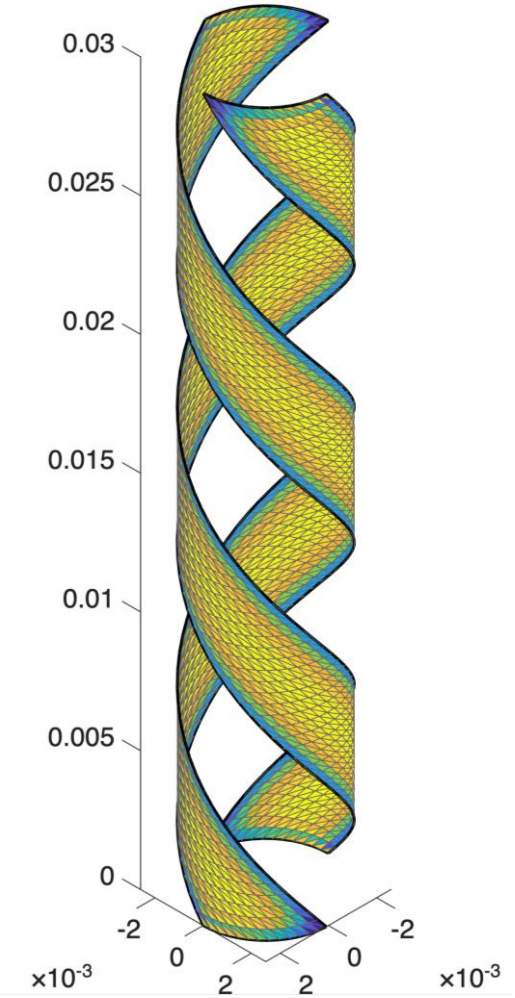
# Development of a Synthetic Database for CORC® Cables

**MAG-S.01.01-T030-D002**

Fabio Freschi
Laura Savoldi

*Department of Energy – Politecnico di Torino*



# Motivation & End-Product

## Why a synthetic database?

- Design-space is large: geometry, tape count/layers, material law parameters, $T$, external $B$ (magnitude + orientation).
- CORC® physics is coupled: self-field + anisotropic $J_c(B_n, B_t, T)$ → non-uniform current sharing.
- Need fast "engineering outputs" for magnet design loops → train metamodels (ANN/surrogates).

## What is stored per case?

- Raw $E-I$ samples around the transition (few points).
- Identified pair $(Ic, n)$ from log-domain regression.
- Tape-wise / segment-wise current distribution (optional feature targets).
- Full metadata: geometry inputs, $J_c$ model + parameters, $B_{ext}$ description, temperature, random seed, convergence + fit diagnostics.

## ML-ready outputs (what the metamodel predicts)

$I_c(B_n, B_t, T, geometry)$

$n(B_n, B_t, T, geometry)$

$E - I$ curve features

Current sharing metrics

→ **A metamodel is only as good as the dataset**: Inputs must encode the same physics used by the solver: geometry, material law, and operating field.
→ trained surrogate enables rapid scans / optimization loops with uncertainty bands

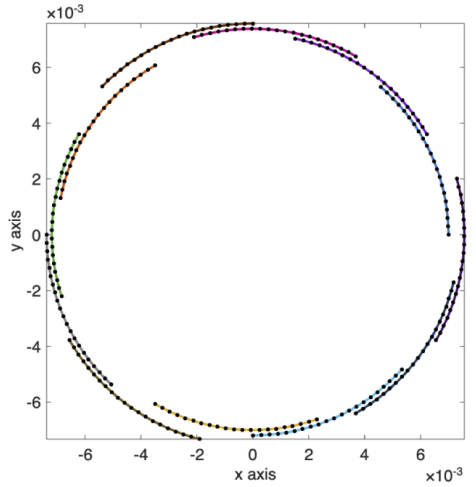
# Simulation engines used to generate the synthetic database

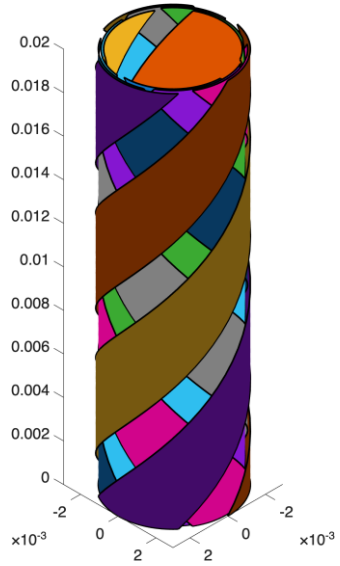
## STAR2d (2-D DC, cross-section)

- Unknowns: segment currents + one shared cable voltage.
- Self-field via Biot–Savart (no air mesh).
- Best for: fast parameter sweeps, sensitivity studies, dataset scale-out.

## STAR3d (3-D stream-function on tape surfaces)

- Captures axial/helical effects (pitch), terminations, non-uniform field along axis.
- More expensive; used selectively as high-fidelity anchors.
- Voltage inferred from power loss; includes automatic outlier rejection in fit.



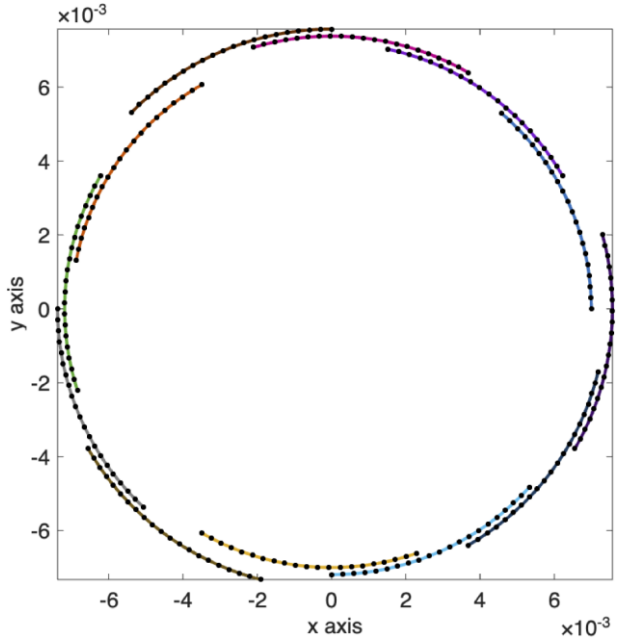




4 layers – 3 tapes per layer

# Step 1 — Geometry parameterization and discretization (STAR2d)

## Geometry pre-processor inputs

- Tape width w, SC thickness (SC-only in DC model).
- No. layers, No. tapes per layer, layer-to-layer spacing.
- Inner/core radius (first layer).
- Target segment size → discretization density.
- Feasibility check: first-layer circumference must fit tapes




## Case instantiation

Geometry params → Discretize tapes → Cache Biot-Savart kernels → Attach material law $J_c(B_n, B_t, T)$

→ fast re-use across current sweeps and parameter scans

# Step 2 — Material law and nonlinear solution

## Constitutive relation (per segment)

- Power-law $E-J$ with criterion $E_c$ and exponent $n$.
- Nonlinear resistivity $\rho(J, B_n, B_t, T)$ derived from the power-law.

$$\rho(B_n, B_t, J, T) = \frac{E_c}{J}\left(\frac{J}{J_c(B_n, B_t, T)}\right)^n$$

- $J_c$ model options: critical-state, isotropic Kim, anisotropic Kim-like, magneto-angular fits.

## Nonlinear system solved

- Unknowns: segment currents $I_k$ + cable voltage V.
- Constraints: $\sum_k I_k = I_{tot}$ and voltage equalization $V_k = V$.
- Self-field update: Biot–Savart superposition each iteration.
- Solver: MATLAB **`fsolve`** (trust-region) with variable scaling.

## Numerical robustness

- **Continuation**: reuse previous current-point solution as initial guess
- **Mesh adaptivity**: refine until ΔV between refinements < 5% (fixed per geometry)
- **Cache geometry kernels** (fixed layout) → cheap iterations

→ stable E–I samples over many decades of E (critical for log-fit)

# Step 3 — Identify $(I_c, n)$ from a small $E$– $I$ sample set

## Key choices for a scalable database

- Use few current points near transition.
- Guarantee bracket around $E_c$ before fitting.
- Fit in log-domain to avoid dominance of high-$E$ points.

## Why log-domain fitting?

- $E$ spans many decades: log-fit treats multiplicative error uniformly.
- Prevents high-$E$ points from biasing the regression upward.
- Window kept tight around $E_c$ so slow $J_c$ variation does not introduce curvature.

Database implication: few solves → reliable labels → massive coverage

## Identification algorithm

Provisional $I_c \approx J_c(B_{ext}) \cdot A_{tot}$

↓

Solve at currents around $I_c$ (continuation)

↓

Check bracket: $E(I_k) < E_c < E(I_{k+1})$

↓

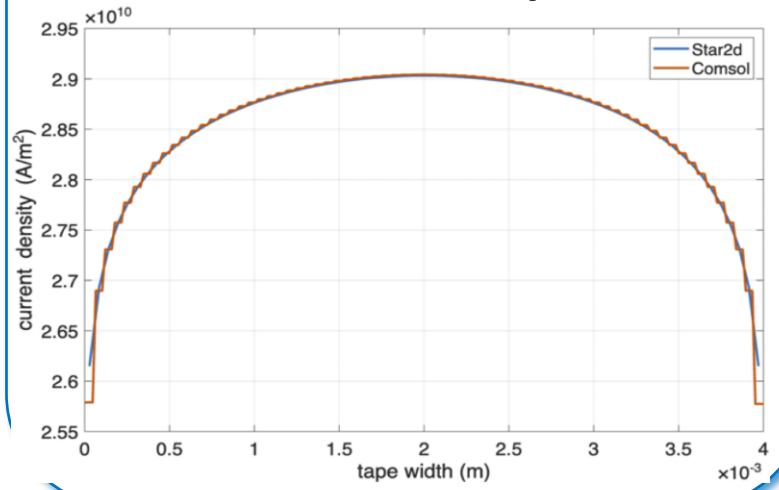
Log–log linear regression near crossing

↓

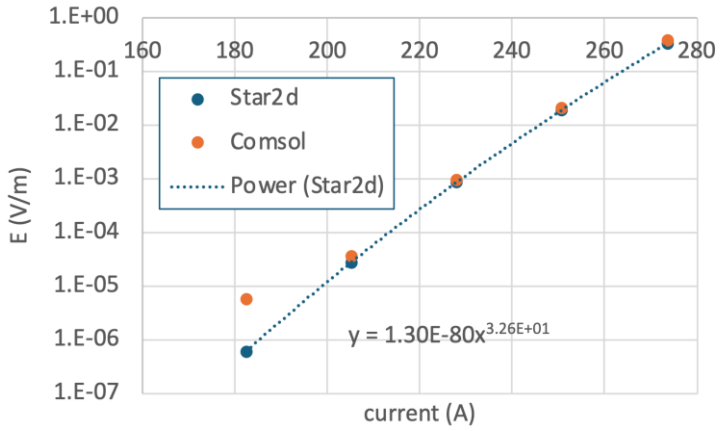
Return: $n$ (slope), $I_c$ from $E = E_c$

# Cross-check vs COMSOL

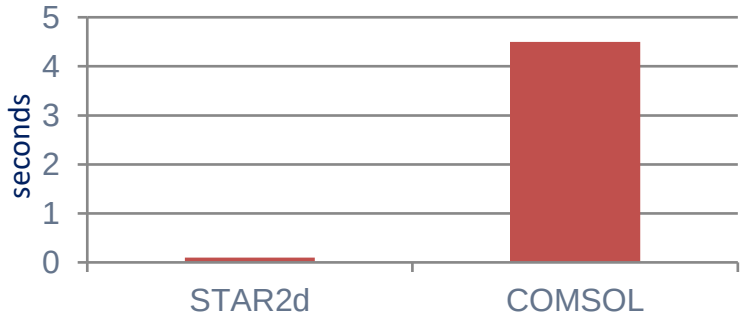
## Local J profile across tape width matches closely




## $I_c$ and $E-I$ behavior consistent near transition

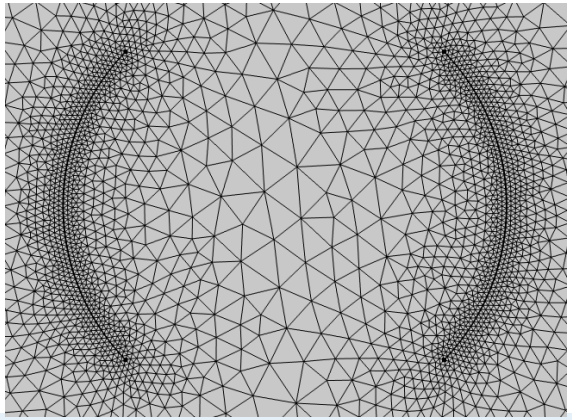



## Faster solution time




## COMSOL

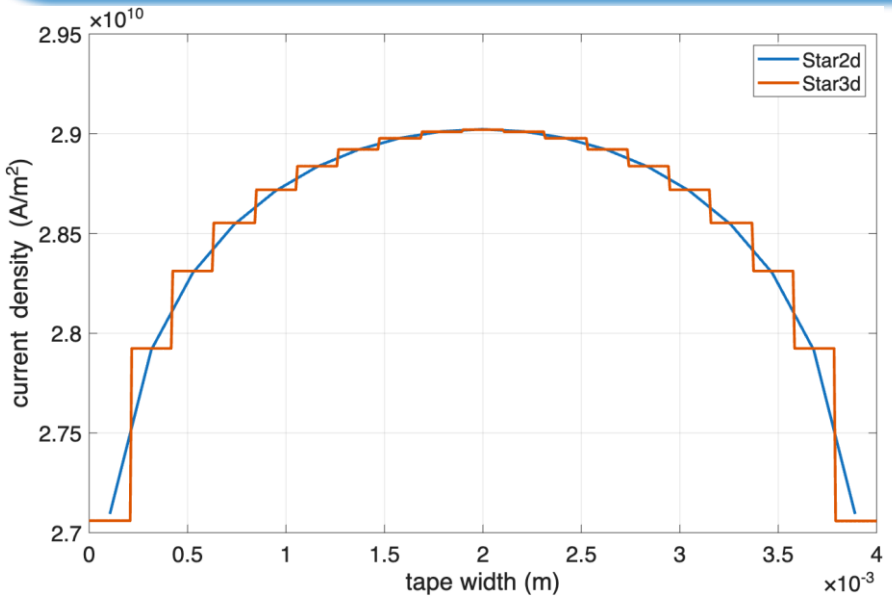
- $T-A$ formulation.
- Discretization of tapes and surrounding air.
- Each tape is discretized with 71 elements, whereas the surrounding air is discretized with 5556 triangular elements for a total of 11305 degrees of freedom.

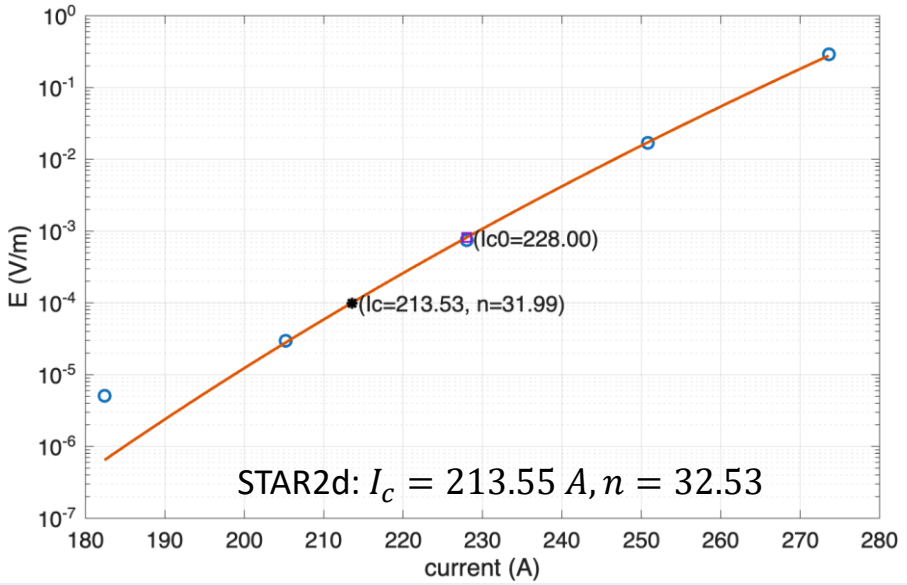


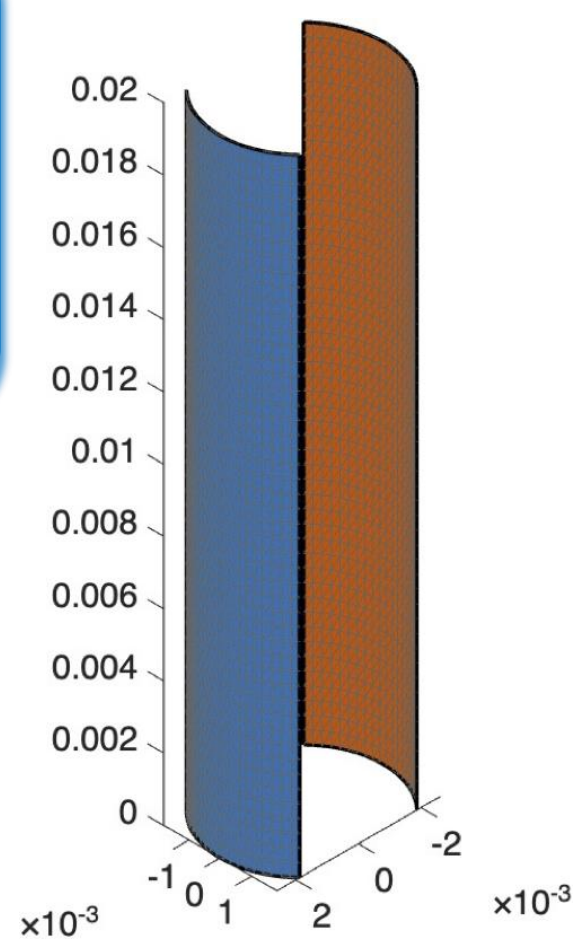
# 3-D extension: STAR3d

**STAR3d**

- ✓ Representation of a thin-surface
- ✓ No air mesh
- ✓ Nodal-based unknowns
- ✓ The calculation kernels for B are cached
- ✓ Connections (inlet/outlet sections) effect
- ✗ Computational expensive but necessary: “true results”



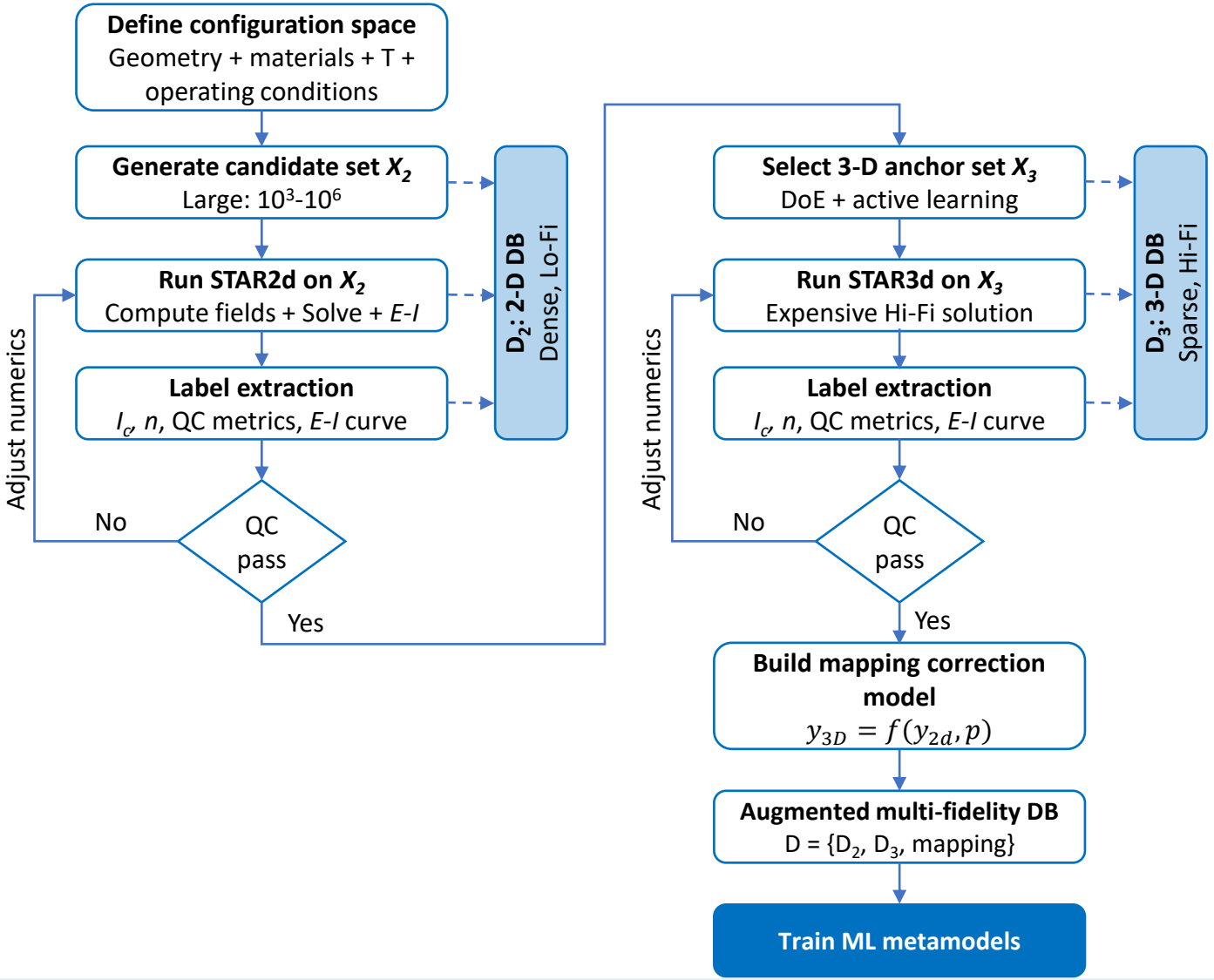



# 2-D/3-D interaction and development of the database

## Role of 2-D simulations

- Essential environment for physics and numerics exploration
- Intermediate training database (parameter sensitivity, extrapolation
- 2-D/3-D combination through ML-based correction model



# Takeaways

**Development of Synthetic Database**

- Physics-consistent features (geometry + $B_n/B_t$ + material law) → metamodels that generalize configurations and operating conditions.
- Robust label extraction: store raw $E-I$ + fitted $(I_c, n)$ + residuals.
- Scale via Star2d; control bias via Star3d anchors (multi-fidelity, residual correction).
- Database records include provenance + QC flags → training is repeatable and defensible.