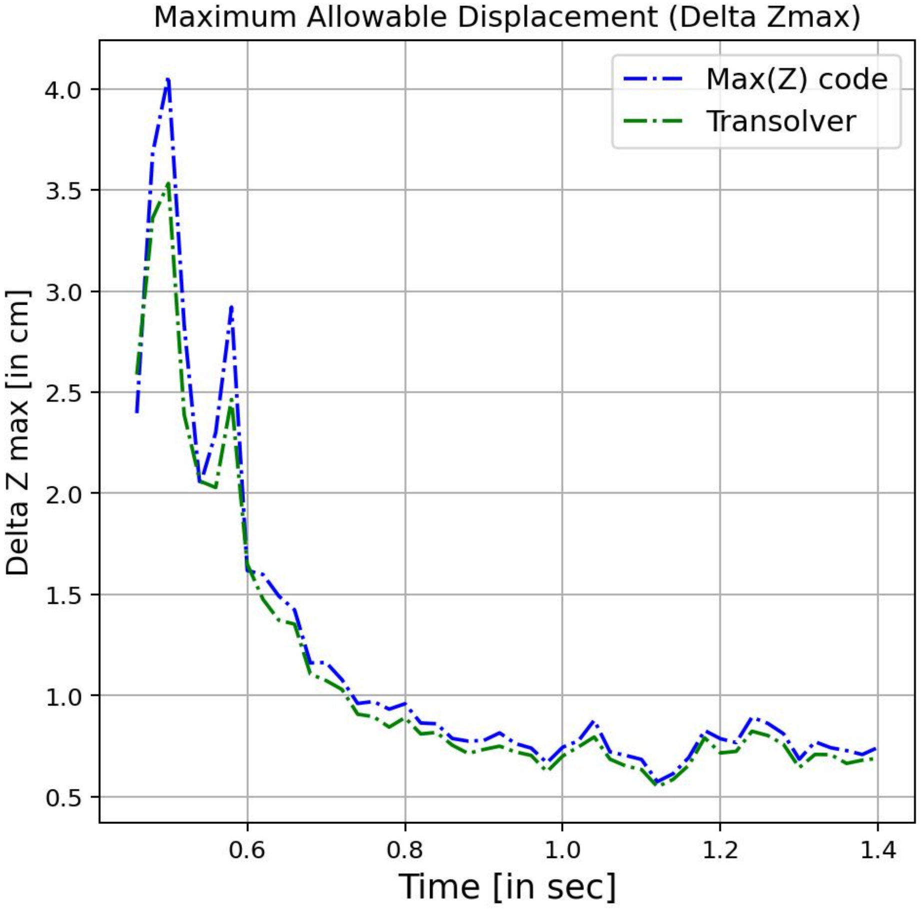
- Transolver surrogate also acts as an fast grad-shafranov equilibrium solver.
- Predictive capabilities of Transolver based surrogate shows an 94-95 % agreement with C-Mod EFIT.
- Ideal Response Time: Achieves predictions in 10-20 ms on multi-core CPUs, enabling near real-time control potential.



Accelerating VDE Predictions: Transolver Achieves High Accuracy in a Fraction of the Time

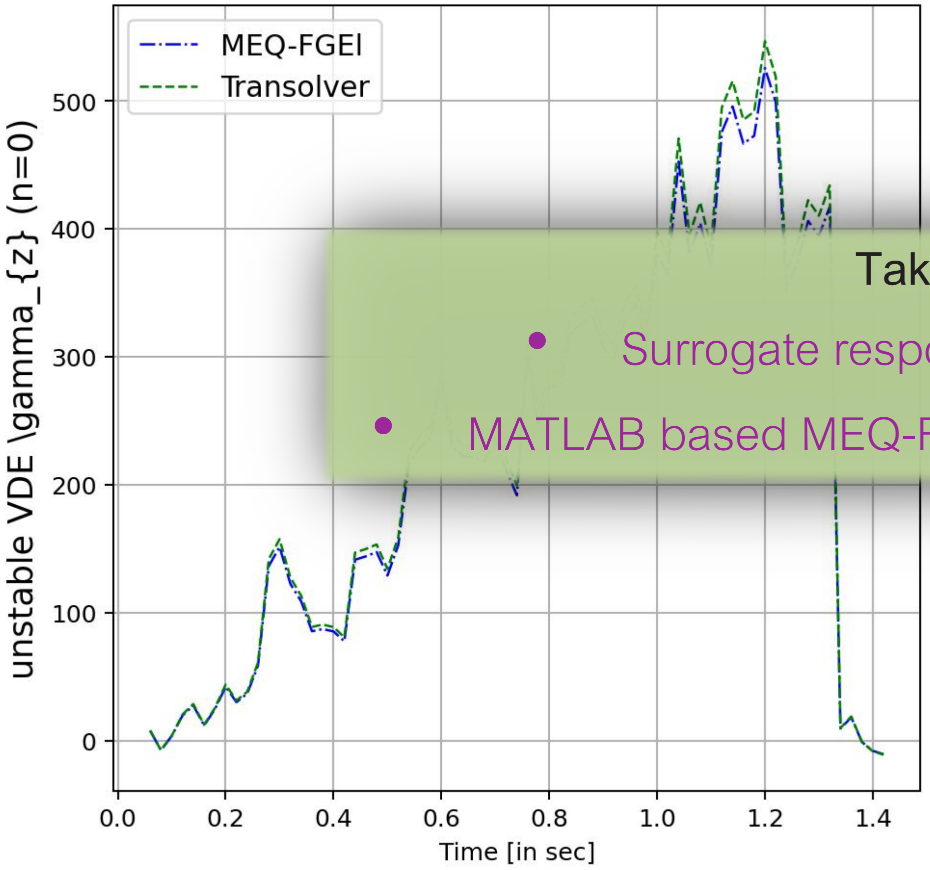
Transolver predicted VDE γ_Z & Max-Z shows an agreement with Physics-based models with relative error of 4-5 %

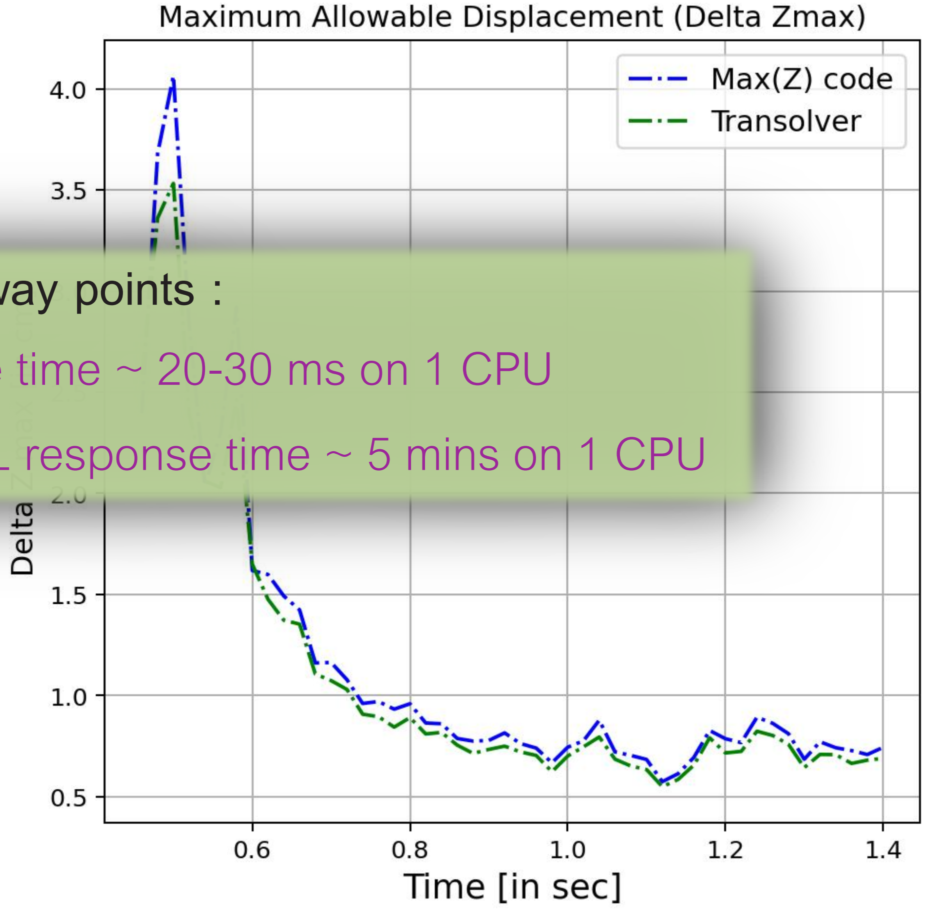




Accelerating VDE Predictions: Transolver Achieves High Accuracy in a Fraction of the Time

Transolver predicted VDE γ_Z & Max-Z shows an agreement with Physics-based models with relative error of 4-5 %

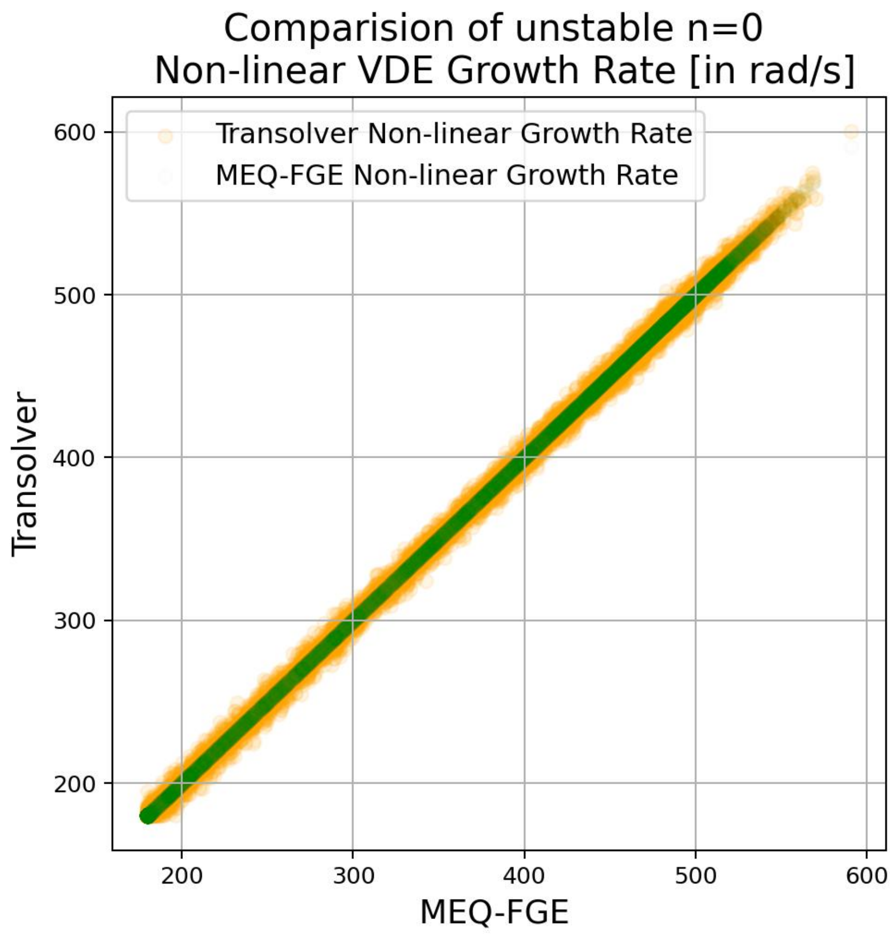


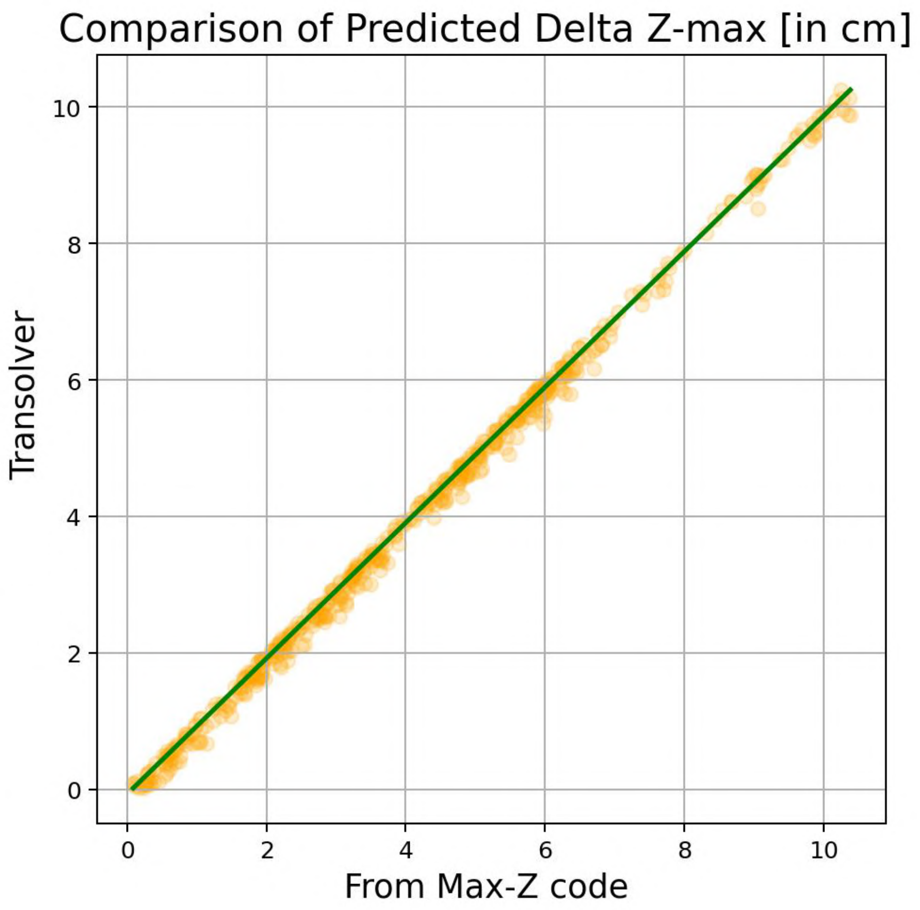


Take away points :

- Surrogate response time ~ 20-30 ms on 1 CPU
- MATLAB based MEQ-FGEL response time ~ 5 mins on 1 CPU

Strong Correlation Between Transolver Surrogate and MEQ-FGE: Consistent Predictions of VDE Dynamics





Future Features/Work

- Measurements of uncertainties, coming both from limited resolution in data and noise (in synthetic database).

- **Train the surrogate for more tokamak databases. Need collaborations to validate this framework in other tokamaks such as DIIID, TCV and etc.**

- Coupling with SPARC ONW and ONSIM framework to test it's predictivity a/c to disruption warning time-scale. Plan also scales to ARC reactor scenarios.

- Inclusion of a predictive model for currents induced in passive conductors during disruption or after current quench.

BONUS SLIDES

Transolver: A Fast Transformer Solver for PDEs on General Geometries

Table 2. Performance comparison on standard benchmarks. Relative L2 is recorded. A smaller value indicates better performance. For clarity, the best result is in bold and the second best is underlined. Promotion refers to the relative error reduction w.r.t. the second best model ($1 - \frac{\text{Our error}}{\text{The second best error}}$) on each benchmark. "/" means that the baseline cannot apply to this benchmark.

Model	Point Cloud	Structured Mesh			Regular Grid	
	Elasticity	Plasticity	Airfoil	Pipe	Navier–Stokes	Darcy
FNO (Li et al., 2021)	/	/	/	/	0.1556	0.0108
WMT (Gupta et al., 2021)	0.0359	0.0076	0.0075	0.0077	0.1541	0.0082
U-FNO (Wen et al., 2022)	0.0239	0.0039	0.0269	0.0056	0.2231	0.0183
geo-FNO (Li et al., 2022)	0.0229	0.0074	0.0138	0.0067	0.1556	0.0108
U-NO (Rahman et al., 2023)	0.0258	0.0034	0.0078	0.0100	0.1713	0.0113
F-FNO (Tran et al., 2023)	0.0263	0.0047	0.0078	0.0070	0.2322	0.0077
LSM (Wu et al., 2023)	0.0218	0.0025	<u>0.0059</u>	0.0050	0.1535	<u>0.0065</u>
Galerkin (Cao, 2021)	0.0240	0.0120	0.0118	0.0098	0.1401	0.0084
HT-Net (Liu et al., 2022)	/	0.0333	0.0065	0.0059	0.1847	0.0079
OFormer (Li et al., 2023c)	0.0183	<u>0.0017</u>	0.0183	0.0168	0.1705	0.0124
GNOT (Hao et al., 2023)	<u>0.0086</u>	0.0336	0.0076	<u>0.0047</u>	0.1380	0.0105
FactFormer (Li et al., 2023d)	/	0.0312	0.0071	0.0060	0.1214	0.0109
ONO (Xiao et al., 2024)	0.0118	0.0048	0.0061	0.0052	<u>0.1195</u>	0.0076
Transolver (Ours)	**0.0064**	**0.0012**	**0.0053**	**0.0033**	**0.0900**	**0.0057**
Relative Promotion	25.6%	29.4%	10.2%	29.7%	24.7%	12.3%

Building a Robust Transolver Surrogate: Leveraging CMOD and SPARC Databases for Comprehensive Training

Plasma Current & Shape Parameters:

- I_c: Coil currents. R-Z grid
- q_A: Plasma current.
- β_p: Poloidal beta, indicating the plasma pressure.
- κ_{areal} : Internal inductance, describing the current distribution in the plasma.
- q_{95}: Safety factors, describing magnetic field line pitch.
- δ**:** Plasma triangularity

Pressure & Temperature Profiles:

- P$'$: Pressure gradient profile .
- FF$'$: Temperature gradient, influencing plasma stability.
- $j_\phi(R, Z)$: Toroidal current density across radial and vertical positions.
- $\psi(R, Z)$: Poloidal flux function, crucial for defining magnetic surfaces.

Machine Constraints:

- Coil Current Limits: Operational limits of the magnetic coils.
- Coil & Passive Conductors Location: Geometric configuration crucial for accurate magnetic equilibrium and stability analysis.
- Voltage Limits
- Actuator Response Times
- Control gains and etc.…

Plasma Constraints:

- mutual inductance matrix $M_{ss}$and etc.
- sensitivity data that shows how small changes in the parameters (e.g. M_{SS}, R_S and etc..) affect the system's behavior.

Building a Robust Transolver Surrogate: Leveraging CMOD and SPARC Databases for Comprehensive Training

Plasma Current & Shape Parameters:

- I_c: Coil currents. R-Z grid
- q_A: Plasma current.
- β_p: Poloidal beta, indicating the plasma pressure.
- κ_{areal} : Internal inductance, describing the current distribution in the plasma.
- q_{95}: Safety factors, describing magnetic field line pitch.
- δ: Plasma triangularity

Pressure & Temperature Profiles:

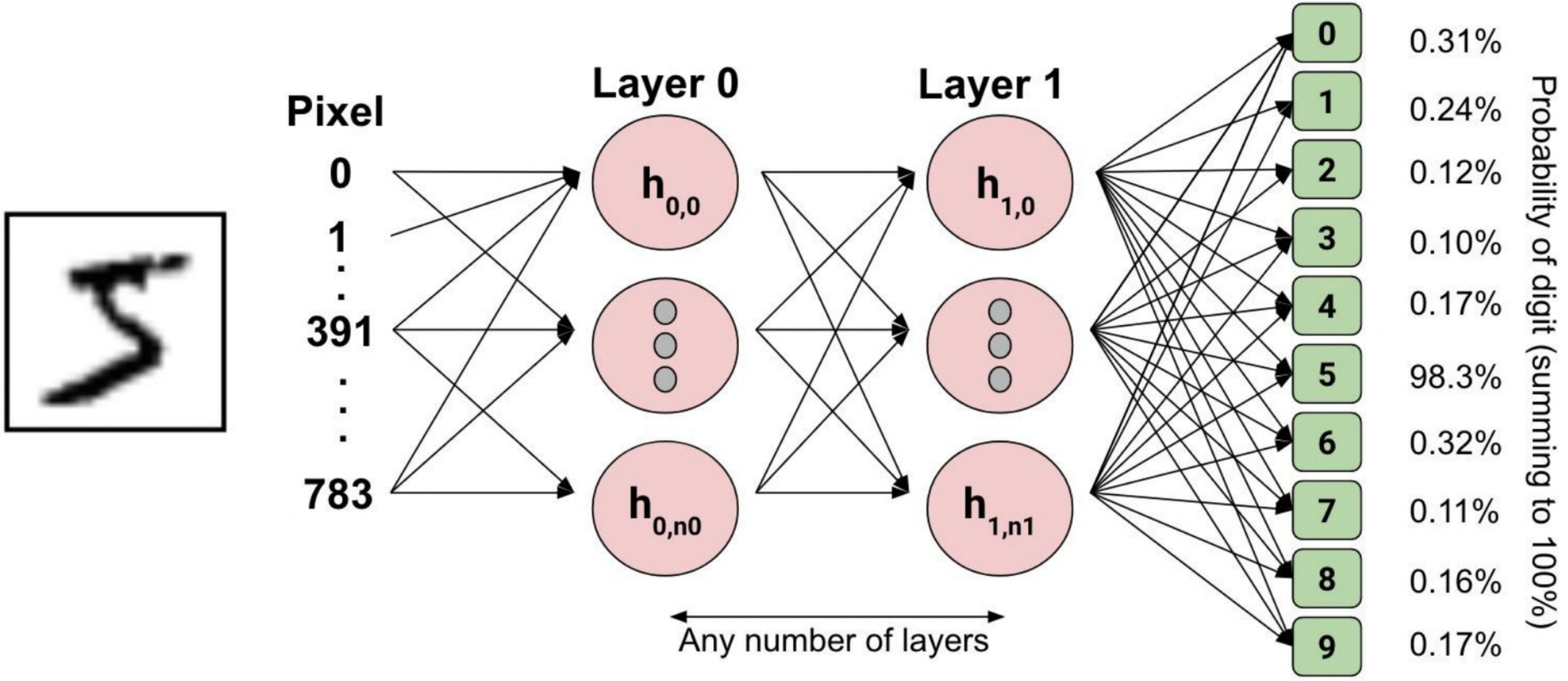
- P': Pressure gradient profile .
- FF': Temperature gradient, influencing plasma stability.
- $j_\phi(R, Z)$: Toroidal current density across radial and vertical positions.
- $\psi(R, Z)$: Poloidal flux function, crucial for defining magnetic surfaces.

Machine Constraints:

- Coil Current Limits: Operational limits of the magnetic coils.
- Coil & Passive Conductors Location: Geometric configuration crucial for accurate magnetic equilibrium and stability analysis.
- Voltage Limits
- Actuator Response Times
- Control gains and etc.…

Plasma Constraints:

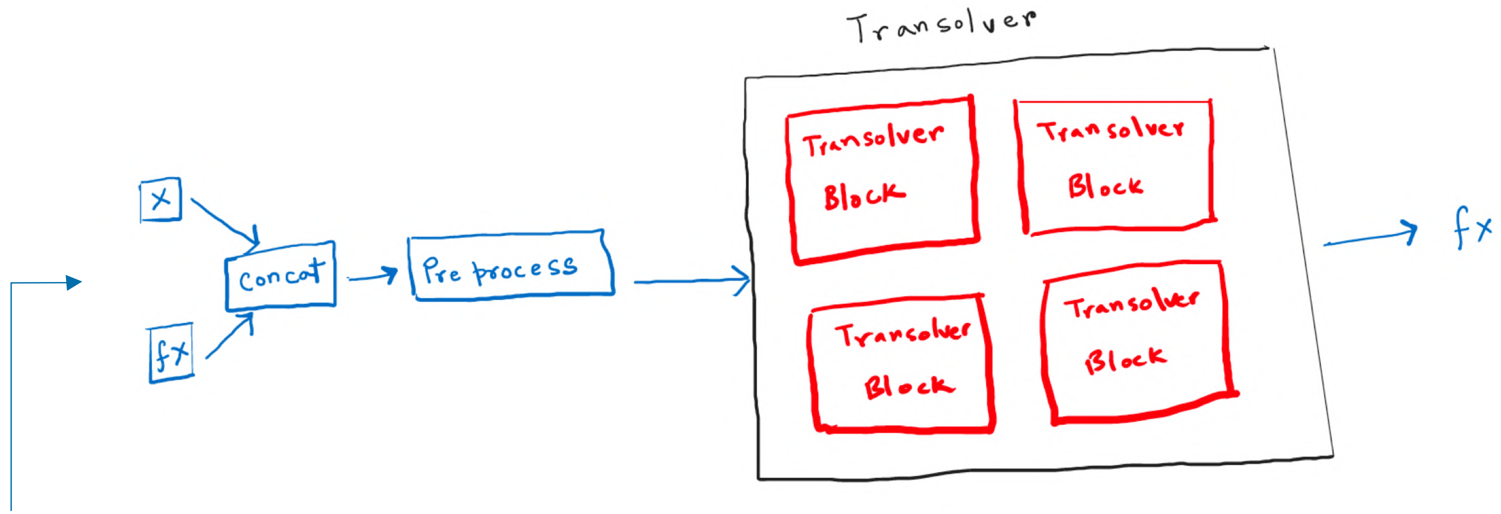
- mutual inductance matrix M_{ss} and etc.
- sensitivity data that shows how small changes in the parameters (e.g. M_{SS}, R_S and etc..) affect the system's behavior.



Pixel
0
1
391
783
Layer 0
h0,0
h0,n0
Layer 1
h1,0
h1,n1
Any number of layers
0
1
2
3
4
5
6
7
8
9
0.31%
0.24%
0.12%
0.10%
0.17%
98.3%
0.32%
0.11%
0.16%
0.17%
Probability of digit (summing to 100%)

AI Transformers for MEQ-FGE

Level 0



I_p ,I_c $\psi(R,Z)$, β_p, l_i , q_a , P', TT', γ_z

Inputs

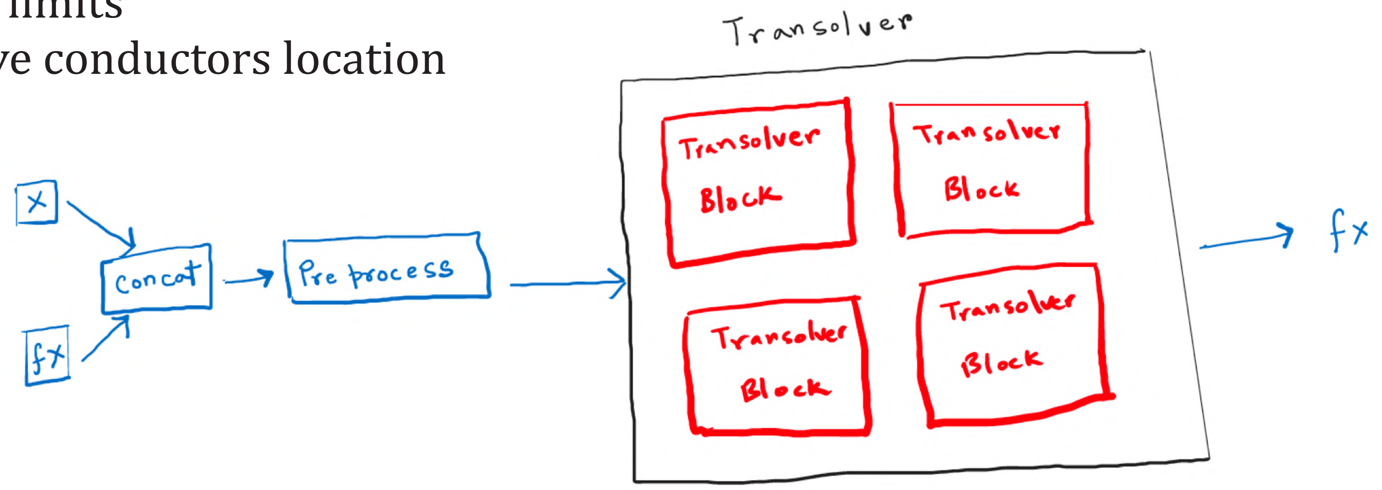
```
fx ← Preprocess ( Concat (x, fx))

for block in Transolver blocks:
    fx ← block (fx)

return fx
```

AI Transformers for MEQ-FGE

Level 0

Key Input Parameters

- $I_a, I_p, \beta_p, l_i, q_A, q_{95}, p', TT', \mathrm{j(R,Z)}, \psi(R,Z)$
- Coil current limits
- Coil & Passive conductors location



$I_p, I_c\ \psi(R,Z), \beta_p, l_i, q_a, P', TT', \gamma_z$

Inputs

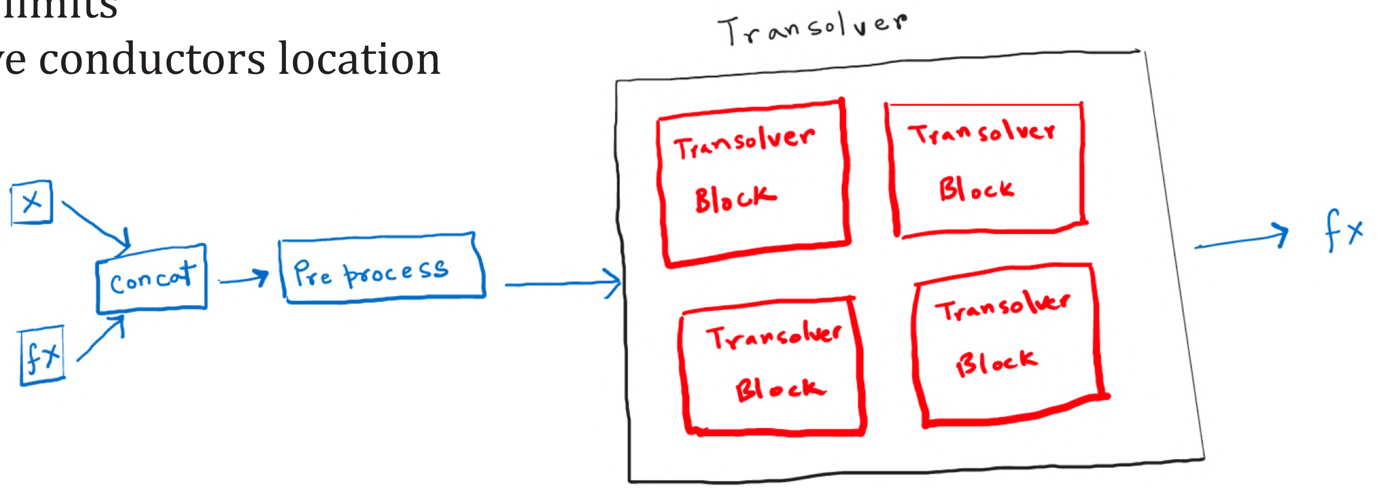
```
fx ← Preprocess ( Concat (x, fx))

for block in Transolver blocks:
    fx ← block (fx)

return fx
```

AI Transformers for MEQ-FGE

Level 0

Key Input Parameters

- $I_a, I_p, \beta_p, l_i, q_A, q_{95}, p', TT'$, j(R, Z), $\psi(R, Z)$
- Coil current limits
- Coil & Passive conductors location



$I_p, I_c\ \psi(R,Z), \beta_p, l_i, q_a, P', TT', \gamma_z$

Inputs

```
fx ← Preprocess ( Concat (x, fx))

for block in Transolver blocks:
    fx ← block (fx)

return fx
```

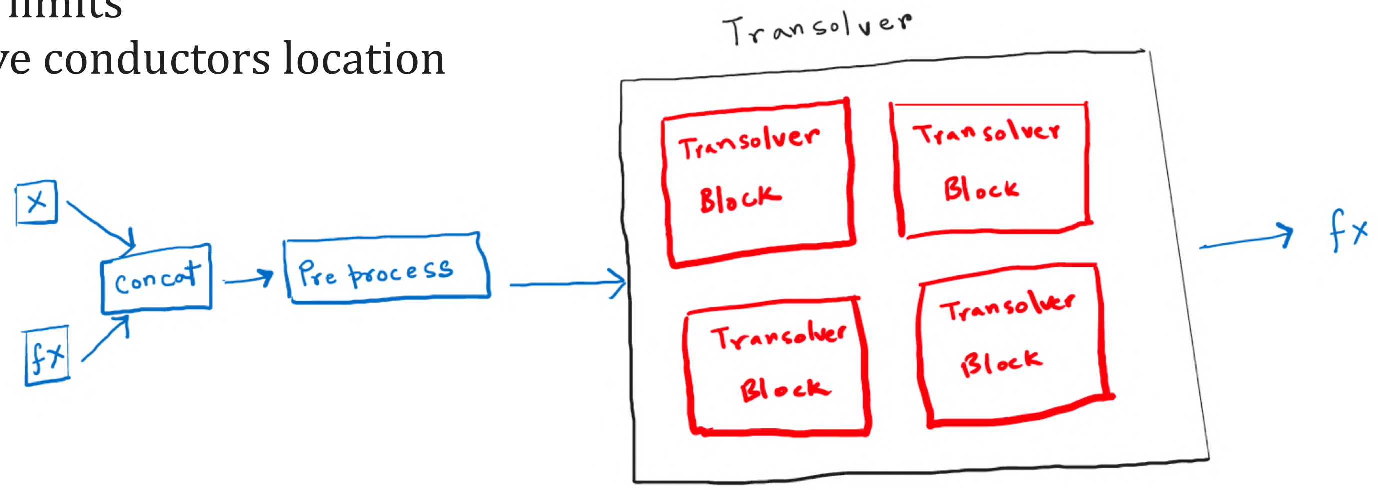
AI Transformers for MEQ-FGE

SPARC

Level 0

Key Input Parameters

- $I_a, I_p, \beta_p, l_i, q_A, q_{95}, p', TT', \mathrm{j(R,Z)}, \psi(R,Z)$
- Coil current limits
- Coil & Passive conductors location



$I_p, I_c\ \psi(R,Z), \beta_p, l_i, q_a, P', TT', \gamma_z$

Inputs

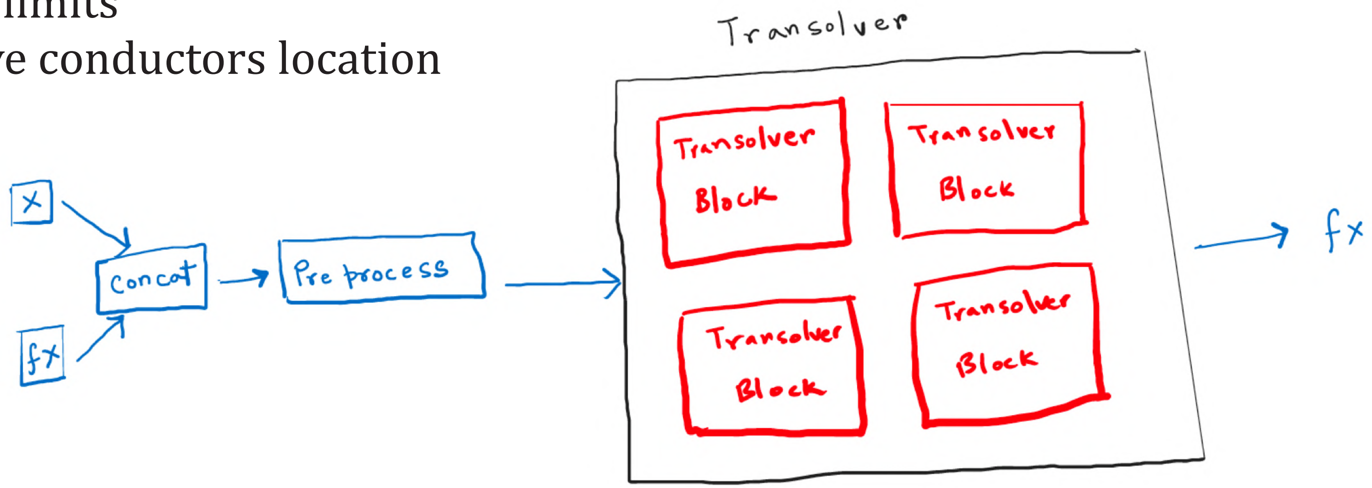
```
fx ← Preprocess ( Concat (x, fx))
for block in Transolver blocks:
    fx ← block (fx)
return fx
```

AI Transformers for MEQ-FGE

Level 0

Key Input Parameters

- $I_a, I_p, \beta_p, l_i, q_A, q_{95}, p', TT'$, j(R, Z), $\psi(R, Z)$
- Coil current limits
- Coil & Passive conductors location

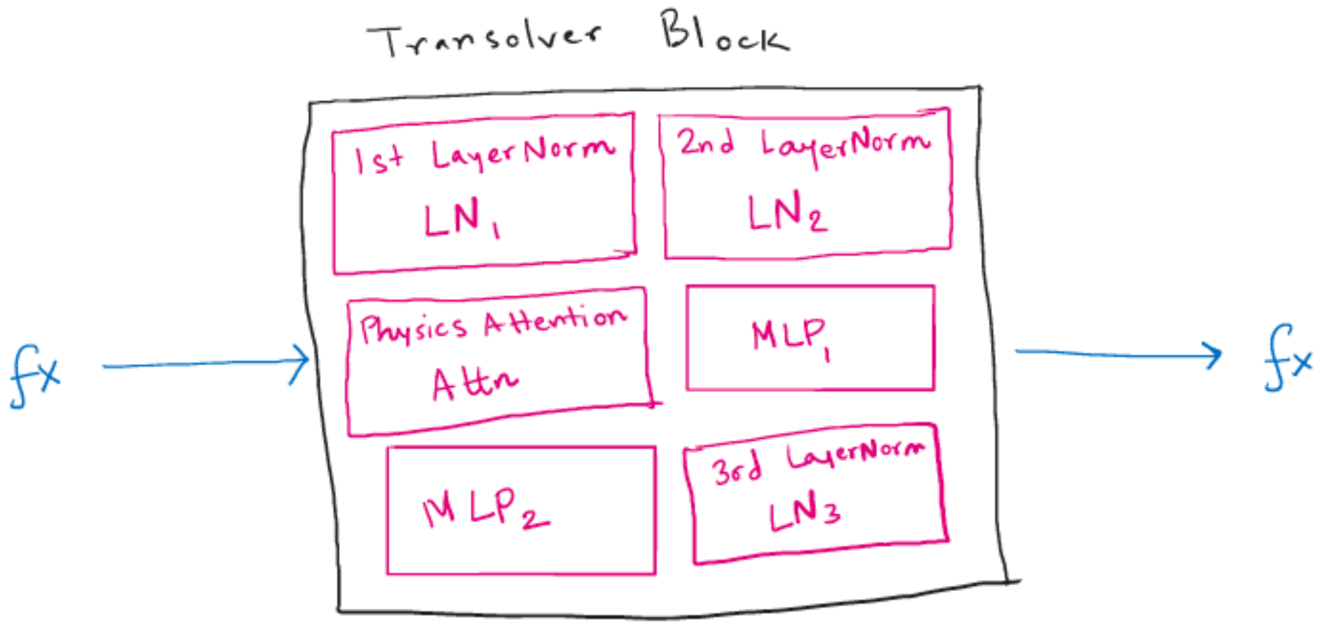


I_p, I_c $\psi(R, Z)$, β_p, l_i, q_a, P', TT', γ_z

Inputs

```
fx ← Preprocess ( Concat (x, fx))

for block in Transolver blocks:
    fx ← block (fx)

return fx
```

Level 1



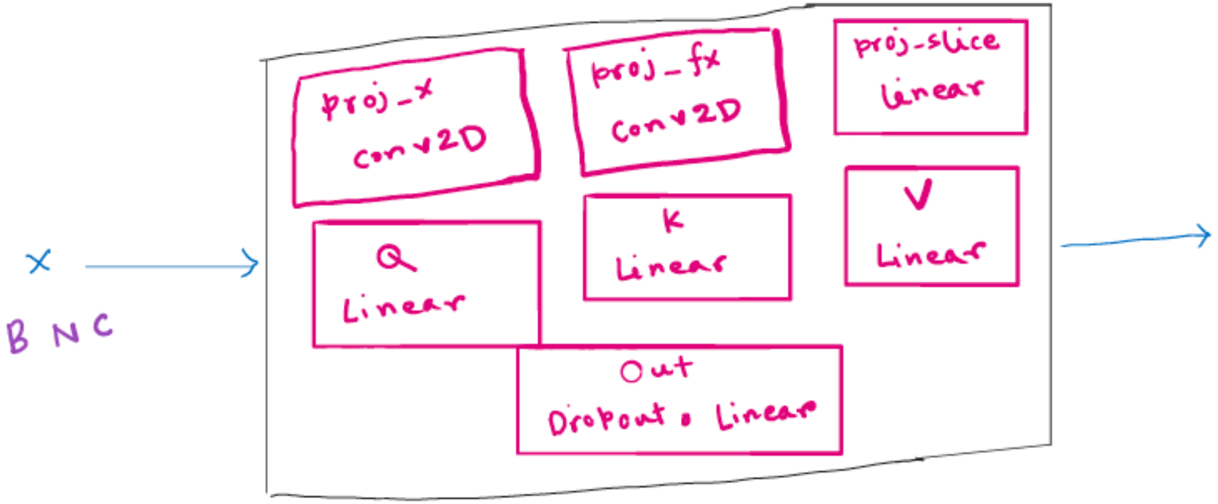
$$fx \mathrel{+}= \mathrm{Attn}(LN_1(fx))$$

$$fx \mathrel{+}= MLP_1(LN_2(fx))$$

$$\text{return } MLP_2(LN_3(fx))$$

Level - 2.3

Physics Attention



Softmax (Normalization function)

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

σ = softmax

$\vec{z}$ = input vector

e^{z_i} = standard exponential function for input vector

K = number of classes in the multi-class classifier

e^{z_j} = standard exponential function for output vector

e^{z_j} = standard exponential function for output vector

$\bar{fx}$ = proj_fx (x)   B H N C

$\bar{x}$ = proj_x (x)   B H N C

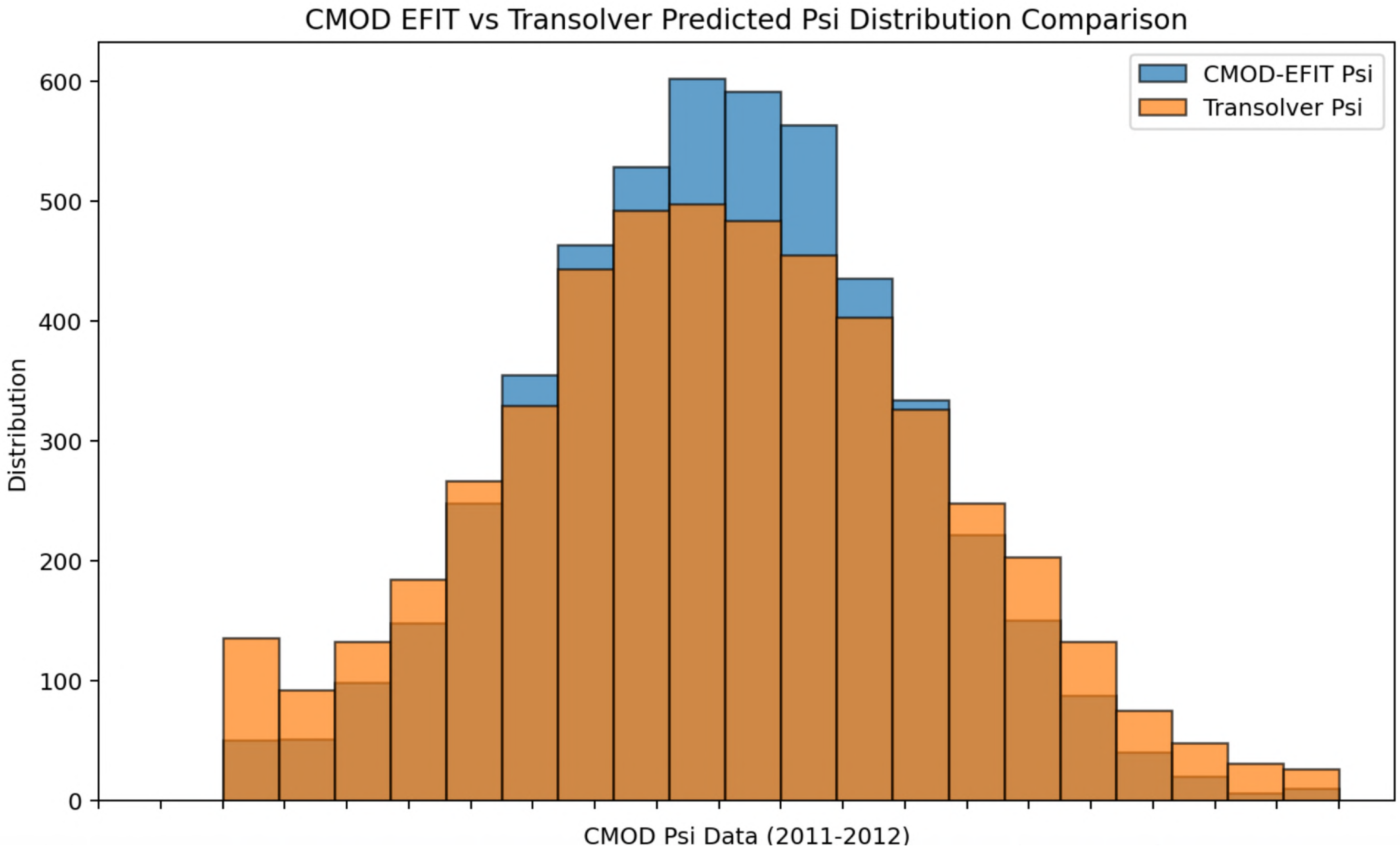
w = softmax (proj_slice ($\bar{x}$))   B H N G

token = Einsum ("bhnc, bhng → bhgc", $\bar{fx}$, w)   Slicing

norm = w.sum (2)   B H G

token /= norm

Flux distribution shows 10-20% difference between EFIT and Transolver predicted C-Mod scenarios



EFIT considered as a ground-truth