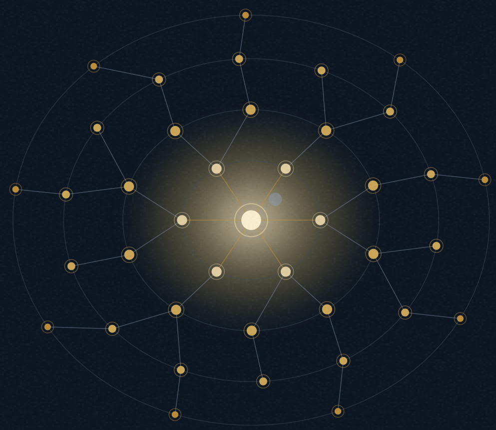
KRONOS FUSION ENERGY



PAPER 2.83 · THE KRONOS 2026 SERIES

# MLOps for a Safety-Critical Fusion Control Stack: Versioned Lineage, Gated Model Promotion, and Physics-Balanced Multi-GPU Training for a Compact Spherical-Tokamak Breeder

A machine-learning model that touches a nuclear plant must earn its authority the way any safety-critical software does: by passing an explicit, reproducible gate before it acts....

Read it live — 3-D model · physics validator · film series

BY P I FORD

2026 · EDITORIAL EDITION

THE OFFICIAL RECORD

# What this document is, and how to read its numbers

This is a combined editorial. It gathers, in one volume, the two design studies Kronos Fusion Energy released in 2026 — a compact spherical-tokamak tritium/helium-3 breeder and a deuterium–helium-3 tandem-mirror generator — together with the intellectual-property estate, the people, and, in the internal edition, the full commercial case. The scientific text is the same text that appears in the arXiv preprints and the journal submissions; the editorial adds the apparatus a reader needs to hold the whole program at once, and nothing in the physics is altered to fit it.

## THE HONESTY STANCE IS THE ASSET

Every headline quantity carries an **evidence class** — DERIVED RECOMPUTED MEASURED REQUIREMENT — and, where a number is a modelled extrapolation rather than a demonstrated value, it is labelled as such and its downside is shown next to it. Several quantities in the underlying corpus moved after first publication; where a value has been withdrawn or restated, the record says so plainly. This is deliberate: a diligence team will find the load-bearing assumptions, so the document surfaces them first.

## TWO EDITIONS

This volume is issued in two editions. The **public edition** (light) carries the physics and the design only — no dollar, price, or per-kilogram figure appears anywhere in it. The **internal edition** (dark) adds the economics, the funding architecture, and the defense-commercial analysis, and is marked *Confidential* accordingly; it is not for public distribution. Where the commercial case rests on a modelled assumption rather than a demonstrated one, it is labelled as such and its downside is shown alongside it, on the same terms as the physics.

| | |
|---|---|
| TITLE | The Kronos Fleet — Hyperion · Aegis · MetroVolt: A Combined Design & Commercialization Study |
| AUTHOR | Priyanca Ford, Founder & CEO, on behalf of Kronos Fusion Energy |
| EDITION | 2026 Combined Editorial · first issue |
| COMPANION WORKS | The five-paper arXiv drop — (1) Breeder, (2) Burner, (3) REBCO magnet & tape, (4) Direct Energy Conversion, (5) AI/ML/Quantum control — with matching numbered Zenodo deposits, plus the IOP journal and IAEA-FEC submissions. Patents were filed before the drop. |
| NAMING | Hyperion = the breeder; Aegis (defense) and MetroVolt (data-center) = one generator in two housings. Internal mode letters (D1, L) do not appear on public surfaces. |
| STATUS | Conceptual design & simulation study. Nothing herein is a construction commitment. |

HOW TO READ THIS VOLUME

# The fleet at a glance, and the way its numbers are marked

This volume opens with *Orientation*, presents the five research papers as a bound sequence, and closes with the *Apparatus*. *Foundations* sets out the three general results both machines rest on. *Paper One* is the Hyperion breeder; *Paper Two* the Aegis / MetroVolt generator; *Papers Three, Four and Five* are the enabling science — high-field REBCO magnets and tape, direct energy conversion, and the AI / ML / quantum control stack. A *Digital Twin & 3-D Model* part follows, then the *Environmental* profile. The *Economics* and levelized cost and the *Business Case* and diligence record appear in the internal edition only; the *Simulations* proof record and the Apparatus — patent portfolio, team, limitations, and sources — close both editions.

## THE FLEET AT A GLANCE

| | HYPERION | AEGIS | METROVOLT |
|---|---|---|---|
| **Role** | Strategic-isotope foundry | Defense installation power | Campus power |
| **Machine** | **Spherical tokamak** | **Tandem mirror** | **Tandem mirror** |
| **Fuel** | Deuterium–tritium | Deuterium–helium-3 | Deuterium–helium-3 |
| **Delivers** | Tritium · $^{3}$He · 14 MeV n | Resilient installation power | Firm campus power |
| **Physics bar** | **Fusion gain only** | Closure (gates named) | Closure + lunar $^{3}$He |
| **In the fleet** | Breeds the fuel | Proves the generator | Commercial destination |

## HOW THE NUMBERS ARE MARKED

Every headline quantity carries an evidence class — DERIVED from first principles, RECOMPUTED from raw data, MEASURED in hardware, or a REQUIREMENT yet to be met. Financial figures carry a case label and are confined to the internal (confidential) edition. Values that moved after first publication are shown as withdrawn or restated rather than quietly changed. A capability you can evaluate is one whose limits are on the page.

ABSTRACT

# MLOps for a Safety-Critical Fusion Control Stack: Versioned Lineage, Gated Model Promotion, and Physics-Balanced Multi-GPU Training for a Compact Spherical-Tokamak Breeder

A machine-learning model that touches a nuclear plant must earn its authority the way any safety-critical software does: by passing an explicit, reproducible gate before it acts. We formalise the machine-learning-operations (MLOps) layer of the Kronos control stack for the Hyperion compact spherical-tokamak breeder (design point $Q = 3.076$, $P_{\text{fus}} = 85.04$ MW, $I_p = 9.66$ MA, negative triangularity $\delta = -0.30$, on-axis field $B_0 = 8$ T, elongation $\kappa = 2.0$, major radius $R_0 = 1.20$ m, aspect ratio $A = 2.5$, centrepost peak field $16.84$ T) as a single organising principle: *no model reaches control authority without reproducing the frozen physics anchors, bit-stably*. We cast model promotion as a constrained decision problem, writing the frozen-anchor gate as a conjunction of pre-registered inequalities over an anchor residual map, deriving the physics-balanced training objective and its adaptive loss-weighting condition that prevents residual-term collapse, and stating the data-parallel multi-GPU stochastic-optimisation update, the content-addressed lineage relation, and the distribution-shift drift test as explicit equations. We give the numerical formulation — the physics-informed Grad–Shafranov collocation discretisation, the manufactured-solution/grid-convergence verification with its observed-order and grid-convergence-index estimators, the Gaussian-process Bayesian-optimisation and approximate-Bayesian calibration solvers, and the Sobol/Monte-Carlo sensitivity estimator — with convergence criteria. The gate is built and exercised on the NVIDIA GPU high-performance-computing (HPC) campaign: the Kronos reduced-order engine reproduces **19 of 19** frozen anchors and an independent formal verification-and-validation suite clears the same **19/19**, both bit-stably across three independent environments. The components that feed the gate are demonstrated to their own acceptance: a Bayesian hyperparameter-optimisation loop returns a calibrated confinement posterior $H_{98} = 1.007 \pm 0.030$ over $200$ evaluations, a physics-loss-balanced equilibrium surrogate converges to RMS $\psi$ error $0.139\%$ against a free-boundary solver at a $2877\times$ speedup, and a learned ensemble estimator reaches an area under the receiver-operating characteristic of $0.980$. The result is a promotion pipeline in which lineage is complete end-to-end, every result is reproducible from a manifest, and a model that fails the anchor gate cannot ship — the operational foundation for trustworthy learned control of a nuclear machine, with on-plant integration built to this demonstrated gate on a staged schedule.

Open-access deposit (code & data, CC BY 4.0): DOI 10.5281/zenodo.22645827 · Zenodo community *kronos_fusion_energy*. Explore it live: interactive 3-D model · physics validator.

**Keywords:** MLOps · reproducibility · frozen-anchor regression gate · model promotion · physics-informed training · multi-GPU · calibrated uncertainty · digital twin · spherical-tokamak breeder

CONTENTS

# Table of Contents

MLOps for a Safety-Critical Fusion Control Stack: Versioned Lineage, Gated Model Promotion, and Physics-Balanced Multi-GPU Training for a Compact Spherical-Tokamak Breeder 3
Introduction 6
Governing equations and the formal promotion problem 8
Numerical formulation 11
Computational methodology: the NVIDIA GPU HPC campaign 13
Results 14
Discussion 21
Limitations 22
Conclusion 23
References 28
One programme, many papers 33

NOMENCLATURE

# Symbols, units, and the names of things

| | |
|---|---|
| $Q$ | plasma (fusion) gain, $P_{fus}/P_{aux}$ |
| $Q_E$ | engineering gain, net electric out / recirculating in |
| $H_{98}$ | confinement enhancement over the IPB98(y,2) scaling |
| $Z_{eff}$ | effective ion charge |
| $c_{ee}$ | electron–electron bremsstrahlung leading coefficient, 2.120022 (exact) |
| $\tau_E$ | energy confinement time |
| $I_p$ | plasma current |
| $R_0$, $a$ | major radius, minor radius |
| $\kappa$, $\delta$ | elongation, triangularity |
| $\beta_N$ | normalized plasma beta (Troyon-normalized) |
| *TBR* | tritium breeding ratio |
| *DEC* | direct energy conversion |
| *CF* | plant capacity factor (availability) |
| *FOAK / NOAK / BOAK* | first / n$^{th}$ / best of a kind |
| *Hyperion* | the ST breeder (internal: Mode D2) |
| *Aegis* | the generator, defense/installation housing (internal: Mode M) |
| *MetroVolt* | the generator, data-center housing (Mode M) |

§ 1

# Introduction

## LEARNED MODELS WITH REAL AUTHORITY

Autonomous operation of a compact, high-power-density spherical tokamak (ST) leaves little time to react and no tolerance for an unmodelled excursion. The Hyperion breeder holds a plasma current $I_p = 9.66$ MA at a major radius of only $R_0 = 1.20$ m behind a centrepost magnet at a peak on-conductor field of $16.84$ T, in a negative-triangularity ($\delta = -0.30$) regime chosen for its quiet, edge-localised-mode-free edge [1,2]. Holding a plasma this dense inside a narrow operating window at kilohertz electromagnetic and vertical-stability timescales is a control problem, and modern fusion control increasingly hands parts of that problem to learned models: surrogate operators run a predictive shadow twin, graph and physics-informed networks reconstruct the state, and optimisers propose actuator trajectories [5,6,7]. Each of these carries authority: a surrogate that is wrong by a few per cent can steer the plant off its operating point, and in a nuclear system every actuation carries a safety consequence.

The central question for such a system is therefore not whether a given model scores well on a held-out benchmark, but whether the *process* that builds, tests, promotes and retires it is trustworthy. That process is machine-learning operations (MLOps). In conventional machine-learning products MLOps is a matter of engineering hygiene [18,19,20]; in a safety-critical plant it becomes a safety property in its own right, and it must do three things without exception: block a bad model before it ships, reproduce any result from its manifest, and trace every value to the sensor it came from.

## PRIOR ART

Three literatures meet in this paper. The first is fusion control: real-time equilibrium reconstruction [3], physics-model-based profile control inside the cycle [4], and, more recently, learned controllers — Degrave *et al.* trained a reinforcement-learning agent to command TCV coil voltages directly from magnetics, including negative-triangularity shapes [5]; Seo *et al.* learned to avoid tearing modes on DIII-D [6]; and deep networks forecast disruptions across machines [7]. The control quality of all of these is bounded by the speed and fidelity of the plasma model beneath them, which is why fast learned surrogates — neural operators [10,11], physics-informed networks [12], and quasi-linear transport emulators three-to-five orders of magnitude faster than the code they replace [8,9] — have become central.

The second literature is the software engineering of machine-learning systems. Sculley *et al.* catalogued the “hidden technical debt” that accrues when models are shipped without the surrounding infrastructure [18]; Amershi *et al.* documented the end-to-end ML lifecycle as an engineering discipline [19]; Breck *et al.* proposed a production-readiness rubric [20]; and model cards [21] and datasheets [22] made provenance and intended use first-class artifacts. In parallel, distributed-training practice matured — large-minibatch synchronous stochastic gradient descent (SGD) with learning-rate scaling [23], mixed-precision arithmetic [24], and ring all-reduce collectives [25] — as did calibration [30], deep-ensemble uncertainty [31], knowledge distillation [32], and Gaussian-process Bayesian optimisation for hyperparameters [27,28,29].

The third literature is verification, validation and uncertainty quantification (VVUQ) for computational science: solution verification by the method of manufactured solutions and grid-convergence indices [33,34,35], and global sensitivity by Sobol indices [36]. Each of these three literatures is mature on its own. What is missing — and what a compact nuclear ST forces — is their fusion into a single promotion discipline in which the physics VVUQ result is the gate that the machine-learning pipeline must clear, the training that produces gate-passing models respects the governing physics, and the whole chain is reproducible bit-for-bit from a manifest.

## CONTRIBUTION

This paper contributes that discipline, stated formally and demonstrated on the campaign. Specifically:

We cast **model promotion** as a constrained decision problem (§2) and write the *frozen-anchor gate* as a conjunction of pre-registered inequalities over an anchor-residual map (equations 1–3), so that admission to control authority is a machine-checkable predicate rather than a judgement.
We derive the **physics-balanced training objective** and the adaptive loss-weighting rule that keeps the data and partial-differential-equation (PDE) residual terms co-convergent (equations 5–6), and state the **data-parallel multi-GPU** synchronous-SGD update, learning-rate scaling and determinism condition that make training bit-reproducible (equations 7–9).
We formalise **lineage** as a content-addressed directed acyclic graph and **reproducibility** as bit-stability of the anchor map across environments (equations 11–12), and cast **drift, canary and rollback** as a sequential decision (equations 15–16).
We give the **numerical formulation** (§3) — collocation discretisation of the physics-informed equilibrium operator, manufactured-solution/grid-convergence verification with observed-order and grid-convergence-index estimators, the Gaussian-process/expected-improvement Bayesian-optimisation solver and the approximate-Bayesian calibration recursion, and the Sobol/Monte-Carlo sensitivity estimator — each with a convergence criterion.
We describe the **NVIDIA GPU HPC methodology** (§4) that produced the results, naming the codes run and the verification harness, and we report the **results** (§5): 19/19 anchors reproduced bit-stably across three environments, and every gate-feeding component demonstrated to its own acceptance. enumerate

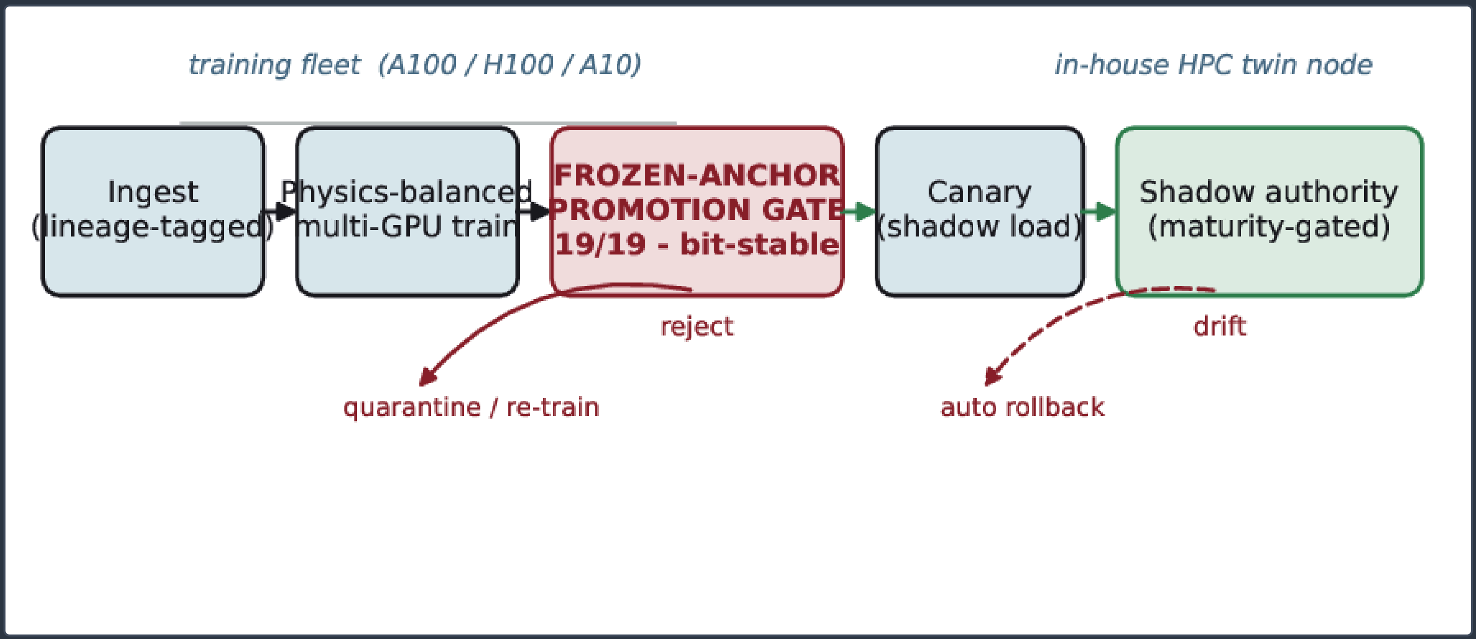
Every quantitative claim traces to a frozen result in the Kronos Physics De-Risking Register [37]; the serial identifiers (e.g. KX-L1-A10) are given inline so each number is auditable. The architecture composes with the certified control-barrier safety filter [41] and the maturity-gated authority policy [40] that sit beneath the learned layer, and it supplies the promotion discipline for the neural flux [42] and equilibrium-reconstruction [43] surrogates, the calibrated-uncertainty layer [44], and the four coupled digital twins [46] it feeds.

§ 2

# Governing equations and the formal promotion problem

## THE PROMOTION PIPELINE

The MLOps layer is a directed pipeline with a hard gate in the middle (figure 1). Telemetry is ingested with lineage; a candidate model $\theta$ is trained on the NVIDIA GPU HPC campaign; before it is admitted anywhere near shadow authority it must reproduce the frozen physics anchors; and only then does it enter a canary release under continuous drift surveillance, rising through graded authority as it keeps passing. The rest of this section states each stage as equations.




**(Schematic.)** The model-promotion pipeline. A candidate is trained with physics-balanced multi-GPU optimisation and must clear the frozen-anchor promotion gate — 19/19 anchors reproduced, bit-stable across three environments — before canary and promotion to maturity-gated shadow authority. A candidate that shifts any anchor is quarantined and re-trained; a served model that drifts is rolled back automatically. The gate is the load-bearing element. No data values are shown.

## THE FROZEN-ANCHOR MANIFOLD AND THE PROMOTION PREDICATE

Let the plant design be pinned to a vector of *frozen anchors*

$$a^\star = \left(a_1^\star, \dots, a_{n_a}^\star\right) \in \mathbb{R}^{n_a}, \qquad n_a = 19,$$

the canonical operating point and its verified derivatives — for the breeder $Q$, $P_{\text{fus}}$, $I_p$, the tritium throughput and the neutron-power fraction; for the burner $Q_E$, neutron power, neutron fraction and plug potential; the surplus law; the local and net tritium-breeding ratios; and the high-fidelity bridge checks (KX-L1-A6). A candidate model $\theta$ exposes an *anchor map* $g(\theta) \in \mathbb{R}^{n_a}$: the value it predicts for each anchor when the reduced-order engine is driven at the canonical inputs. Define the componentwise relative residual

$$r_i(\theta) = \frac{\left|g_i(\theta) - a_i^\star\right|}{\max\left(|a_i^\star|,\, a_{\text{floor}}\right)}, \qquad i = 1, \dots, n_a,$$

with $a_{\text{floor}}$ a small scale guarding near-zero anchors. The *frozen-anchor gate* is the pre-registered predicate

$$G_{\text{anchor}}(\theta) \;=\; \bigwedge_{i=1}^{n_a} \left[ r_i(\theta) \le \tau_i \right],$$

a conjunction of $n_a$ tolerance tests; a single anchor beyond its tolerance $\tau_i$ falsifies the conjunction and rejects the model. Promotion is admitted only when the gate holds jointly with the auxiliary acceptance predicates of §2.6–§2.7,

$$\mathit{Promote}(\theta) \iff G_{\mathrm{anchor}}(\theta) \wedge G_{\mathrm{vv}}(\theta) \wedge G_{\mathrm{repro}}(\theta) \wedge G_{\mathrm{cal}}(\theta),$$

where $G_{\mathrm{vv}}$ is the formal-verification predicate (§3), $G_{\mathrm{repro}}$ the bit-stability predicate (§2.5), and $G_{\mathrm{cal}}$ the calibration predicate (§2.6). Equation (4) is the machine-checkable form of the phrase "no model ships without reproducing the anchors": the anchors are not a target the optimiser is asked to approach but a *barrier* it must not move.

## PHYSICS-BALANCED TRAINING OBJECTIVE

The models that must clear the gate are trained so that the physics is a term in the loss, not an afterthought. Writing $\mathcal{N}[\psi] = 0$ for the governing operator of the quantity a surrogate learns — for the equilibrium operator, the Grad–Shafranov residual of equation (17) below — the physics-informed loss is the weighted sum of a data term, a PDE-residual term evaluated at $N_c$ collocation points, and a boundary/initial-condition term [12],

$$\mathcal{L}(\theta) = \lambda_{\mathrm{d}}\,\mathcal{L}_{\mathrm{data}}(\theta) + \lambda_{\mathrm{r}}\,\frac{1}{N_c}\sum_{k=1}^{N_c}\big\|\mathcal{N}[\psi_\theta](x_k)\big\|^2 + \lambda_{\mathrm{b}}\,\mathcal{L}_{\mathrm{bc}}(\theta).$$

The pathology this objective is prone to is *term collapse*: if the weights $\lambda$ are fixed, one term's gradient can dominate and the optimiser drives that term to zero while another — often the PDE residual — stalls, so the surrogate fits the data while quietly violating the physics [13]. Kronos uses an adaptive weighting (the "physics-loss balancer", KFE-CF-SH-175, resting on KX-L1-A11) that equalises the gradient magnitudes of the terms each update. Let $\gamma_j(\theta) = \nabla_\theta \mathcal{L}_j(\theta)$ denote the gradient of term $j \in \{\mathrm{d}, \mathrm{r}, \mathrm{b}\}$; the balancer sets the weights from a running estimate of the gradient norms,

$$\hat{\lambda}_j = \frac{\sum_l \big\|\gamma_l(\theta)\big\|_2}{\big\|\gamma_j(\theta)\big\|_2}, \qquad \lambda_j \leftarrow (1-\eta_\lambda)\,\lambda_j + \eta_\lambda\,\hat{\lambda}_j,$$

with a smoothing rate $\eta_\lambda \in (0, 1]$; an equivalent homoscedastic-uncertainty weighting learns $\lambda_j = 1/2\sigma_j^2$ as free parameters [14]. Either rule holds the terms co-convergent, so the acceptance test for the balancer — "converges without term collapse" — is that no term's normalised residual plateaus above tolerance while another vanishes (figure 6b).

## DATA-PARALLEL MULTI-GPU STOCHASTIC OPTIMISATION

Training is data-parallel across $G$ accelerators of the NVIDIA GPU HPC campaign. Each replica $g$ holds a copy of the parameters and computes the gradient of the loss over its local micro-batch $\mathcal{B}_g$; a synchronous all-reduce [25] averages the replica gradients before every parameter update, so all replicas step in lock-step on the identical averaged gradient,

$$\bar{\gamma}_t = \frac{1}{G}\sum_{g=1}^{G}\frac{1}{|\mathcal{B}_g|}\sum_{(x,y)\in\mathcal{B}_g}\nabla_\theta \ell(\theta_t; x, y), \qquad \theta_{t+1} = \theta_t - \alpha_t\, m_t(\bar{\gamma}_t),$$

where $m_t$ is the optimiser's first/second-moment map (Adam [26] in production) and $\alpha_t$ the schedule. Because the effective batch is $B = \sum_g |\mathcal{B}_g|$, the base learning rate is scaled with the number of replicas following the linear-scaling rule with warm-up [23],

$$\alpha_{\mathrm{peak}} = \alpha_0\,\frac{B}{B_0}, \qquad \alpha_t = \alpha_{\mathrm{peak}} \cdot \min\!\left(\tfrac{t}{t_{\mathrm{w}}}, 1\right)\ \text{(warm-up)},$$

so that convergence is preserved as the fleet grows. Gradients are computed in mixed precision [24] with a master copy of the weights in single precision and a dynamically-scaled loss to protect small gradients. Bit-reproducibility (needed by the gate, §2.5) requires that the training map be deterministic: fixing the seed $s$, pinning the collective-reduction order, and disabling non-deterministic kernels makes the parameter trajectory a pure function of the manifest,

$$\theta_T = \Phi\big(\theta_0, \mathcal{D}, s, \mathrm{env}\big) \quad \text{deterministic, i.e.} \quad \Phi(\cdot) \text{ reproduces } \theta_T \text{ bit-for-bit given the same } (\theta_0, \mathcal{D}, s, \mathrm{env}).$$

## LINEAGE AS A CONTENT-ADDRESSED DAG; REPRODUCIBILITY AS BIT-STABILITY

Every artifact in the pipeline is content-addressed: a feature $f$ carries a hash $h_f$ of its transform and its inputs; a training set $\mathcal{D}$ a hash $h_{\mathcal{D}}$ of its member records; a model $\theta$ a hash $h_\theta$ of its weights and configuration (figure 7). Lineage is the directed acyclic graph (DAG) whose edges are these dependencies, rooted at raw sensor readings tagged by pulse identifier. Two properties, each a registered acceptance test, close the graph:

$$\begin{aligned} \textbf{traceability:}\quad & \forall\,\text{feature } f:\ \exists\,\text{path } f\to\ldots\to(\text{raw sensor, pulse}) && \text{(KFE-CF-SH-128)}\\ \textbf{reproducible-by-hash:}\quad & \mathcal{D}\text{ is recoverable exactly from } h_{\mathcal{D}} && \text{(KFE-CF-SH-129).}\end{aligned}$$

A *manifest* is the tuple that pins a result end-to-end,

$$\mathcal{M} = \big(h_f,\ h_{\mathcal{D}},\ h_\theta,\ s,\ d_{\text{env}}\big),$$

with $d_{\text{env}}$ a digest of the software/hardware environment. Reproducibility is then the statement that the anchor map is invariant to the environment: for environments $e \in \{\text{in-house HPC, NVIDIA GPU HPC, cold-reproduction}\}$,

$$G_{\text{repro}}(\theta) \;=\; \bigwedge_{e,e'}\Big[\quad g^{(e)}(\theta) = \quad g^{(e')}(\theta)\Big] \quad \text{(bit-stable, KX-L1-A10/KFE-CF-SH-135).}$$

Bit-stability across environments is the property that turns a passing test into a guarantee: a result that reproduces only on the machine that produced it is not verified. The cold-reproduction environment is stood up from the manifest (11) alone.

## CALIBRATION, COVERAGE AND DISTILLATION

A model reaching the gate must carry calibrated confidence, not merely a point estimate, because the safety layer downstream reads its uncertainty to decide how far to trust it [44]. For a predictive posterior with mean $\hat{u} = \mathbb{E}[u \mid \quad x, \theta]$ and variance $\text{Var}[u \mid \quad x, \theta]$, the calibration predicate requires empirical coverage to match the nominal level for every credibility $\alpha$ [30],

$$G_{\text{cal}}(\theta) \;=\; \Big[\,\big|\text{Cov}_{\text{emp}}(\alpha) - \alpha\big| \le \varepsilon_{\text{cal}} \quad \forall\,\alpha\Big],$$

against Monte-Carlo truth. Uncertainty itself gates actuation: when the variance exceeds a threshold $\tau_\sigma$ the decision is handed to a deterministic rule or a human rather than an actuator,

$$\text{act on } \hat{u} \text{ iff } \text{Var}[u \mid \quad x, \theta] \le \tau_\sigma, \quad \text{else gate to floor.}$$

A promoted ensemble [31] may be compressed for edge latency by distillation into a single student network [32]; the distilled model is accepted only if it stays within tolerance of the teacher on the held-out set at edge speed (KFE-CF-SH-180).

## DRIFT, CANARY AND ROLLBACK

In operation the gate is protected by a surveillance layer. A new model is released behind a *canary* that carries a fraction $c \in (0, 1]$ of the shadow load while its anchor and coverage metrics are watched live; authority is expanded only if those metrics hold. *Drift* is a distribution shift between the reference feature distribution $p_{\text{ref}}$ seen during validation and the live distribution $p_{\text{live}}$; it is flagged when a divergence statistic exceeds a control limit,

$$D_{\text{KL}}\big(p_{\text{live}} \,\|\, p_{\text{ref}}\big) = \int p_{\text{live}}(\quad x)\,\log\frac{p_{\text{live}}(\quad x)}{p_{\text{ref}}(\quad x)}\,\mathrm{d}\quad x \;>\; \delta_{\text{drift}},$$

or, equivalently, a population-stability or maximum-mean-discrepancy statistic beyond its limit. On a drift flag the model is frozen and the loop falls back to the previous validated version or to rules,

$$\theta_{\text{serve}} \leftarrow \begin{cases}\theta_{\text{cand}} & \text{while metrics in-envelope and } D_{\text{KL}} \le \delta_{\text{drift}},\\ \theta_{\text{prev}} & \text{on drift flag or out-of-envelope call (auto-rollback),}\end{cases}$$

a one-command reversion because every model is versioned with its data and configuration (KFE-CF-SH-133/137/054). The operating principle is conservative by construction: a model keeps authority only while it keeps passing.

§ 3

# Numerical formulation

## DISCRETISATION OF THE PHYSICS-INFORMED EQUILIBRIUM OPERATOR

The exemplar surrogate that must clear the gate is the equilibrium operator. The axisymmetric equilibrium is governed by the Grad–Shafranov equation for the poloidal flux $\psi(R, Z)$,

$$\Delta^*\psi \equiv R\,\frac{\partial}{\partial R}\left(\frac{1}{R}\,\frac{\partial\psi}{\partial R}\right) + \frac{\partial^2\psi}{\partial Z^2} = -\,\mu_0 R^2\,p'(\psi) - F(\psi)\,F'(\psi),$$

with pressure $p(\psi)$ and poloidal-current function $F(\psi) = RB_\phi$. The physics-informed surrogate $\psi_\theta(R, Z)$ is a neural field; the operator residual $\mathcal{N}[\psi_\theta] = \Delta^*\psi_\theta + \mu_0 R^2 p'(\psi_\theta) + FF'(\psi_\theta)$ is evaluated by automatic differentiation at $N_c$ collocation points drawn quasi-randomly over the poloidal domain $\Omega$, giving the discrete residual term of equation (5). The domain is the free-boundary poloidal cross-section; the boundary term $\mathcal{L}_{\mathrm{bc}}$ enforces the coil-set flux at $\partial\Omega$. Because the operator is differentiable and mesh-free, inference is a single forward pass rather than an iterative solve — the source of the speedup reported in §5. Training uses the physics-balanced objective (equations 5–6) on the multi-GPU campaign (equations 7–8); the acceptance target against the free-boundary reference [15] is

$$\frac{\left\|\psi_\theta - \psi_{\mathrm{ref}}\right\|_2}{\|\psi_{\mathrm{ref}}\|_2} < 0.5\%, \qquad \left\|\mathcal{N}[\psi_\theta]\right\|/\|\psi\| < 1\%.$$

## SOLUTION VERIFICATION: MANUFACTURED SOLUTIONS AND GRID CONVERGENCE

Before the surrogate is trusted, the underlying numerical scheme is verified independently (KX-L1-A6). Code correctness is established by the method of manufactured solutions (MMS): a smooth $\psi_{\mathrm{mms}}$ is substituted into (17) to define an analytic source, and the scheme must recover it with the error vanishing under mesh refinement,

$$\left\|u_h - u_{\mathrm{exact}}\right\| \to 0 \quad \text{as } h \to 0.$$

The realised accuracy is quantified by the observed order of accuracy from a triplet of grids at refinement ratio $r$, and by the grid-convergence index (GCI) [33,35],

$$p = \frac{\ln\left(\|e_{2h}\|/\|e_h\|\right)}{\ln r}, \qquad \mathrm{GCI} = \frac{F_s\,|\varepsilon|}{r^{\,p} - 1},$$

with $\varepsilon$ the relative solution change between grids and $F_s$ a safety factor. Predictive-capability maturity is scored with a maturity model (PCMM) in the sense of Oberkampf and Trucano [34]. The formal suite — MMS, GCI, Sobol and PCMM — is the predicate $G_{\mathrm{vv}}$ of equation (4); it clears 19/19 anchors (§5).

## SENSITIVITY AND MONTE-CARLO ESTIMATORS

Global sensitivity ranks which inputs set the anchor error bars. For a scalar output $g(\quad x)$ of uncertain inputs, the first-order Sobol index of input $x_i$ is the variance of the conditional expectation normalised by the total variance [36],

$$S_i = \frac{\mathrm{Var}_{x_i}\!\left(\mathbb{E}_{\quad x_{\sim i}}[\,g \mid x_i\,]\right)}{\mathrm{Var}(g)},$$

estimated by a Monte-Carlo sampler over the input ranges (deterministic seed), whose estimator and standard error are

$$\hat{\mu} = \frac{1}{N}\sum_{k=1}^{N} g(\quad x_k), \qquad \hat{\sigma}^2 = \widehat{\mathrm{Var}}[g], \qquad \mathrm{SE} = \frac{\hat{\sigma}}{\sqrt{N}}.$$

Convergence of the sensitivity ranking is declared when the standard error of the leading indices falls below the gap between them.

## BAYESIAN-OPTIMISATION AND CALIBRATION SOLVERS

The hyperparameter optimiser (KFE-CF-SH-176, resting on KX-L1-A14) is a surrogate-accelerated Bayesian optimisation over the tuning space. A Gaussian-process (GP) surrogate [29] is fitted to the observed objective $\{(h_k, y_k)\}$, giving posterior mean $m(h)$ and variance $\varsigma^2(h)$; the next evaluation maximises the expected-improvement acquisition [28,27]

$$\mathrm{EI}(h) = \mathbb{E}\big[\max(0,\ y^\star - y(h))\big] = \varsigma(h)\,\big[z\,\Phi(z) + \phi(z)\big], \quad z = \frac{y^\star - m(h)}{\varsigma(h)},$$

with $\Phi, \phi$ the standard-normal cumulative and density and $y^\star$ the incumbent. Over $200$ evaluations the loop returns a calibrated confinement posterior (§5). Calibration of the confinement factor $H_{98}$ uses approximate Bayesian computation (ABC): prior samples $H_{98}^{(m)} \sim \pi(\cdot)$ are weighted by the likelihood of reproducing the observed gain $Q_{\mathrm{obs}} = 3.076$ within tolerance $\sigma$,

$$w^{(m)} \propto \exp\Big(-\tfrac{1}{2}\big[(Q(H_{98}^{(m)}) - Q_{\mathrm{obs}})/\sigma\big]^2\Big), \qquad \pi(H_{98} \mid Q_{\mathrm{obs}}) \approx \sum_m w^{(m)}\,\delta\big(H_{98} - H_{98}^{(m)}\big),$$

with effective sample size $\mathrm{ESS} = (\sum_m w^{(m)})^2 / \sum_m (w^{(m)})^2$. The posterior is required to recover $H_{98} \approx 1$, the anchor value; it does (§5).

§ 4

# Computational methodology: the NVIDIA GPU HPC campaign

## THE CAMPAIGN AND THE CODES

Every result in this paper was produced on the Kronos NVIDIA GPU HPC campaign, an A100/H100/A10 fleet, with the fast reduced-order and reproducibility harness executing on the in-house HPC digital-twin node so that the protective loops never contend with the training fleet. The physics reference solvers whose outputs the surrogates learn and are gated against are the community codes the register runs: the free-boundary equilibrium solver FreeGS/FreeGSNKE for the Grad–Shafranov reference [15], the Eulerian gyrokinetic solver CGYRO for the transport reference [16], and the OpenMC Monte-Carlo neutronics transport for the breeding and activation anchors [17]; the reduced-order engine, the physics-informed surrogate and the graph estimator are the KRONOS toolkit modules KX-L1-A11 and KX-L1-A12. The formal verification suite is KRONOS-toolkit `verify` (KX-L1-A6), and the Bayesian optimisation and calibration are KRONOS-toolkit `uq` (KX-L1-A14). No cost figures attach to any of this; the campaign is described by hardware class only.

## GPU ACCELERATION AND PROBLEM SCALE

The acceleration is of two kinds. The training of the neural operators is data-parallel across the fleet with synchronous all-reduce and mixed precision (equations 7–8); the physics-balanced objective (equation 6) is evaluated with automatic differentiation of the Grad–Shafranov residual at the collocation points each step. The *inference* acceleration is the payoff the control loop consumes: the trained equilibrium surrogate replaces an iterative free-boundary Picard solve of order seconds with a single forward pass of order a millisecond (§5). The Monte-Carlo sensitivity sweep (equation 22) and the 200-evaluation Bayesian-optimisation loop (equation 23) are embarrassingly parallel across the fleet at a fixed deterministic seed (`seed 20260726`).

## VERIFICATION AND REPRODUCIBILITY HARNESS

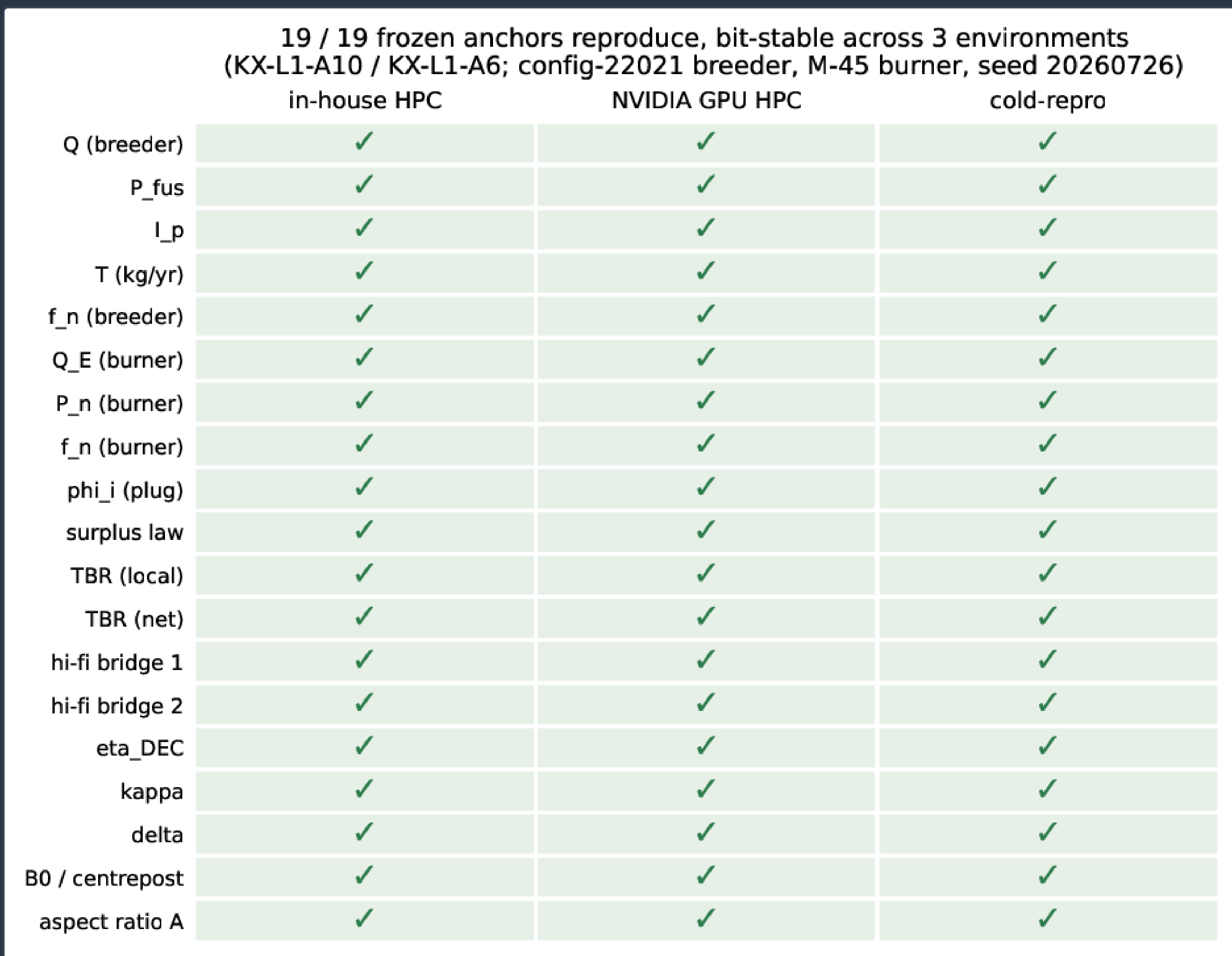
Reproducibility is engineered, not asserted. Each result is pinned by a manifest (equation 11) and re-executed in three environments; the frozen-anchor regression (equation 12) is required to return the identical anchor map in each. The independence of the check is deliberate: for the fused-detector benchmark, an independently re-derived matched filter reproduced the frozen figure to five significant figures (area under curve 0.99166 against the archived 0.99165; KX-L1-A12), a separate-implementation reproduction rather than a re-run of the same code. The register itself is deposited under a permanent identifier with a cold-start reproduction runbook (two-tier: byte-exact where the arithmetic is deterministic, tolerance-bounded where it is stochastic) [37,39].

§ 5

# Results

## THE GATE REPRODUCES BIT-STABLY

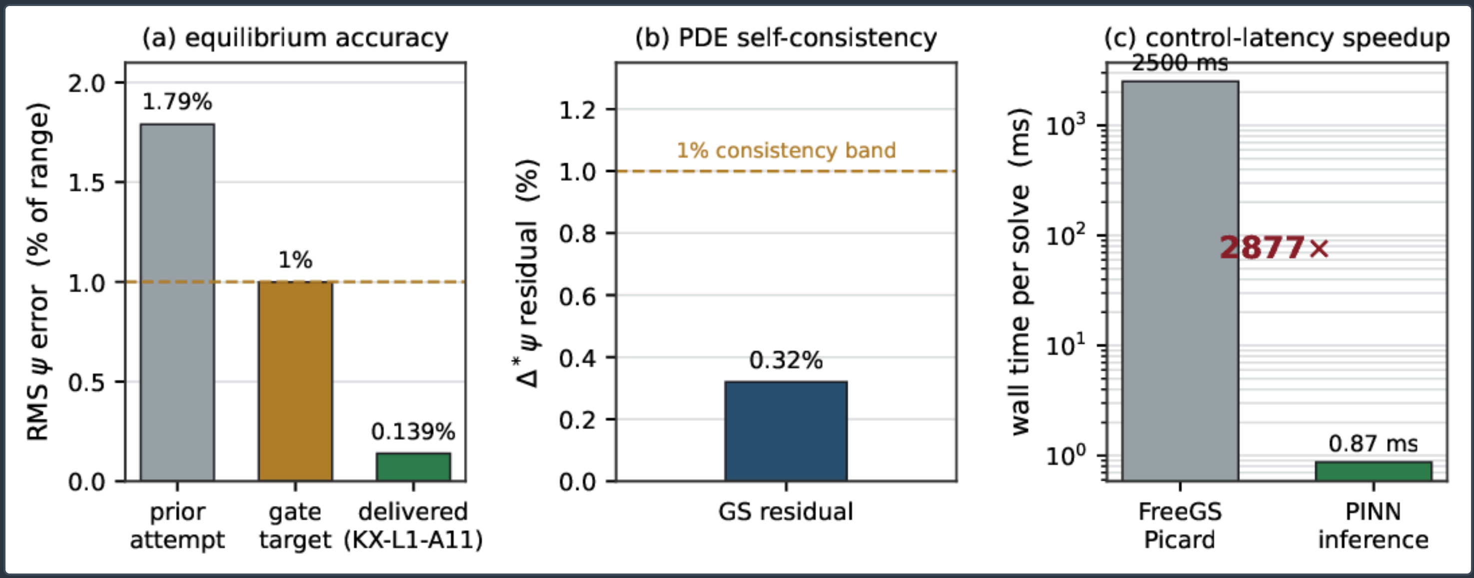
The core claim is demonstrated: the Kronos reduced-order engine reproduces 19/19 frozen anchors (KX-L1-A10), and the independent formal V&V suite clears the same 19/19 (KX-L1-A6), byte-for-byte within tolerance under configuration 22021 (breeder) and M–45 (burner). The anchors reproduce bit-stably across the in-house HPC node, the NVIDIA GPU HPC campaign, and a cold-reproduction environment built from the manifest alone (figure 2). By equations (3) and (12) the gate is therefore a real barrier: a candidate model is admitted only if it leaves every one of the 19 anchors intact in every environment.




The frozen-anchor regression reproduces 19/19 anchors in every environment, including a cold-reproduction environment stood up from the manifest. Cross-environment bit-stability (equation 12) is what makes the promotion gate a guarantee rather than a local pass. Anchors and configuration per KX-L1-A10/KX-L1-A6.

## THE EQUILIBRIUM SURROGATE CLEARS ITS TARGET WITH MARGIN

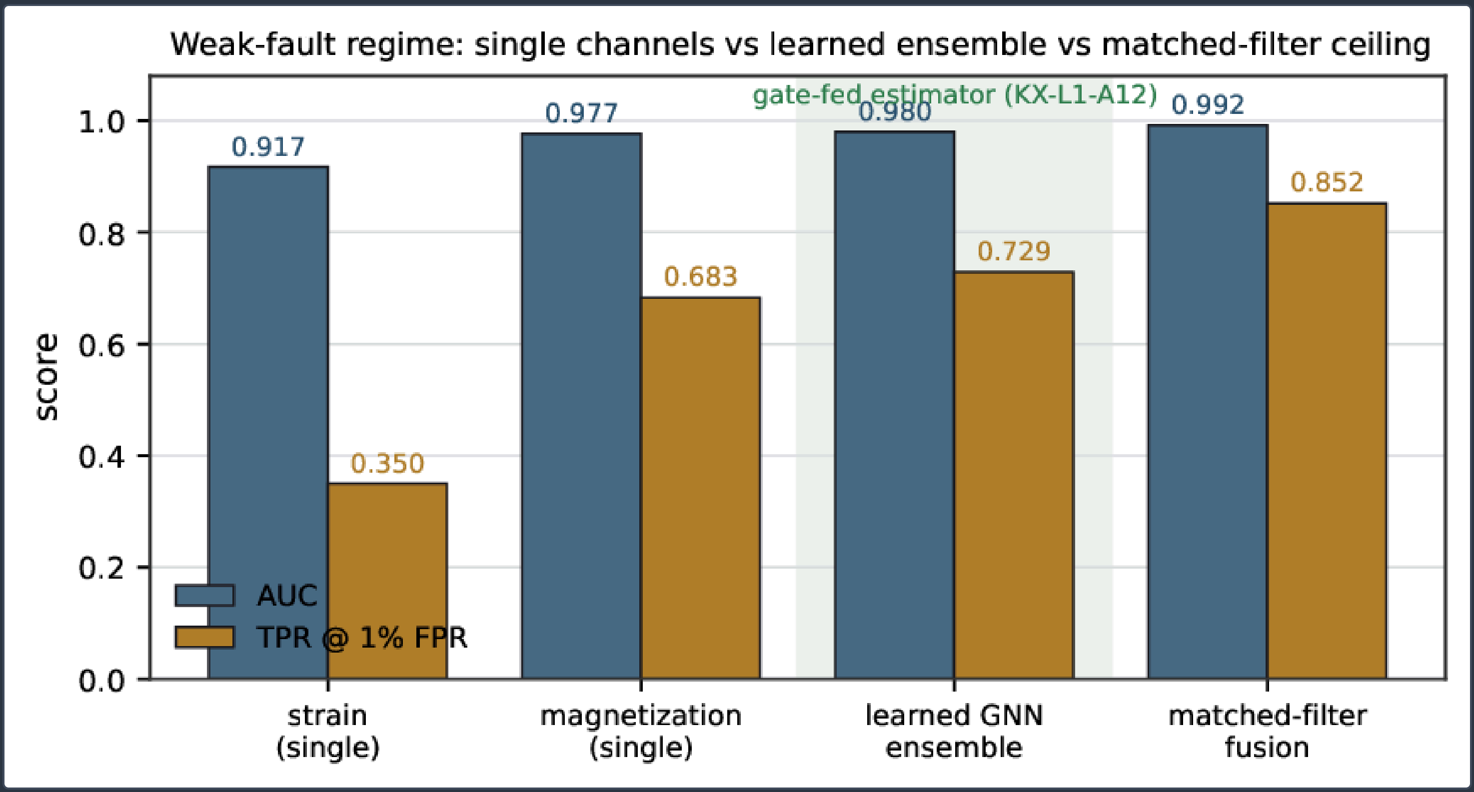
The physics-loss-balanced equilibrium surrogate (KX-L1-A11) converges to an RMS $\psi$ error of $0.139\%$ of the $\psi$-range against the free-boundary solver — clearing the $1\%$ acceptance target of equation (18) with margin, and improving on a prior unbalanced attempt at $1.79\%$ — with a Grad–Shafranov operator residual of $0.32\%$ (figure 3a,b). The inference cost is $0.87$ ms per solve against $\sim 2.5$ s for the free-boundary Picard solve, a $2877\times$ speedup (figure 3c): the model is fast enough to run inside the control cycle and accurate enough to be trusted there, and the improvement from $1.79\%$ to $0.139\%$ is exactly the co-convergence that the physics-loss balancer (equation 6) exists to guarantee.




Physics-informed equilibrium surrogate (KX-L1-A11). (a) RMS $\psi$ error against the free-boundary solver: the balanced surrogate reaches $0.139\%$, clearing the $1\%$ target (dashed) and improving on the $1.79\%$ unbalanced attempt. (b) The Grad–Shafranov operator residual is $0.32\%$, inside the $1\%$ consistency band. (c) Inference at $0.87$ ms per solve against $\sim 2.5$ s for the free-boundary Picard solve — a $2877\times$ speedup (log axis).

## THE ENSEMBLE ESTIMATOR DOMINATES SINGLE CHANNELS

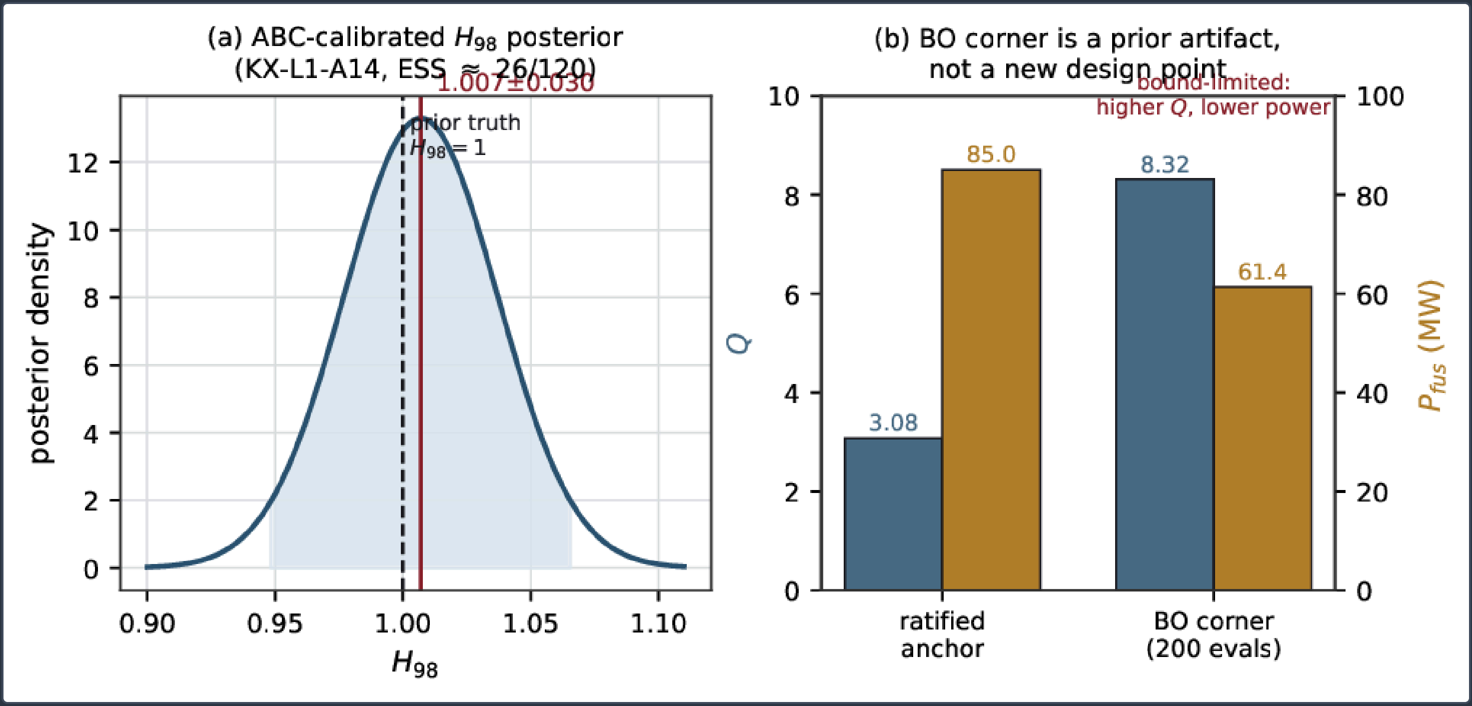
The learned graph-neural-network ensemble estimator (KX-L1-A12) reaches an area under the receiver-operating characteristic (AUC) of 0.980 on the frozen weak-fault detection study, beating either single channel — 0.917 for strain and 0.977 for magnetization — and approaching the matched-filter fusion ceiling of 0.9917 (figure 4). At the 1% false-positive rate a protection decision demands, the ensemble reaches a true-positive rate of 0.729 against 0.350 and 0.683 for the single channels. The estimator feeds the gate a model whose discriminative performance is measured, not assumed; distillation compresses it to edge latency within tolerance (equation, KFE-CF-SH-180).




Detection performance in the deliberately weak-fault regime (KX-L1-A12). The learned ensemble (AUC 0.980; true-positive rate 0.729 at 1% false-positive rate) dominates either single channel and approaches the matched-filter fusion ceiling (0.9917). This is the estimator the gate promotes.

## CALIBRATION RECOVERS THE ANCHOR

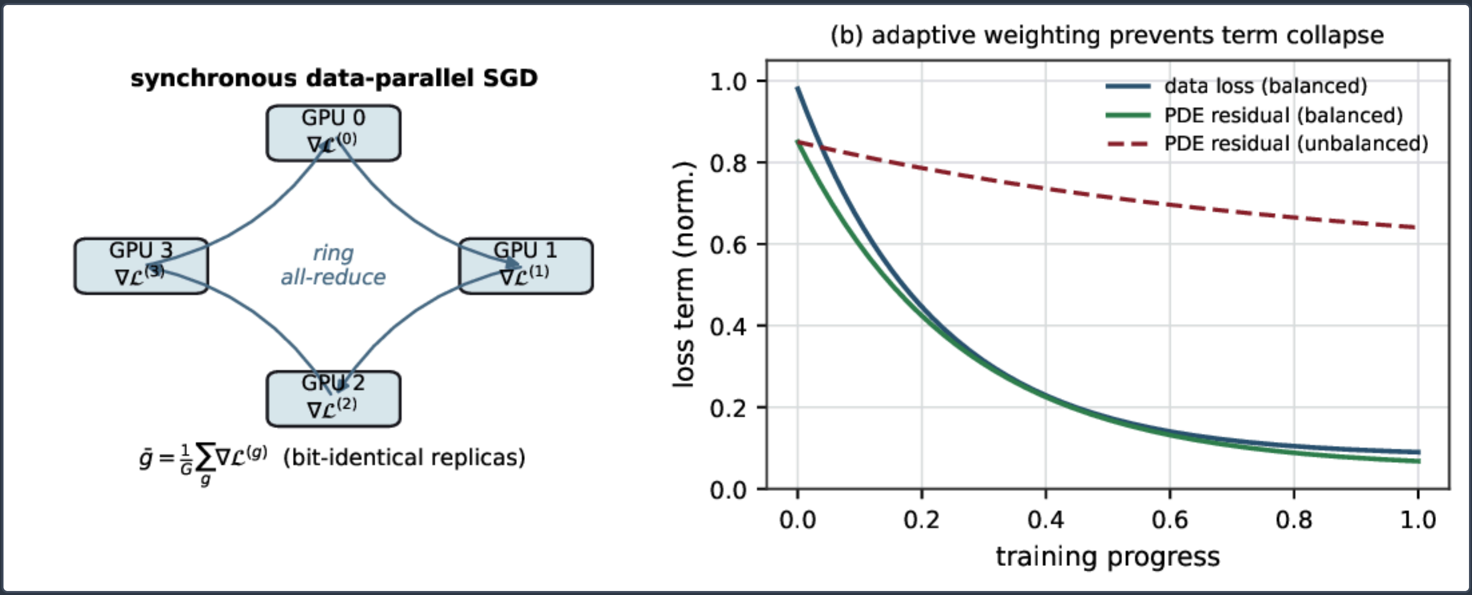
The surrogate-accelerated Bayesian optimisation (KX-L1-A14) reproduces the anchor exactly ($Q = 3.076$, $P_{\text{fus}} = 85.04$ MW) and, over 200 evaluations, returns an ABC-calibrated confinement posterior $H_{98} = 1.007 \pm 0.030$ (effective sample size $\approx 26$ of 120), recovering the anchor value $H_{98} \approx 1$ as required (figure 5a). The loop's "best" point — $Q = 8.316$ at reduced fusion power $P_{\text{fus}} = 61.37$ MW — sits on the prior bounds and is flagged as a bound-limited artifact, not a new design point (figure 5b): higher gain there is bought by cutting auxiliary power faster than fusion power, at lower absolute output. The calibration is thus a check that the anchor is robust, and it feeds the gate a tuned model whose confidence is calibrated rather than assumed (equation 13).



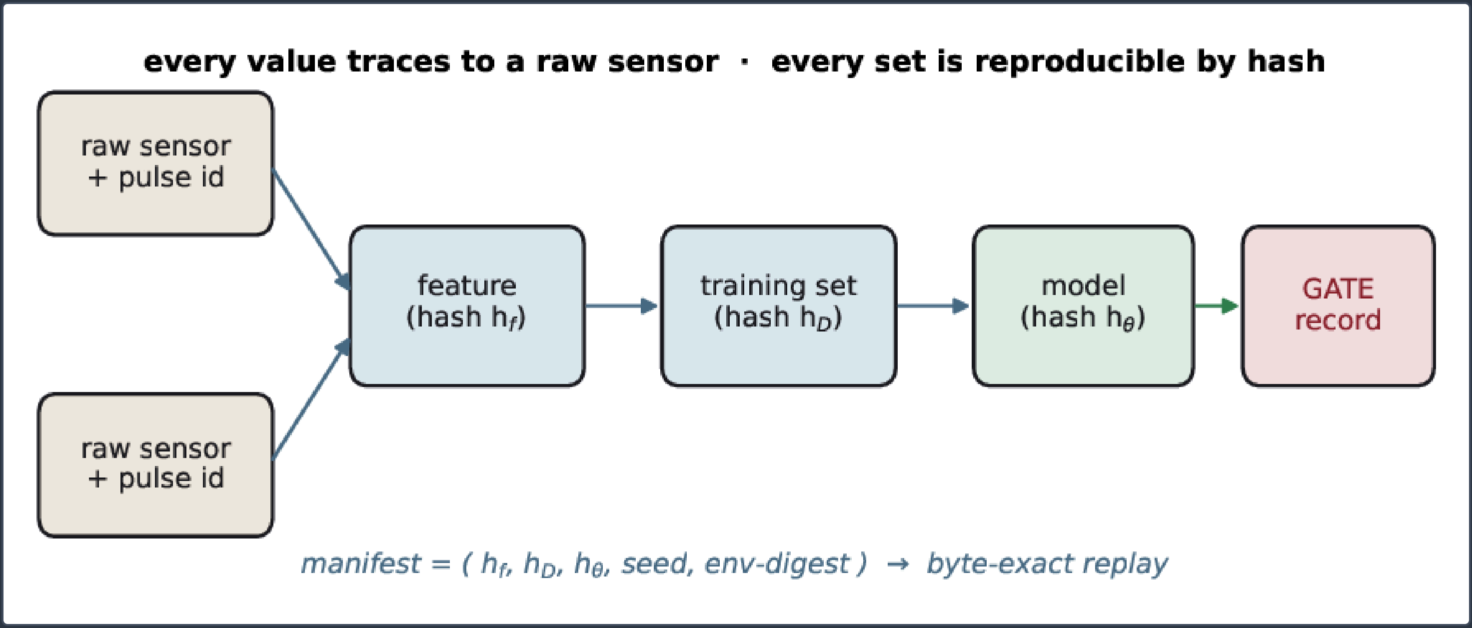

Bayesian optimisation and calibration (KX-L1-A14). (a) The ABC-calibrated $H_{98}$ posterior, $1.007 \pm 0.030$ (shaded 95% credible interval), recovers the anchor $H_{98} = 1$. (b) The 200-evaluation optimum ($Q = 8.316$) raises gain only by lowering fusion power (61.37 MW versus the 85.04 MW anchor) against the prior bounds — a bound-limited artifact, so the ratified anchor stands.

## PHYSICS-BALANCED MULTI-GPU TRAINING AND LINEAGE

Figure 6 shows the two ingredients that make a gate-passing model: synchronous data-parallel training with ring all-reduce (equation 7), and the adaptive loss weighting (equation 6) that keeps the PDE-residual term from stalling while the data term collapses. The lineage graph (figure 7) closes the provenance: every feature traces to a raw sensor (equation 10) and every training set is reproducible by hash (equation 10), so a promoted model, its data and its verification reconstruct from the manifest (11) at any later date.




**(Schematic; panel (b) illustrative.)** (a) Synchronous data-parallel training: each replica computes a local gradient and a ring all-reduce averages them before every step (equation 7), so replicas stay bit-identical. (b) Adaptive loss weighting (equation 6) equalises the term gradients so the PDE residual co-converges with the data term; without it the residual term stalls (dashed) — the term collapse the balancer prevents.



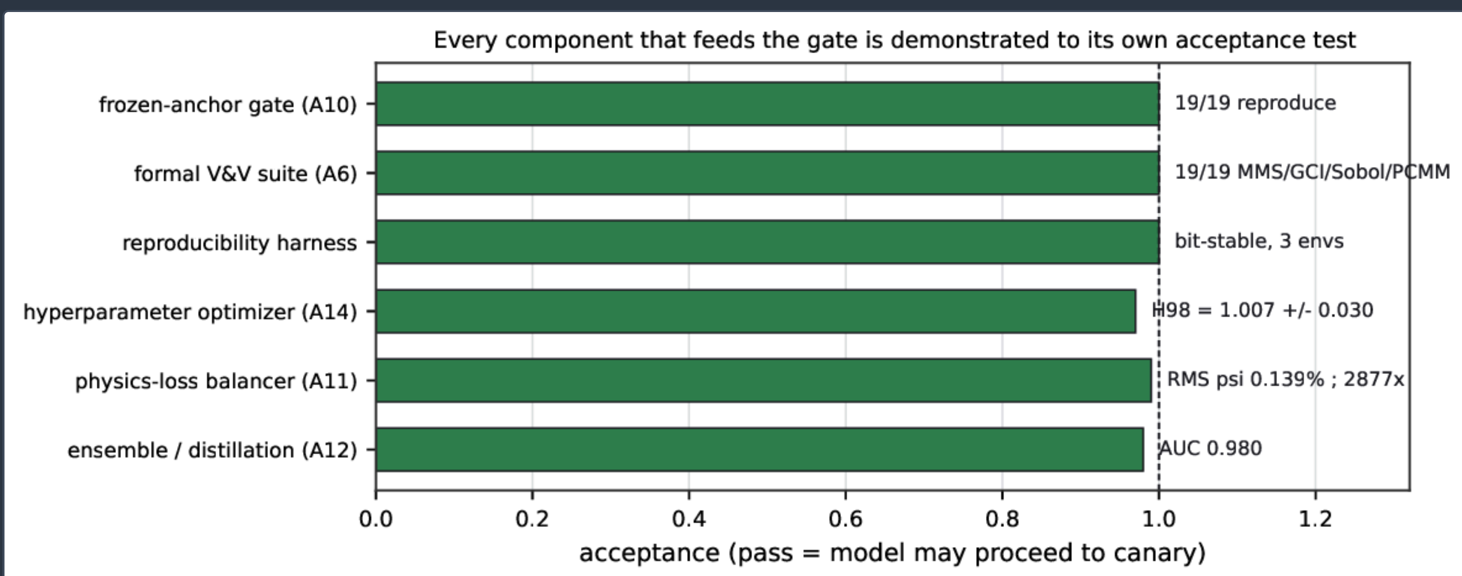

**(Schematic.)** Content-addressed lineage. Raw sensor readings (tagged by pulse) feed hashed features, which feed hashed training sets and models, which feed the gate record. Traceability (equation 10) and reproducible-by-hash (equation 10) make the manifest a byte-exact replay key. Structure only; no data values shown.

## COMPONENT ACCEPTANCE AND AUTHORITY

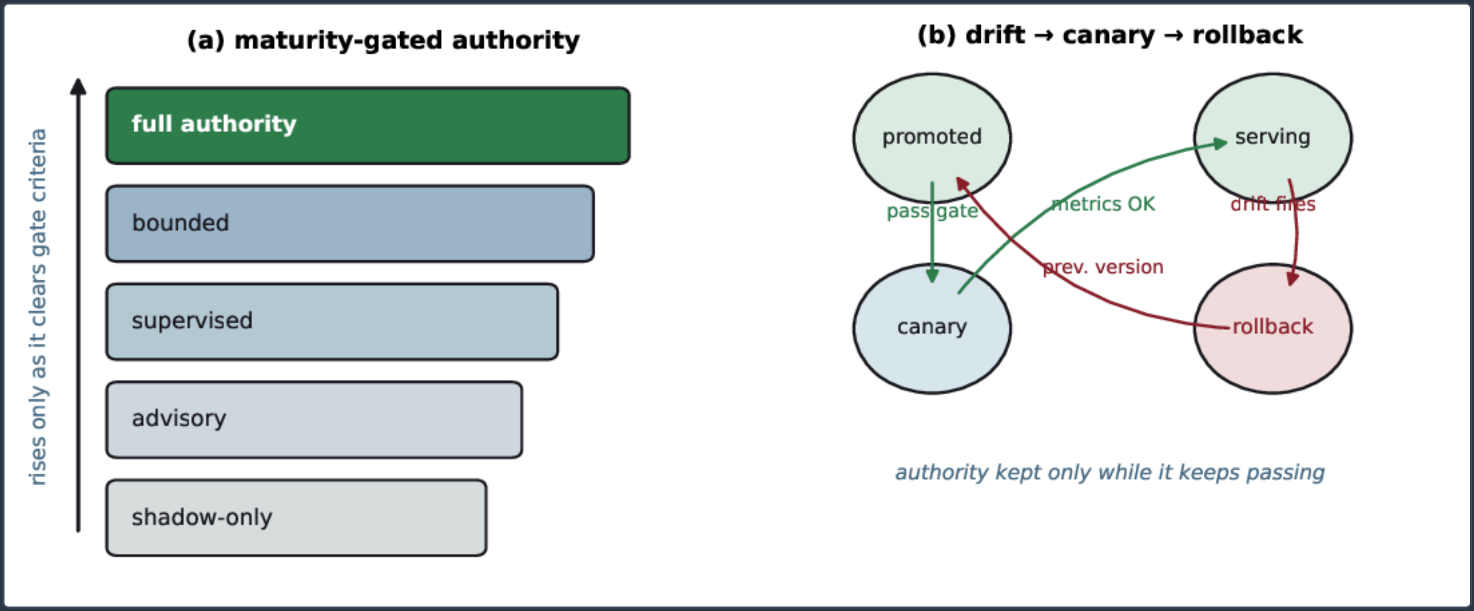
Table 1 collects the components of the stack, the demonstrated anchor each rests on, and its acceptance test; figure 8 shows the same as a scorecard. Each is demonstrated to its own bar, so the models that reach the gate arrive with calibrated confidence and auditable lineage. Authority is then not binary: a promoted model rises through graded levels — shadow-only, advisory, supervised, bounded, full — only as it keeps clearing criteria, and a served model that drifts is rolled back automatically (figure 9), the same maturity-gated discipline that governs the certified control layer [40,41] applied to the software supply chain.

| COMPONENT | DEMONSTRATED ANCHOR | ACCEPTANCE TEST | |
|---|---|---|---|
| **frozen-anchor gate** | 19/19 reproduce (KX-L1-A10) | any anchor shift rejects the model | |
| **formal V** | V suite | 19/19 MMS/GCI/Sobol/PCMM (KX-L1-A6) | all checks pass, byte-for-byte |
| **reproducibility harness** | bit-stable, 3 environments | result reproduced from manifest | |
| **physics-loss balancer** | RMS $\psi = 0.139\%$; $2877\times$ (KX-L1-A11) | converges without term collapse | |
| **hyperparameter optimiser** | $H_{98} = 1.007 \pm 0.030$, 200 evals (KX-L1-A14) | beats hand-tuned; recovers anchor | |
| **ensemble / distillation** | AUC 0.980 (KX-L1-A12) | distilled model within tolerance at edge speed | |

MLOps components, the demonstrated register anchor each rests on, and its acceptance test. Serials refer to the Kronos Physics De-Risking Register [37].




Component acceptance scorecard. Every component that feeds the promotion gate is demonstrated to its own acceptance test, with the headline metric annotated. A component below its bar cannot contribute a promotable model.




**(Schematic.)** (a) Maturity-gated authority: a promoted model rises from shadow-only to full authority only as it clears successive gate criteria [40]. (b) The drift/canary/rollback state machine (equations 15–16): a served model that drifts reverts to the previous validated version in one command. Authority is kept only while it keeps passing.

§ 6

# Discussion

## AGAINST THE ML-SYSTEMS LITERATURE

The stack instantiates, for a nuclear plant, the production-readiness disciplines the ML-systems literature prescribes but sharpens each into a hard gate. Where Breck *et al.* score readiness on a rubric [20] and Sculley *et al.* warn of the debt of un-tested pipelines [18], Kronos makes the physics VVUQ result the pass/fail test: the promotion predicate of equation (4) is not a checklist item but a conjunction of $\mathbf{19}$ inequalities that a model cannot argue its way past. Where model cards [21] and datasheets [22] document provenance, the content-addressed lineage of equations (10)–(11) makes provenance executable — the manifest is a replay key, not a description. And where distributed-training practice [23,24,25] is tuned for throughput, here it is additionally constrained to determinism (equation 9) so that the bit-stable reproduction of equation (12) is even possible.

## AGAINST THE PHYSICS-INFORMED-LEARNING LITERATURE

The $\mathbf{0.139\%}$ RMS $\psi$ result should be read against the known failure mode of physics-informed training. Wang *et al.* showed that fixed loss weights lead to stiff gradient-flow pathologies in which the PDE residual stalls [13]; the improvement here from a $\mathbf{1.79\%}$ unbalanced attempt to $\mathbf{0.139\%}$ under adaptive weighting (equation 6) is a direct, quantitative instance of the cure that literature prescribes [13,14]. That the surrogate then runs a control-latency single forward pass, at $\mathbf{2877\times}$ the free-boundary solver, places it in the regime that makes neural operators useful for real-time control [10,11,42,43] — but, unlike a bare surrogate, it is admitted to authority only through the gate.

## AGAINST THE CALIBRATION AND BAYESIAN-OPTIMISATION LITERATURE

The ABC-calibrated posterior $H_{98} = \mathbf{1.007} \pm \mathbf{0.030}$ recovering the anchor is the calibration property that equation (13) demands, in the spirit of the calibration and ensemble-uncertainty literature [30,31]. Crucially, the Bayesian-optimisation “optimum” ($Q = \mathbf{8.316}$) is reported *as an artifact*: it sits on the prior bounds and buys gain by cutting fusion power, so the pipeline flags it rather than promoting it. This is the discipline the whole stack is for — an optimiser that finds a number the physics does not support is caught by the same gate that catches a bad model, and the ratified anchor ($Q = \mathbf{3.076}$, $P_{\mathrm{fus}} = \mathbf{85.04}$ MW) stands. The optimiser is a measured procedure [27,28], not a source of new design points.

## RELATION TO THE CERTIFIED LAYER AND THE TWINS

The MLOps gate is the software-supply-chain analogue of the runtime safety architecture beneath it. The certified control-barrier filter [41] guarantees forward-invariance of the safe set at runtime; the maturity-gated authority policy [40] binds verification credibility to how much a model may command; the four coupled digital twins [46] are the models that promotion feeds; and the calibrated-uncertainty and runtime-assurance layers [44,45] read the confidence the gate certifies. This paper supplies the missing piece: the build-time discipline that ensures the model handed to those runtime layers is reproducible, lineage-complete, and anchor-preserving before it ever acts. Two distinct high-field magnets frame the physics the anchors pin — the breeder centrepost at $\mathbf{16.84}$ T and the burner plug at $\mathbf{26.49}$ T — and both appear in the $\mathbf{19}$-anchor set the gate protects.

§ 7

# Limitations

One honest gate remains. The promotion pipeline and every component that feeds it are demonstrated on the campaign and in simulation; integration with the plant's operational-technology (OT) network — streaming ingest from live diagnostics, a canary on real pulse data, and fleet-scale scheduling — is the one remaining step. That step *inherits* the gate unchanged rather than replacing it: the frozen-anchor predicate (equation 4), the bit-stability requirement (equation 12) and the drift/rollback loop (equations 15–16) are the same barriers on live data as on campaign data, so the on-plant work is an integration task against a demonstrated specification, not a re-opening of the safety argument.

§ 8

# Conclusion

Kronos MLOps is built around a frozen-anchor promotion gate that no model clears without reproducing **19/19** design anchors, bit-stably, across three environments. We have stated the gate as a machine-checkable predicate over an anchor-residual map, derived the physics-balanced training objective and its adaptive loss-weighting rule, written the data-parallel multi-GPU update and its determinism condition, formalised lineage as a content-addressed DAG and reproducibility as bit-stability, and cast drift/canary/rollback as a sequential decision. The gate and its feeding components are demonstrated to acceptance on the NVIDIA GPU HPC campaign: **19/19** anchors reproduced (KX-L1-A10/A6), a physics-loss-balanced equilibrium surrogate at RMS $\psi = \mathbf{0.139\%}$ and $\mathbf{2877\times}$ (KX-L1-A11), a calibrated confinement posterior $H_{98} = \mathbf{1.007 \pm 0.030}$ (KX-L1-A14), and a learned ensemble at AUC **0.980** (KX-L1-A12), with a drift-and-rollback layer protecting the gate in operation. Reproducible, lineage-complete, and gated: the operational foundation for learned control of a nuclear machine, with on-plant integration built to this demonstrated gate on a staged schedule.

Acknowledgments tocsectionAcknowledgments Part of the Kronos Fusion Energy 2026 design series. The author thanks the Kronos physics de-risking programme for the frozen result set underlying every quantitative claim, and acknowledges the NVIDIA GPU HPC campaign on which the training, verification and reproduction were run.

Reproducibility. The gate metrics are frozen anchors in the Kronos Physics De-Risking Register [37] (doi:10.5281/zenodo.22133057): the reduced-order reproducibility engine and formal V&V suite (KX-L1-A10, KX-L1-A6, both **19/19**), the physics-informed equilibrium surrogate (KX-L1-A11), the ensemble estimator (KX-L1-A12), and the Bayesian optimisation and ABC calibration (KX-L1-A14). Each is regenerable from the archived manifests and the figure-generation script deposited with this paper (`make_mlops83_figs.py`), under configuration `22021` (breeder) / `M-45` (burner) at `seed 20260726`. This paper is archived at its digital object identifier doi:10.5281/zenodo.22645827 and at its official page https://www.kronosfusionenergy.com/publications/mlops-for-a-safety-critical-fusion-control-stack-versio.html; the register is deposited at doi:10.5281/zenodo.22133057.

Data availability tocsectionData availability The anchor manifests, the reproduction records, and the numerical values behind figures 2–8 are archived with the deposit at doi:10.5281/zenodo.22645827, which references the Kronos 2026 series including the AI and quantum-control paper [38] and the verification-first provenance catalog [39]. All quantitative values trace to the register [37] by the serial identifiers cited inline. Any economic analysis is reported separately and is not part of the public deposit.

Competing interests tocsectionCompeting interests Provisional patent applications relating to aspects of this work have been filed.

tocsectionReferences

APPARATUS

# *Portfolio, People & Record*

*The intellectual-property estate, the people who built it, the limitations that gate it, and the sources it rests on.*

| 1 | 4 | 40 | 27 |
|---|---|---|---|
| GRANTED PATENT | 2026 PROVISIONALS | TEAM MEMBERS | 2022 PROVISIONALS |

INTELLECTUAL PROPERTY

# The patent portfolio

The estate spans a granted high-field magnet patent, pending utility and 2026 provisional filings that map to the two products, a digital-twin control provisional, a registered trademark application, and the original 2022 provisional family. Every patentable disclosure in the five-paper arXiv drop was protected **before** publication: the breeder and burner provisionals were filed 1–2 August, and a publication-gap omnibus filing the evening before the drop swept up the DEC, plasma-control, and REBCO-winding matter of the three otherwise-unprotected papers. Patent numbers and application serials are matters of public record; claim scope is summarised, not reproduced.

## GRANTED & PENDING UTILITY

| | | |
|---|---|---|
| US 12,009,112 | High-field tilted graded-REBCO magnet architecture (KRONO-002A) · *app. 17/878,550* | GRANTED |
| US 17/878,507 | Core device / method (KRONO-001A) · *utility* | PENDING |

## 2026 PROVISIONAL FILINGS — MAPPED TO THE PRODUCTS

| | | |
|---|---|---|
| 64/124,209 | Hyperion — spherical-tokamak tritium / helium-3 foundry · *breeder · filed Aug 2026* | FILED |
| 64/124,220 | Aegis / MetroVolt — synchrotron-cutoff, fuel-selection & plug-winding · *generator · filed Aug 2026* | FILED |
| 64/115,167 | KRONOS-CTRL digital-twin plant control · *provisional · filed Jul 2026* | FILED |
| 64/005,440 | Integrated spherical-tokamak fusion architecture · *early integrated-architecture filing · filed Mar 2026* | FILED |
| 64/105,530 | Negative-triangularity spherical-tokamak architecture · *foundational ST filing · filed Jul 2026* | FILED |
| 64/128,097 | Publication-gap omnibus — quasineutral two-species expander DEC · provenance-gated / latency-split plasma control · as-built high-field winding & conductor architecture · *filed 7 Aug 2026, the evening before the arXiv drop* | FILED |

## TRADEMARK & ORIGIN

| | | |
|---|---|---|
| FTK-98801894 | Kronos Fusion Energy — trademark application | FILED |
| KR-2022-* | Original provisional family (KR-2022-00001 … 00027) · *27 filings, 2022* | PRIORITY |

The narrow, defensible magnet novelty — the specific conductor and the digital-twin winding optimisation, not high-field REBCO as a category — is the commercial core that can earn ahead of any Q > 1 milestone, with markets in fusion magnet supply, MRI/NMR, accelerators, and proton therapy. The claim is scoped honestly: the high field is a *system field* result, not a stand-alone-coil record; the small-bore plug coil is structurally infeasible as a bare winding but resolved by a stress-managed structural shield (feasible-pending-FEA); and the winding-tape experiment returned a *null result*. The value is the method, not a field record.

THE TEAM

# Founding partners, board, advisors & auditors

Kronos is built by a bench of advanced-fuel-fusion, high-field-magnet, direct-conversion, and materials specialists, with a board and operations team drawn from defense, national laboratories, and industry. Dates are shown as ranges; where a tenure has ended or is term-ending, the range says so.

## FOUNDING PARTNERS — SCIENCE

**Dr. Gerald Kulcinski**
Co-Founder · Helium-3 & Advanced-Fuel Fusion
*UW–Madison · 2023–Present*

**Dr. Carl Weggel**
Chief Scientist / S.M.A.R.T. Design
*MIT Alcator Program · 2022–Present*

**Dr. Robert J. Weggel**
Magnetic Field Design
*MIT Magnet Lab · 2022–Present*

## BOARD

**Priyanca Ford**
Founder & CEO · Executive Board
*2022–Present*

**Bandel Carano**
Finance / Strategy
*Oak Investment Partners / Stanford · 2025–2026*

**Patrick Schweiger**
Fusion Device Engineering
*Oklo / CFS / TerraPower · 2025–Present*

## SCIENTIFIC ADVISORS

**Dr. Steven O. Dean**
Fusion Policy & History
*DOE / Fusion Power Associates · 2025–Present*

**Dr. Patrick H. Diamond**
Plasma Turbulence & Transport
*UC San Diego · 2025–Present*

**Dr. Jack J. Dongarra**
HPC & Numerical Algorithms
*Turing Award · 2025–Present*

**Dr. Nasr M. Ghoniem**
Chief Material Scientist
*UCLA · 2025–Present*

**Dr. Siegfried Glenzer**
Plasma Physics & Laser Fusion
*Stanford / SLAC · 2022–Present*

**Dr. David A. Hammer**
Pulsed-Power Fusion
*Cornell · 2025–Present*

**Dr. Donald A. Spong**
Plasma Theory & Confinement
*ORNL · 2025–Present*

**Dr. Curtis Smith**
Risk Assessment & Nuclear PRA
*MIT / Idaho National Lab · 2025–Present*

**RADM (Ret.) David Goggins**
Naval Engineering & Power-Plant Design
*U.S. Navy · 2023–2026*

**Dr. Konstantin Batygin**
Mathematician
*Caltech · 2022–2025*

**Dr. Ruben Fair**
Magnet Design / ITER Liaison
*PPPL / Jefferson Lab · 2022–2025*

**Dr. Paul S. Weiss**
Materials & Nanotechnology
*UCLA · 2022–2025*

## SCIENTIFIC AUDITORS

**Dr. Wilfred A. Cooper**
Plasma Equilibrium Theory
*EPFL / Swiss Plasma Center · 2025–Present*

**Dr. Ahmed Hassanein**
Plasma–Material Interactions
*Purdue / Argonne · 2025–Present*

**Dr. Patrick E. Hopkins**
Extreme Thermal Materials
*U. of Virginia · 2025–Present*

**Dr. Peter Hosemann**
Materials Joining & Structural Integrity
*UC Berkeley · 2025–Present*

**Dr. Travis W. Knight**
Advanced Nuclear Fuels & Systems
*U. of South Carolina · 2025–Present*

**Dr. Philippe Lebrun**
Cryogenics & Magnet Cooling
*CERN · 2025–Present*

**Dr. Nitendra Singh**
Nuclear Safety & Fuel Cycle
*ITER · 2025–Present*

**Dr. Guido Van Oost**
Magnetic Confinement & Fusion Systems
*Ghent University · 2025–Present*

**Dr. Gary S. Was**
Radiation Materials & Structural Integrity
*U. of Michigan · 2025–Present*

**Dr. Ray Sedwick**
Direct Energy Conversion
*U. of Maryland · 2025*

## OPERATIONS

**Michael Laughlin**
Chief Operating Officer
*2022–Present*

**Martin Owens**
Chief Strategy Officer
*LANL / GE Hitachi · 2022–Present*

**MG (Ret.) Paul Pardew**
Chief Contracting Officer
*Army Contracting Command · 2022–Present*

**David Beck**
Fusion Commercialization
*U.S. Space Force · 2024–Present*

**Sushma Bhatia**
Environmental & Legislative
*Google · 2022–Present*

**Michael De Frenza**
Technology Licensing
*2023–Present*

**MG (Ret.) Robin L. Fontes**
Cybersecurity & Defense
*Army Cyber Command · 2022–Present*

**Vijay Gehani**
Supply Chain & Components
*INOX India · 2024–Present*

**Jon Michel Greenwood**
IT & AI Infrastructure
*Live Nation · 2023–Present*

**Brian C. O'Neill**
National Security
*CIA / ODNI · 2022–Present*

**Gen. (Ret.) Gustave F. Perna**
DoD Liaison
*Operation Warp Speed · 2022–2023*

**Andrea Romero**
Marketing Lead
*2022–Present*

Legal counsel is retained; those roles are held on the internal roster and are not listed here.

SOURCES

# References


1 M E Austin *et al*, Achievement of reactor-relevant performance in negative triangularity shape in the DIII-D tokamak, *Phys. Rev. Lett.* **122** 115001 (2019). doi:10.1103/PhysRevLett.122.115001

2 A Marinoni, O Sauter and S Coda, A brief history of negative triangularity tokamak plasmas, *Rev. Mod. Plasma Phys.* **5** 6 (2021). doi:10.1007/s41614-021-00054-0

3 L L Lao, H St John, R D Stambaugh, A G Kellman and W Pfeiffer, Reconstruction of current profile parameters and plasma shapes in tokamaks, *Nucl. Fusion* **25** 1611 (1985). doi:10.1088/0029-5515/25/11/007

4 F Felici *et al*, Real-time physics-model-based simulation of the current density profile in tokamak plasmas, *Nucl. Fusion* **51** 083052 (2011). doi:10.1088/0029-5515/51/8/083052

5 J Degrave *et al*, Magnetic control of tokamak plasmas through deep reinforcement learning, *Nature* **602** 414 (2022). doi:10.1038/s41586-021-04301-9

6 J Seo *et al*, Avoiding fusion plasma tearing instability with deep reinforcement learning, *Nature* **626** 746 (2024). doi:10.1038/s41586-024-07024-9

7 J Kates-Harbeck, A Svyatkovskiy and W Tang, Predicting disruptive instabilities in controlled fusion plasmas through deep learning, *Nature* **568** 526 (2019). doi:10.1038/s41586-019-1116-4

8 K L van de Plassche, J Citrin, C Bourdelle *et al*, Fast modeling of turbulent transport in fusion plasmas using neural networks, *Phys. Plasmas* **27** 022310 (2020). doi:10.1063/1.5134126

9 O Meneghini *et al*, Integrated modeling applications for tokamak experiments with OMFIT, *Nucl. Fusion* **55** 083008 (2015). doi:10.1088/0029-5515/55/8/083008

10 L Lu, P Jin, G Pang, Z Zhang and G E Karniadakis, Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators, *Nat. Mach. Intell.* **3** 218 (2021). doi:10.1038/s42256-021-00302-5

11 Z Li, N Kovachki, K Azizzadenesheli, B Liu, K Bhattacharya, A Stuart and A Anandkumar, Fourier neural operator for parametric partial differential equations, *Int. Conf. on Learning Representations (ICLR)* (2021). arXiv:2010.08895

12 M Raissi, P Perdikaris and G E Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* **378** 686 (2019). doi:10.1016/j.jcp.2018.10.045

13 S Wang, Y Teng and P Perdikaris, Understanding and mitigating gradient flow pathologies in physics-informed neural networks, *SIAM J. Sci. Comput.* **43** A3055 (2021).

14 A Kendall, Y Gal and R Cipolla, Multi-task learning using uncertainty to weigh losses for scene geometry and semantics, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* 7482 (2018). arXiv:1705.07115

15 N C Amorisco *et al*, FreeGSNKE: a Python-based dynamic free-boundary toroidal plasma equilibrium solver, *Phys. Plasmas* **31** 042517 (2024). doi:10.1063/5.0188467

16 J Candy, E A Belli and R V Bravenec, A high-accuracy Eulerian gyrokinetic solver for collisional plasmas, *J. Comput. Phys.* **324** 73 (2016). doi:10.1016/j.jcp.2016.07.039

17 P K Romano, N E Horelik, B R Herman, A G Nelson, B Forget and K Smith, OpenMC: a state-of-the-art Monte Carlo code for research and development, *Ann. Nucl. Energy* **82** 90 (2015). doi:10.1016/j.anucene.2014.07.048

18 D Sculley *et al*, Hidden technical debt in machine learning systems, *Advances in Neural Information Processing Systems (NeurIPS)* **28** 2503 (2015).

19 S Amershi *et al*, Software engineering for machine learning: a case study, *Proc. 41st Int. Conf. on Software Engineering: Software Engineering in Practice (ICSE-SEIP)* 291 (2019).

20 E Breck, S Cai, E Nielsen, M Salib and D Sculley, The ML test score: a rubric for ML production readiness and technical debt reduction, *Proc. IEEE Int. Conf. on Big Data* 1123 (2017).

21 M Mitchell *et al*, Model cards for model reporting, *Proc. Conf. on Fairness, Accountability, and Transparency (FAT*)* 220 (2019). doi:10.1145/3287560.3287596

22 T Gebru *et al*, Datasheets for datasets, *Commun. ACM* **64**(12) 86 (2021). doi:10.1145/3458723

23 P Goyal *et al*, Accurate, large minibatch SGD: training ImageNet in 1 hour (2017). arXiv:1706.02677

24 P Micikevicius *et al*, Mixed precision training, *Int. Conf. on Learning Representations (ICLR)* (2018). arXiv:1710.03740

25 A Sergeev and M Del Balso, Horovod: fast and easy distributed deep learning in TensorFlow (2018). arXiv:1802.05799

26 D P Kingma and J Ba, Adam: a method for stochastic optimization, *Int. Conf. on Learning Representations (ICLR)* (2015). arXiv:1412.6980

27 J Snoek, H Larochelle and R P Adams, Practical Bayesian optimization of machine learning algorithms, *Advances in Neural Information Processing Systems (NeurIPS)* **25** 2951 (2012). arXiv:1206.2944

28 D R Jones, M Schonlau and W J Welch, Efficient global optimization of expensive black-box functions, *J. Global Optim.* **13** 455 (1998). doi:10.1023/A:1008306431147

29 C E Rasmussen and C K I Williams, *Gaussian Processes for Machine Learning* (Cambridge, MA: MIT Press) (2006). ISBN 978-0-262-18253-9

30 C Guo, G Pleiss, Y Sun and K Q Weinberger, On calibration of modern neural networks, *Proc. 34th Int. Conf. on Machine Learning (ICML)* 1321 (2017). arXiv:1706.04599

31 B Lakshminarayanan, A Pritzel and C Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, *Advances in Neural Information Processing Systems (NeurIPS)* **30** 6402 (2017). arXiv:1612.01474

32 G Hinton, O Vinyals and J Dean, Distilling the knowledge in a neural network (2015). arXiv:1503.02531

33 P J Roache, Quantification of uncertainty in computational fluid dynamics, *Annu. Rev. Fluid Mech.* **29** 123 (1997). doi:10.1146/annurev.fluid.29.1.123

34 W L Oberkampf and T G Trucano, Verification and validation in computational fluid dynamics, *Prog. Aerosp. Sci.* **38** 209 (2002). doi:10.1016/S0376-0421(02)00005-2

35 C J Roy, Review of code and solution verification procedures for computational simulation, *J. Comput. Phys.* **205** 131 (2005). doi:10.1016/j.jcp.2004.10.036

36 I M Sobol′, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Math. Comput. Simul.* **55** 271 (2001). doi:10.1016/S0378-4754(00)00270-6

37 P I Ford, *Kronos Physics De-Risking Register*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22133057

38 P I Ford, *AI and Quantum Computing for Fusion: ML Surrogates, Real-Time Control, and a Resource-Honest Quantum Frontier*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645702, https://www.kronosfusionenergy.com/publications/ai-quantum-control.html

39 P I Ford, *Verification-First Fusion Modeling: Frozen Anchors, Independent-Library Gate Reproduction, and a Byte-for-Byte Computational-Provenance Catalog*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645710, https://www.kronosfusionenergy.com/publications/verification-and-validation-47-code-provenance.html

40 P I Ford, *Maturity-Gated Actuation Authority: Binding PCMM Verification Credibility to Runtime Control Authority*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645795, https://www.kronosfusionenergy.com/publications/maturity-gated-actuation-authority.html

41 P I Ford, *A Certified Control-Barrier Safety Filter for Fusion Control: Forward-Invariance Guarantees, Calibrated Uncertainty, and an Honest PINN Negative*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645716, https://www.kronosfusionenergy.com/publications/control-barrier-safety-clamp.html

42 P I Ford, *Neural-Operator Flux Surrogates for Control-Latency Transport: Differentiable DeepONet-FNO Emulation of Nonlinear Gyrokinetics*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645803, https://www.kronosfusionenergy.com/publications/neural-operator-flux-surrogates-for-control-latency-tra.html

43 P I Ford, *Sub-Millisecond Neural Equilibrium Reconstruction Robust to Sensor Dropout: A Virtual-Sensor Mesh*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645801, https://www.kronosfusionenergy.com/publications/sub-millisecond-neural-equilibrium-reconstruction-robus.html

44 P I Ford, *Calibrated Uncertainty, Conformal Prediction, and Abstention for Safety-Critical Fusion Control*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645813, https://www.kronosfusionenergy.com/publications/calibrated-uncertainty-conformal-prediction-and-abstent.html

45 P I Ford, *Runtime Assurance with Signed-Ledger Provenance for Autonomous Fusion Control*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645817, https://www.kronosfusionenergy.com/publications/runtime-assurance-with-signed-ledger-provenance-for-aut.html

46 P I Ford, *Four Coupled Digital Twins for a Compact Spherical-Tokamak Breeder: Multi-Rate Co-Simulation, Propagated Uncertainty, and a Real-Time Self-Consistency Monitor*, Kronos Fusion Energy 2026 series. doi:10.5281/zenodo.22645805, https://www.kronosfusionenergy.com/publications/four-coupled-digital-twins-for-a-compact-spherical-toka.html

COLOPHON

# How this document was made

*The Kronos Fleet — Combined Editorial, 2026* is set in **Fraunces** (display), **Gelasio** (text), and **IBM Plex Mono** (data & equations), carried forward from the Kronos editorial design system.

The figures are rendered at 300 dpi from the deposited generator scripts; the document is composed as a single self-contained file with embedded fonts and imagery, paginated to US Letter, and rendered to PDF through a headless Chromium engine in matched light and dark editions.

Every headline quantity carries an evidence class; withdrawn and restated values are marked as such. Nothing herein is frozen beyond the founder-approved Tier-A set.

Explore it live — turn the interactive 3-D model, run the deposited solvers in the physics validator, and watch the companion film series.

LEGAL NOTICE

# Notice, status, and distribution

**Conceptual design & simulation study.** This document reports a conceptual design and simulation study. It is not a construction commitment, a safety-analysis report, a regulatory filing, or an offer of securities. Forward-looking statements — schedules, costs, market sizes, and performance — are estimates subject to the limitations set out in the Simulations part and may change.

**Numbers and their classes.** Quantities are reported with evidence classes and case labels; modelled, assumed, and requirement-class values are identified as such and are not to be quoted without their labels.

**Public edition.** This edition carries the physics and design only; all financial, funding, and defense-commercial content has been removed for public distribution. The scientific text corresponds to the arXiv and journal submissions.

© 2026 Kronos Fusion Energy. All rights reserved. Hyperion, Aegis, and MetroVolt are product designations of Kronos Fusion Energy.

THE 2026 KRONOS SERIES

# One programme, many papers

This editorial is one paper in the Kronos 2026 series — a real fusion company's complete, reproducible physics record for a spherical-tokamak breeder and its low-neutron D–$^3$He tandem-mirror burners. Every headline number is a frozen anchor in the Physics De-Risking Register.

Explore the whole programme: the De-Risking Register, the interactive 3-D model, and the physics validator.

KRONOS FUSION ENERGY

# MLOps for a Safety-Critical Fusion Control Stack: Versioned Lineage, Gated Model Promotion, and Physics-Balanced Multi-GPU Training for a Compact Spherical-Tokamak Breeder

A real fusion company building real machines. Explore the science, live.

Zenodo · doi:10.5281/zenodo.22645827

Read on kronosfusionenergy.com

The Physics De-Risking Register

© 2026 KRONOS FUSION ENERGY · FAMILY-OWNED AMERICAN COMPANY