

## Correcting Sensor-Induced Distribution Drift with Wasserstein Adversarial Learning

**Authors:** Ali, Ms. Saraa, Bocharnikov, Mr. Vladimir, Ratnikov, Dr. Fedor, Hushchyn, Dr. Mikhail, Ryzhikov, Dr. Artem, Derkach, Dr. Denis, Ali, Ms. Saraa

**Date:** 2026-03-12T03:45:27+00:00

### Abstract

The quality of recorded data depends on the stability of the sensor system that acquires it. Sensor motion and aging can degrade the performance and stability of downstream data-driven methods. We present a Wasserstein-GAN-inspired approach for unsupervised inference of physically interpretable transformation parameters that map a changed detector response distribution back to a nominal reference distribution. In contrast to standard generative modeling, the generator is used as a learnable calibration transformation whose trainable weights represent the sought parameters, while the critic provides a distributional distance signal via the Wasserstein objective. We validate the approach on a tracking-detector toy model with controlled layer shifts and demonstrate its application on high-granularity Geant4-simulated calorimeter data with cell-wise aging effects. The method recovers aging coefficients for individual cells with correlation to ground truth and improves agreement between calibrated and reference energy-sum distributions, while exhibiting the expected degradation at increasing channel-to-channel noise levels. These results indicate that adversarial distribution matching can serve as a data-driven component of calibration strategies in settings where direct labels for degradation parameters are unavailable.

### Full Text

### Preamble

Correcting Sensor-Induced Distribution Drift with Wasserstein Adversarial Learning Saraa Ali,<sup>1</sup>, \* Vladimir Bocharnikov,<sup>1</sup> Fedor Ratnikov,<sup>1</sup> Mikhail Hushchyn,<sup>1</sup> Artem Ryzhikov,<sup>1</sup> and Denis Derkach<sup>1</sup> <sup>1</sup>Laboratory of Methods for Big Data Analysis, HSE University, Pokrovsky Boulevard 11, Moscow 109028, Russia The quality of recorded data depends on the stability of the sensor

system that acquires it. Sensor motion and aging can degrade the performance and stability of downstream data-driven methods. We present a Wasserstein-GAN-inspired approach for unsupervised inference of physically interpretable transformation parameters that map a changed detector response distribution back to a nominal reference distribution. In contrast to standard generative modeling, the generator is used as a learnable calibration transformation whose trainable weights represent the sought parameters, while the critic provides a distributional distance signal via the Wasserstein objective. We validate the approach on a tracking-detector toy model with controlled layer shifts and demonstrate its application on high-granularity Geant4-simulated calorimeter data with cell-wise aging effects. The method recovers aging coefficients for individual cells with correlation to ground truth and improves agreement between calibrated and reference energy-sum distributions, while exhibiting the expected degradation at increasing channel-to-channel noise levels. These results indicate that adversarial distribution matching can serve as a data-driven component of calibration strategies in settings where direct labels for degradation parameters are unavailable.

Keywords: Machine Learning, Deep Learning, Generative Adversarial Neural Networks, Calibration, High Energy Physics

## INTRODUCTION

The integrity of any data-driven analysis pipeline depends on the quality and stability of the measurements provided by the underlying sensor system [1]. In experimental environments, sensor responses are rarely stationary: mechanical drifts, environmental conditions, and long term material aging can introduce systematic distortions that accumulate over time [2]. Even when these effects are small at the level of individual channels, they can propagate to downstream reconstruction and inference, leading to degraded resolution, biased estimates, or reduced robustness of machine-learning-based decision systems [3].

Modern detectors [4] operate over long time scales and under intense radiation exposure [5-7], and their response must be continuously monitored and corrected to preserve the fidelity of physics measurements [6-8]. Calibration [6, 8, 9] and alignment [9, 10] procedures are therefore a core part of detector operations. However, comprehensive recalibration may be resource-intensive and is not always feasible at the frequency that changing detector conditions would ideally require [8, 11, 12]. A key difficulty is that the parameters governing detector degradation are often not directly observable.

In many realistic scenarios, one has access to data recorded under nominal (reference) conditions and data recorded under changed conditions, but lacks event-level correspondences or labeled degradation targets [13, 14]. This motivates approaches that infer detector-change parameters by exploiting distributional differences between datasets rather than relying on supervised targets or explicit likelihood models [15, 16].

From a machine learning perspective, such non- [17], also stationarities correspond to distribution shift referred to as dataset shift or domain shift, where the data- generating distribution differs between two conditions. While drift monitoring is typically formalized as a decision problem \* Corresponding author tharaaali@hse.ru (i.e., whether and when drift occurs), detector systems require estimating the underlying, physically meaningful transformation parameters responsible for the shift (e.g., alignment offsets, attenuation coefficients, or calibration constants) [9, 10, 18].

In detector systems, the objective is not merely to flag a change but to infer interpretable parameters that can be used to correct the data [6, 9, 11].

In this work, we propose a novel interpretable calibration approach to tackle sensor degradation as a distribution- matching problem [19]. We introduce a Wasserstein-inspired adversarial parameter-search framework [20] for unsupervised recovery of interpretable detector degradation parameters from distribution shift, and demonstrate it on a tracker misalignment toy model and a calorimeter aging task.<sup>2</sup> We thus position detector calibration and alignment within a broader class of distribution-shift and distribution-alignment problems. The proposed method is formulated and tested in an offline setting, however, it can in principle be adapted to online calibration scenarios.

- II. BACKGROUND A. HEP detector systems: tracking alignment and calorimeter aging Modern high energy physics (HEP) experiments rely on large scale, high-granularity detector systems designed to reconstruct particle trajectories and measure their energies with high precision. These detectors are typically composed of multiple sub systems, including tracking detectors and calorimeters, each optimized for specific measurement tasks.

## 2 The

implementation available wasserstein-parameter-recovery. publicly <https://github.com/tharaaali/>

a. Tracking Detectors and Alignment Tracking detectors infer charged-particle trajectories from spatial hit measurements recorded in many individual sensor elements. Precise reconstruction requires that the relative positions and orientations of these elements are known in a global reference frame: even when the intrinsic hit resolution is at the level of tens of microns, construction and assembly uncertainties can be much larger and can bias fitted track parameters if uncorrected [10, 21]. In this context, misalignment is the difference between the true installed detector geometry and the assumed geometry used in reconstruction, and it can systematically bias reconstructed track parameters through biased hit positions in the fit [22]. b. Calorimeters, Aging, and Calibration While tracking detectors measure charged-particle trajectories, calorimeters measure particle energy through total absorption [23–27]. A calorimeter is segmented into cells, each producing a signal proportional to deposited energy. Accurate energy reconstruction requires uniform response

across all cells. In an ideal detector, identical particles impinging on different cells should produce identical signals at large statistics. In practice, several effects break this uniformity [23, 24]. Manufacturing tolerances, electronic variations, temperature fluctuations, and radiation damage introduce cell to cell response variations. Over time, radiation exposure leads to material degradation and signal attenuation—commonly referred to as detector aging. Calibration [23, 24] procedures aim to restore uniformity by determining correction coefficients for each channel. Calibration is typically established using a combination of hardware procedures (e.g., electronic pulse injection, laser or radioactive sources), test beam measurements, and in-situ physics constraints from well understood processes [23, 24]. Over time, calorimeter response can drift due to aging and radiation damage in active materials and readout components, altering cell wise gains or attenuation and necessitating periodic recalibration to maintain uniformity and energy scale stability [23]. Because the detector comprises many channels, maintaining calibration becomes a large scale parameter estimation problem, where physically meaningful parameters (e.g., per-cell response coefficients) must be updated from data under realistic constraints on dedicated calibration time and resources [23, 24].

**B. Applications of Generative Machine Learning Models in High Energy Physics**  
Deep generative models have become a central component of modern machine-learning toolkits for HEP because they can learn high-dimensional detector- and event-level distributions and provide efficient sampling procedures [28]. A major driver is fast detector simulation: surrogate generative models can approximate the output of computationally expensive Monte Carlo pipelines while retaining statistical features needed for downstream analyses [28, 29]. At the same time, high fidelity deployment requires careful validation and uncertainty treatment, since mismodeling can bias physics conclusions even when one dimensional projections look reasonable [30].

Recent reviews organize state of the art surrogate models by architecture [31] class, including GAN based models [32], variational autoencoders (VAEs), normalizing flows [33], diffusion/score-based models [34], and their hybrids, as well as by the data representation and conditioning strategy used in different detectors and benchmarks [28].

In practice, GANs have been widely explored for calorimeter and detector-response emulation, with substantial gains in generation speed compared to traditional simulation chains [29, 35–39]. However, their use also raises questions of statistical fidelity and uncertainty propagation, motivating rigorous validation protocols [30].

Generative models are also used as components of analysis workflows, especially in weakly supervised and unsupervised settings. Autoencoder-based approaches, including conditional VAEs [40], have been studied for anomaly detection at the LHC [41], where deviations from the learned manifold may indicate rare or unexpected processes. More broadly, generative modeling can support model-independent searches and background modeling under limited supervision [28].

Beyond event generation and simulation acceleration, GANs have also been explored as tools for detector calibration and parameter tuning. In this context, the adversarial objective is used not to generate new samples, but to enforce consistency between observed data distributions and reference expectations [42, 43]. Related ideas have been investigated using likelihood-based methods and supervised neural networks for parameter estimation, including generator tuning and detector-response modeling [44, 45]. Track-based alignment provides a classical example of large-scale parameter inference in HEP, where geometric constants are extracted from reconstructed tracks through iterative optimization and must address scalability and parameter degeneracies [46–48].

This motivates distribution level strategies [49] for parameter inference that do not require event level correspondences.

Adversarial objectives, and in particular Wasserstein-based critics, provide a practical way to quantify and minimize mismatch between high dimensional distributions [50, 51].

In this work, we adopt this perspective by using a deterministic transformation with trainable weights corresponding to calibration or alignment parameters, optimized using a distribution-matching objective.

**C. Generative Adversarial Networks** A GAN is composed of two components trained simultaneously: a generator  $G$ , which produces samples intended to resemble data drawn from a target distribution, and a discriminator  $D$ , which evaluates whether a given sample originates from the true data distribution or from the generator. Rather than relying on an explicit likelihood, GANs learn by enforcing statistical indistinguishability between generated and real samples [32].

During training, the generator maps latent variables  $z$ , sampled from a predefined distribution  $p_z(z)$ , to synthetic samples  $G(z)$ . The discriminator receives either real samples  $x \sim p_{\text{data}}(x)$  or generated samples  $G(z)$  and outputs a scalar score reflecting its confidence in the sample's authenticity.

The two networks are optimized in opposition: the discriminator improves its ability to separate real and generated data, while the generator adapts to produce samples that are increasingly difficult to distinguish.

This adversarial interaction is commonly formulated as a minimax optimization problem [32]:

$$E_x p_{\text{data}}(x)[\log D(x)] + E_z p_z(z)[\log(1 - D(G(z)))]$$
, where the equilibrium corresponds to the generator reproducing the target data distribution. Although this formulation has proven effective in many applications, it is known to suffer from optimization instabilities and sensitivity to model design and hyperparameters, particularly when the real and generated distributions have limited overlap [20].

The Wasserstein GAN (WGAN) is proposed to address the optimization chal-

allenges inherent to the original GAN formulation by replacing the discriminator-based divergence with a distance metric grounded in optimal transport theory [20].

Instead of estimating a probability of authenticity, the discriminator is reformulated as a critic  $C$ , which assigns real-valued scores to samples. The difference between expected critic scores for real and generated data provides an estimate of the Wasserstein-1 (Earth Mover's) distance between their distributions.

The WGAN objective is given by  $E_{x \sim p_{\text{data}}(x)}[C(x)] - E_{z \sim p_z(z)}[C(G(z))]$ , where  $C$  denotes the set of 1-Lipschitz functions. Enforcing this constraint is essential for the validity of the Wasserstein distance and is commonly achieved through techniques such as weight clipping [20], gradient penalty regularization [51], or spectral normalization [52].

By providing more informative gradients and a loss function that correlates with distributional mismatch, WGAN improves training stability and robustness compared to the original GAN framework [20]. Importantly, the critic output offers a meaningful quantitative measure of how far two distributions are from one another, making WGAN a natural choice for problems centered on distribution alignment rather than explicit sample generation.

D. Alignment and calibration as distribution matching parameter inference  
Tracking alignment and calorimeter calibration share the same mathematical structure: unknown physical parameters (geometric shifts or response coefficients) must be inferred from data, while only indirect observables (hit positions or energy deposits) are accessible.

In tracking, misalignment alters the spatial distribution of residuals and reconstructed track parameters [46, 47]. In calorimetry, imperfect equalization and long term detector aging modify the statistical distribution of energy deposits and shower shapes, motivating periodic calibration and monitoring to preserve the stability of reconstructed energies [23].

In both cases, detector deformations induce a distributional shift between a nominal (reference) dataset and a distorted (changed/damaged) dataset. Traditional procedures often rely on explicit residual-based objectives [46, 53], or on external reference signals and well controlled physics samples for calorimeter calibration [23]. This motivates complementary approaches that operate directly at the level of data distributions: instead of minimizing residuals event-by-event, one may infer the underlying physical parameters by matching the full high-dimensional distribution of observables between a reference and a shifted dataset.

While Wasserstein GANs (WGANs) are most often employed for synthetic data generation, their formulation as a distribution matching framework enables alternative modes of use beyond sample synthesis [20]. In particular, the generator can be used as a deterministic transformation whose trainable weights represent transformation parameters applied to observed data [13, 14].

In this setting, the adversarial objective enforces agreement between transformed and reference distributions, and the learned parameters implicitly capture the systematic effects responsible for their mismatch. This perspective is especially relevant for detector studies, where effects such as response changes due to radiation exposure and aging [23], as well as geometric misalignment [46, 47], manifest as structured changes in measured data distributions [44, 45]. By minimizing the Wasserstein distance between transformed and nominal detector responses, the WGAN framework enables unsupervised inference of physically interpretable parameters without requiring labeled data or explicit likelihood models.

In the present work, we adopt this paradigm by embedding detector-related parameters within the generator and using the critic exclusively as a distributional consistency measure.

This shifts the role of adversarial learning from data synthesis to parameter recovery, providing a flexible and scalable approach for data-driven detector calibration and alignment studies.

**III. METHOD AND ARCHITECTURE** We propose a data-driven framework for inferring physically interpretable parameters that characterize systematic changes between two data taking regimes. The method is designed for scenarios in which a reference dataset, corresponding to an undamaged or nominal detector state, is available alongside a second dataset exhibiting distortions induced by aging, misalignment, or other systematic effects. Rather than explicitly modeling these effects through parametric likelihoods, we formulate the problem as one of distribution alignment and parameter recovery. To achieve this aim, we adopt the adversarial training principle but reinterpret the role of the generator: instead of producing new data samples, it consists of a parameterized transformation applied directly to existing data, the transformation parameters are treated as trainable model variables. From a conventional perspective, the task of the generator is to correct the damaged detector responses. Specifically, we consider two distributions: a reference distribution  $p_{ref}$ , representing undamaged or nominal detector responses, and a target distribution  $p_{chg}$ , corresponding to damaged or systematically altered data. A deterministic transformation  $T\theta(\cdot)$ , parameterized by  $\theta$ , is applied to samples drawn from one of the distributions. The objective is to adjust  $\theta$  such that the transformed distribution aligns statistically with the opposing distribution. Depending on the chosen formulation, this corresponds to matching  $T\theta(p_{chg})$  to  $p_{ref}$ , or equivalently matching  $p_{chg}$  to  $T\theta(p_{ref})$ .

Inspired by the WGAN algorithm, a critic network is employed to quantify the discrepancy between the transformed and target distributions. By minimizing the Wasserstein distance estimate, the model directly optimizes the transformation parameters that best explain the observed distributional shift. This yields a stable optimization procedure and a scalar objective that correlates

with distributional mismatch.

Fig. 1 [Figure 1: see original paper] illustrates the overall architecture. While implementation details vary between applications, the core components a parameterized transformation, an adversarial critic, and a Wasserstein based objective remain unchanged throughout this work. parameterized by  $\theta$ , where  $\theta$  has a direct physical interpretation. a. Tracker toy model In the tracker toy example, we consider three measurement planes and use a single spatial coordinate. Let  $h = (h_1, h_2, h_3) \in \mathbb{R}^3$  denote the hit coordinates on planes  $i = 1, 2, 3$ . A misalignment is injected by shifting plane  $i = 2$  by an unknown  $\Delta x_2$ . The correction transformation applies an opposite shift to the second component:

$T_{\Delta x_2}(h) = (h_1, h_2 + \Delta x_2, h_3)$ . b. Calorimeter aging Let  $e \in \mathbb{R}^N$  denote the vector of cell-wise energy deposits for an event in the changed dataset, with  $e_i$  the deposit in cell  $i$ . Aging is modeled as a cell-wise attenuation with coefficients  $A_i \in (0, 1]$  such that  $e(i)_{aged} = A_i e(i)$  true. The calibration transformation rescales cell energies to undo the attenuation:

A. Operational principle The method operates by iteratively updating the transformation parameters  $\theta$  to minimize the discrepancy between transformed and reference data distributions. At each training step, minibatches of samples are drawn from the reference and changed datasets. The transformation  $T\theta$  is applied to a set of samples, and the critic evaluates both transformed and untransformed samples to estimate the Wasserstein distance between their distributions. The transformation parameters are optimized to minimize this distance, thereby driving the transformed distribution toward statistical consistency with the reference. Importantly, no direct supervision on  $\theta$  is required, and no explicit correspondence between individual samples in the two datasets is assumed. The inferred parameters are learned from population-level distributional agreement. This adversarial parameter-search framework is symmetric by construction. Depending on the direction of the applied transformation, it can be used either to recover undamaged detector responses from altered data or to impose controlled distortions on nominal data. Consequently, the same architecture supports both calibration-oriented inference and controlled simulation studies within a unified framework.

B. Parameterized transformation model Let  $y \in \mathbb{R}^d$  denote a generic detector observation vector, with  $y_{chg}$  for changed data and  $y_{ref}$  for reference data. We introduce a deterministic transformation  $T\theta : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $T\alpha(e)_i = \alpha_i e_i = (\alpha_1, \dots, \alpha_N)$ ,  $\alpha_i \in (0, 1]$ .

Ideally,  $\alpha_i$  converges to the true aging coefficient  $A_i$ . While these two cases differ in dimensionality and physical interpretation, both can be treated uniformly within the proposed framework by treating  $\theta$  as trainable parameters of a deterministic transformation module  $T\theta$ , optimized through adversarial distribution matching.

C. Adversarial objective We introduce a critic  $C\phi$  parameterized by  $\phi$ , constrained to be 1-Lipschitz. The WGAN objective for aligning corrected changed

data with the reference data is  $E_y \text{pref} [C\phi(y)] - E_y \text{pchg} [C\phi(T\theta(y))]$ , where  $C$  denotes the set of 1-Lipschitz functions. The critic is trained to maximize the objective, while  $\theta$  is trained to minimize it, driving  $T\theta(\text{pchg})$  toward  $\text{pref}$ .

In both cases, the critic is trained to maximize the objective (increase the estimated Wasserstein distance), while  $\theta$  is trained to minimize it (reduce the distance by improving the transformation).

Fig. 1. Architecture of the proposed WGAN-inspired adversarial parameter-search framework.

Algorithm 1 Adversarial parameter search  
 Require: Reference dataset  $D_{\text{ref}}$ , changed dataset  $D_{\text{chg}}$ , critic  $C\phi$ , transformation  $T\theta$ , clipping parameter  $c$ , number of critic updates  $n_{\text{critic}}$   
 Ensure: Learned parameters  $\theta$   
 1: Initialize critic parameters  $\phi$   
 2: Initialize transformation parameters  $\theta$  (e.g.,  $\alpha_i \leftarrow 1$  for calorimeter;  $\Delta x_2 \leftarrow 0$  for tracker)  
 3: for each training iteration do  
 for  $n_{\text{critic}}$  critic updates do  
 Sample minibatch  $x_{\text{ref}} \sim D_{\text{ref}}$  Sample minibatch  $x_{\text{chg}} \sim D_{\text{chg}}$  Correct changed samples:  $\tilde{x} \leftarrow T\theta(x_{\text{chg}})$  Critic loss (to minimize by gradient descent):  
 (cid:16)  $LC = -E[C\phi(x_{\text{ref}})] - E[C\phi(\tilde{x})]$  (cid:17) Update  $\phi$  by descending  $\nabla\phi LC$  Clip critic weights:  $\phi \leftarrow \text{clip}(\phi, -c, c)$  end for  
 Sample minibatch  $x_{\text{chg}} \sim D_{\text{chg}}$  Correct changed samples:  $\tilde{x} \leftarrow T\theta(x_{\text{chg}})$  Generator/transformation loss:

$LT = -E[C\phi(\tilde{x})]$  Update  $\theta$  by descending  $\nabla\theta LT$  16: end for 17: return  $\theta$  IV.

EXPERIMENTS A. Case study I: Tracker Toy Model As a first validation of the proposed adversarial parameter-search framework, we consider a simplified tracking-detector toy model. The purpose of this study is not to provide a realistic alignment solution, but rather to demonstrate that the method can reliably recover a known geometric misalignment parameter in a controlled and interpretable setting.

The toy detector consists of three tracking layers (Fig. 2 [Figure 2: see original paper]) Fig. 2. Tracker toy model, a misalignment is injected by shifting the second plane by  $\Delta x_2$ , producing a systematic change in hit patterns (and residual distributions) relative to the nominal geometry. that record one-dimensional hit coordinates. Two types of trajectories are simulated: straight tracks and inclined tracks.

In the reference configuration, all three planes are perfectly aligned, yielding consistent geometric relations across layers.

Two datasets are constructed: • Reference dataset ( $D_{\text{ref}}$ ): generated with all tracker layers at their nominal positions. • Changed dataset ( $D_{\text{chg}}$ ): generated by shifting the second tracking plane by a constant offset  $\Delta x_2$  while keeping the remaining planes fixed. a. Generator as a single parameter offset model In this setup, we have a reference (undamaged) distribution  $\text{pref}$  and a changed/distorted distribution  $\text{pchg}$ . Samples from the changed dataset are passed through the generator, which in this case is as a deterministic correction transform  $T\theta$  with a single trainable scalar parameter  $\theta = (\text{cid:100})\Delta x_2$ . Each sample is a hit tuple  $h = (h_1, h_2, h_3) \in \mathbb{R}^3$ , where  $h$  is the measured x coord-

dinate on plane  $\{1, 2, 3\}$ . The transformation acts on the second coordinate and outputs a corrected sample  $\tilde{h} = T(\Delta x_2)$  ( $h = (h_1, h_2 + \Delta x_2, h_3)$ ).

Refenced Data Transformed DataCriticCalculatewassersteindistanceGeneratorTransition Parameter  $\theta$  Changed DataTracker 1Tracker 2Tracker 3Straight track  $\rightarrow x_1 x_2 x_3$ Inclined track  $\rightarrow x_1 \neq x_2 \neq x_3$ Tracker 1Tracker 3Straight track  $\rightarrow x_1 x_2 x_3$ Inclined track  $\rightarrow x_1 \neq x_2 \neq x_3$ Tracker 2 $\Delta x_2$

$T(\Delta x_2)$  Training therefore directly yields  $\tilde{h} = T(\Delta x_2)$  as the inferred mis- alignment correction. The critic  $C\phi$  is a neural network trained with the WGAN objective to maximize the separation between reference samples  $h_{ref}$  and corrected samples  $\tilde{h} = T(\Delta x_2)$ , providing an estimate of the Wasserstein distance. The parameter  $\Delta x_2$  is updated in the opposite direction to minimize this estimated distance, thereby aligning  $\tilde{h}$  with  $h_{ref}$ . At convergence,  $\Delta x_2$  approaches  $-\Delta x_2$ , since this value best aligns the corrected and reference hit distributions. To study the sensitivity to a systematic distortion, a fixed shift value of  $\Delta x_2 = 0.01$  is applied to the detector coordinates. Each dataset contains only hit coordinates and does not include labels or track level correspondences between the reference and changed samples, ensuring a fully unsupervised setup based exclusively on distribution-level information. To assess robustness against measurement uncertainty, independent Gaussian noise is added to the hit coordinates in the changed dataset. The noise standard deviation  $\sigma$  is varied over several orders of magnitude, ranging from negligible detector noise to values comparable to the applied shift.

B. Case study II: Calorimeter Aging This case study investigates whether the proposed adversarial parameter search framework can recover cell wise aging effects in a fully unsupervised manner by exploiting distributional differences between undamaged and aged calorimeter responses.

Simulation and data representation Monte Carlo simulations are performed using the GEANT4 toolkit [54].

We generate 105 events of single  $\pi^-$  mesons with an exponential kinetic-energy spectrum and a flat transverse impact distribution; the particle momentum is oriented perpendicular to the upstream face of the calorimeter. The detector geometry corresponds to a highly granular sampling calorimeter with alternating 20 mm lead absorber plates and 5 mm polystyrene scintillator tiles, segmented into a transverse cell grid. To reduce edge effects and shower leakage, we restrict the analysis to the central  $16 \times 16$  cells in each sensitive layer.

Fig. 3 [Figure 3: see original paper] shows the corresponding transverse hit-activity map aggregated over the dataset.

Each event is represented as a 3D tensor  $X \in \mathbb{R}^{L \times H \times W}$ , where  $L$  is the number of sensitive layers and  $(H, W)$  correspond to the selected transverse cell grid;  $X_{u,v}$  is the energy deposit in cell  $(u, v)$ . We denote by  $h_{ref}$  and  $h_{chg}$  the

distributions of undamaged and aged events, respectively. b. Aging model and datasets Two complementary datasets are constructed: • Undamaged dataset (Dref): events  $X_{pref}$  simulated without aging effects. • Changed dataset (Dchg): events  $X_{chg}$  obtained by applying cell-wise signal attenuation to emulate radiation damage.

Fig. 3. Hit activity map in the central  $16 \times 16$  cell region (counts summed over events).

We model aging through a cell-wise coefficient field  $A \in (0, 1]^{L \times H \times W}$ , so that an undamaged event  $X$  is mapped to an aged event by element-wise multiplication:

$X_{chg} = T_{age} A(X) = A \odot X$ ,  $(X_{chg})_{u,v} = A_{u,v} X_{u,v}$ , where  $\odot$  denotes the Hadamard product. In our simulation setup, the coefficients are generated as  $A_i = f(n_i) + N(0, \epsilon)$ , where  $i$  indexes calorimeter cells,  $n_i$  is the total hit count accumulated by cell  $i$  over the dataset,  $f(\cdot)$  is a monotone decay model assigning smaller coefficients to more frequently activated cells, and  $\epsilon$  controls cell-to-cell miscalibration noise (unless stated otherwise,  $\epsilon = 0.01$ ). Fig. 4 [Figure 4: see original paper] illustrates the effect of aging by comparing the per event total deposited energy,  $E_{sum} = \sum_i E(i)$ , for undamaged and aged samples, where a noticeable shift is observed. c. Generator as a calibration field The generator is the parameterized transformation  $T_\theta$  whose trainable parameters are the calibration coefficients themselves:  $\theta = A \in (0, 1]^{L \times H \times W}$ .

Operationally,  $A$  is implemented as a learnable tensor matching the calorimeter granularity (a 3D parameter field). Given an aged event  $X_{chg}$ , the generator applies an element wise correction and outputs a calibrated event  $X = T_\theta(A)(X_{chg}) = X_{chg} \oslash A$ ,  $(X)_{u,v} = (X_{chg})_{u,v} / A_{u,v}$  where  $\oslash$  denotes element wise division. To ensure numerical stability, we constrain  $A_{u,v} \in [\alpha_{min}, 1]$  via clamping after each update. Fig. 5 [Figure 5: see original paper]. Absolute error of the estimated tracker shift  $|\Delta x - \hat{\Delta x}|$  as a function of the Gaussian noise standard deviation  $\sigma$ . Accurate recovery is observed for low and moderate noise levels, while the estimation error increases rapidly once the noise becomes comparable to the applied shift.

Fig. 6 [Figure 6: see original paper]. RMSE between inferred and injected aging coefficients. expected from cell wise attenuation. At the same time, the distribution shape changes modestly, indicating that a one dimensional global observable alone provides limited sensitivity to spatially varying aging and that additional information must be exploited from the multi cell response patterns.

The training dynamics of the WGAN-inspired calibration model are summarized in Fig. 6, which shows the RMSE between the predicted and injected aging coefficients as a function of epoch. The error decreases during training and reaches a stable plateau after roughly 650-700 epochs, indicating consistent convergence under the chosen optimization setup.

Fig. 7 [Figure 7: see original paper] compares the inferred coefficients to the

injected values on a cell by cell basis. The estimates follow a clear positive correlation, with most cells clustered near the diagonal, showing that the model recovers the dominant aging trend from distribution-level constraints. The largest deviations are observed for the most strongly attenuated cells, where reduced signal amplitudes and the injected stochastic perturbations limit the information available for parameter identification. We do not interpret these deviations as a failure of calibration, but rather as an expected limitation of inferring fine grained, per cell fluctuations in an un-

The critic network  $C\phi$  is trained with the WGAN objective to estimate the Wasserstein distance between the corrected changed distribution  $T$  (cid:98)A(pchg) and the reference distribution  $pref$ . Training alternates between updating the critic to maximize the estimated Wasserstein discrepancy and updating the generator parameters (cid:98)A to minimize it. At convergence, (cid:98)A provides an estimate of the underlying aging field and thus yields cell-wise calibration coefficients directly as learned generator parameters.

Although aging primarily affects the global energy scale, global energy-sum projections alone are insufficient for identifying cell wise effects. Instead, the information that constrains (cid:98)A is encoded in the high-dimensional structure of cell level depositions and their correlations, which are matched implicitly by the critic during adversarial training. The overall architecture of the WGAN-inspired calibration model is shown in Fig. 1.

V. RESULTS AND DISCUSSION a. Tracker toy model Fig. 5 reports the absolute estimation error  $|(cid:99)\Delta x - \Delta x|$  as a function of the Gaussian noise standard deviation  $\sigma$  for injected misalignment magnitudes. For low and moderate noise levels, the inferred shift remains close to the true value, with errors below 10–3. As  $\sigma$  approaches the scale of the injected shift, the error increases rapidly, consistent with the loss of identifiable geometric information in the hit tuples. Overall, the toy study demonstrates that the proposed distribution matching objective can recover a simple, physically interpretable alignment parameter when the distortion is not dominated by measurement noise. b. Calorimeter aging We first examine the effect of simulated aging on the calorimeter response. Fig. 4 compares the total reconstructed energy (Esum) distributions for the undamaged and damaged datasets. The damaged distribution exhibits an overall suppression of the energy scale, as

020406080100Esum [GeV]02000400060008000100001200014000Entries / binEnergy Sum Distribution - Undamaged vs DamagedUndamagedDamaged10–510–410–310–210–1Noise std \$ 0.0000.0010.0020.003|  $x^2 - \sigma^2$ |0200400600800Epoch / Step0.020.030.040.050.060.070.080.090.10RMSERMSE vs Epoch

8 Fig. 7. Cell wise comparison of true and inferred aging coefficients.

Fig. 8 [Figure 8: see original paper]. Energy-sum distributions for undamaged and calibrated calorimeter responses. = Echg Eref pervised setting with finite statistics and noisy coefficients.

The proposed method achieves a mean reconstruction error of  $\text{RMSE} = 0.0176 \pm 0.0005$  and a mean coefficient of determination of  $R^2 = 0.896 \pm 0.007$ . A simple per cell baseline that estimates aging as the ratio of mean responses, (cid:98)Amean , yields a worse agreement with the injected coefficients of with  $\text{RMSE} = 0.0562 \pm 0.0008$ . This comparison indicates that the proposed distribution level adversarial calibration extracts significantly more information from the full multi cell response than can be captured by matching only per cell mean energies. To assess the practical effect of the inferred parameters, we apply the learned correction to the damaged dataset and recompute the energy sum distribution. Fig. 8 shows that the calibrated distribution moves closer to the undamaged reference over the full energy range.

This indicates that the recovered coefficients compensate the dominant radiation-induced attenuation at the level of a global summary observable, while not guaranteeing perfect cell by Fig. 9 [Figure 9: see original paper]. Effect of stochastic coefficient noise  $\epsilon$  on the RMSE of aging-coefficient estimation. cell recovery of the injected stochastic component.

Finally, we evaluate sensitivity to uncertainty in the aging coefficients by varying the standard deviation  $\epsilon$  of the additive Gaussian perturbation in the aging model over  $\epsilon \in [0, 0.05]$ .

As shown in Fig. 9, the RMSE between inferred and injected coefficients increases monotonically with  $\epsilon$ . This trend is expected: larger cell to cell stochastic variations reduce the identifiability of individual attenuation factors from distribution matching alone, because multiple coefficient configurations can yield similar aggregate shower statistics.

In the deterministic case ( $\epsilon = 0$ ), the model attains its lowest error,  $\text{RMSE} = 0.01573$ , which reflects the achievable accuracy for recovering the global aging trend  $f(\text{ni})$  under the given simulation and training setup. As  $\epsilon$  increases, the attainable accuracy degrades gradually, reaching  $\text{RMSE} = 4.7 \times 10^{-2}$  at  $\epsilon = 0.05$ .

**VI. CONCLUSION** We have presented a Wasserstein-inspired adversarial framework for unsupervised inference of sensor related transformation parameters, formulated as a distribution matching problem between reference and changed data regimes. By reinterpreting the generator as a deterministic transformation whose trainable weights correspond to physically meaningful parameters, the approach shifts the use of adversarial learning from sample synthesis to parameter recovery. We validated the method in two complementary settings. In a simplified tracking toy model, the framework recovers injected geometric shifts in a low dimensional and fully controlled scenario, serving as a proof of concept for alignment style parameter inference. We then studied calorimeter aging using high granularity Geant4-simulated data with cell wise signal attenuation. In this setup, the inferred aging coefficients show good agreement with the injected values and,

when applied as a calibration correction, improve the consistency between the calibrated and undamaged energy response distributions, without requiring labeled calibration targets or event-level correspondences. The results also illustrate a key limitation of purely distribution based inference: as channel to channel stochastic variations increase and become comparable to the underlying systematic effect, the identifiability of individual parameters deteriorates due to reduced discriminative information in the observed distributions. This suggests that adversarial calibration is best viewed as a complementary technique that reduces reliance on supervision, rather than a replacement for established workflows. Future work will focus on extending the framework to time dependent degradation, more realistic detector geometries and readout effects, and hybrid strategies that incorporate conventional calibration constraints or external information.

[1] Sulemana Fatoama Abdulai, Tarek Zayed, Ibrahim Yahaya Wuni, Maxwell Fordjour Antwi-Afari, and Abdul-Mugis Yusuf. Cross-industry review of autonomous alignment technologies: Adaptation potential for modular construction. *Journal of Cleaner Production*, 495:145101, 2025. [2] Philip Stahmann, Maximilian Nebel, and Christian Janiesch.

Ai-based real-time anomaly detection in industrial engineering:

A structured literature review, taxonomy, and research agenda.

*Computers & Industrial Engineering*, 207:111236, 2025. [3] Long Chen, Chenbin Xia, Zhehui Zhao, Haoran Fu, and Yunmin Chen. Ai-driven sensing technology: Review. *Sensors*, 24(10), 2024. [4] T Golling. Alignment of the Silicon Tracking Detector using Survey Constraints. Technical report, CERN, Geneva, 2006. [5] Erik Butz. Calibration, alignment and long-term performance of the cms silicon tracking detector, May 2009. [6] Tanmay Mudholkar and on behalf of the CMS Collaboration.

Cms ecal monitoring and calibration in lhc run ii. *Journal of Physics: Conference Series*, 1162(1):012002, jan 2019. [7] G. Aad, B. Abbott, J. Abdallah, et al. ATLAS data quality operations and performance for 2015-2018 data-taking. *Journal of Instrumentation*, 15(04):P04003, 2020. [8] Tapio Lampén and (for the CMS Collaboration). Cms silicon strip tracker calibration workflow and tools. *Journal of Physics: Conference Series*, 331(3):032044, dec 2011. [9] M Verducci. ATLAS conditions database experience with the LCG COOL conditions database project. *J. Phys.: Conf. Ser.*, 119:042031, 2008. [10] Ana Ovcharova. Track-Based Alignment of the Inner Detector of ATLAS. Technical report, CERN, Geneva, 2012. [11] PJ Laycock, MA Chelstowska, TC Donzelmann, J Guenther, A Nairz, R Nikolaidou, E Shabalina, J Strandberg, A

Taffard, S Wang, and on behalf of the ATLAS Collaboration. Atlas data preparation in run 2. *Journal of Physics: Conference Series*, 898(4):042050, oct 2017. [12] B Pinto, A Amorim, P Pereira, M Elsing, R Hawkings, J Schieck, S Garcia, A Schaffer, H Ma, and A Anjos. Align- Journal of ment data streams for the atlas inner detector.

Physics: Conference Series, 119(2):022030, jul 2008. [13] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Ger- main, Hugo Larochelle, Franc\_ois Laviolette, Mario March, and Victor Lempitsky. Domain-adversarial training of neural net- works. *Journal of Machine Learning Research*, 17(59):1-35, nal reference signals, with the aim of reducing recalibration overheads in long running experiments.

VII. ACKNOWLEDGMENTS The work was supported by the grant for research cen- ters in the field of AI provided by the Ministry of Economic Development of the Russian Federation in accordance with the agreement 000000C313925P4E0002 and the agreement with HSE University № 139-15-2025-009. The computation for this research was performed using the computational re- sources of HPC facilities at HSE University. [14] Jun- Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros.

Unpaired image-to-image translation using cycle-consistent adversarial networks. In 2017 IEEE International Conference on Computer Vision (ICCV), pages 2242-2251, 2017. [15] Kyle Cranmer, Johann Brehmer, and Gilles Louppe. The fron- tier of simulation-based inference. *Proceedings of the National Academy of Sciences of the United States of America*, 117(48):30055-30062, December 2020. [16] Gilles Louppe, Joeri Hermans, and Kyle Cranmer. Adversarial variational optimization of non-differentiable simulators.

The 22nd international conference on artificial intelligence and statistics, pages 1438-1447. PMLR, 2019. [17] Joaquin Quinonero-Candela, Masashi Sugiyama, Anton Schwaighofer, and Neil D. Lawrence, editors. *Dataset Shift in Machine Learning*. Neural Information Processing Series.

The MIT Press, Cambridge, MA, 2008. Hardcover edition. [18] Fabian Hinder, Valerie Vaquet, and Barbara Hammer. One or two things we know about concept drift—a survey on monitor- ing in evolving environments. part a: detecting concept drift.

Frontiers in Artificial Intelligence, Volume 7 - 2024, 2024. [19] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Ger- main, Hugo Larochelle, Franc\_ois Laviolette, Mario March, and Victor Lempitsky. Domain-adversarial training of neural net- works. *Journal of machine learning research*, 17(59):1-35, [20] Martin Arjovsky, Soumith Chintala, and L´eon Bottou. Wasser- stein generative adversarial networks. In *International confer- ence on machine learning*, pages 214-223. Pmlr, 2017. [21] ATLAS Collaboration. ATLAS inner detector. Technical design report. ATLAS. CERN, Geneva, 1997. [22] ATLAS Collab- oration. ATLAS detector and physics perfor- mance. Technical design report.

ATLAS. CERN, Geneva, [23] C. W. Fabjan and F. Gianotti. Calorimetry for particle physics.

Reviews of Modern Physics, 75:1243–1286, 2003. [24] S. S. Korpachov and M. V. Chadeeva. Study of the uniformity of the response of scintillation tiles for high-granularity calorimeters. Brief Communications on Physics of Lebedev Institute, (10):52–56, 2018. [25] A. A. Alves, E. Ben-Haim, D. Bettinardi, et al. The lhcb detector at the lhc. Journal of Instrumentation, 3:S08005, 2008. [26] K. Aamodt, A. Abrahantes Quintana, R. Acosta, et al. The alice experiment at the cern lhc. Journal of Instrumentation, 3:S08002, 2008. [27] G. Aad, B. Abbott, J. Abdallah, et al. The atlas experiment at the cern large hadron collider. Journal of Instrumentation, 3:S08003, 2008. [28] Baran Hashemi and Claudius Krause. Deep generative models for detector signature simulation: A taxonomic review. Reviews in Physics, 12:100092, 2024. [29] Luke de Oliveira, Michela Paganini, and Benjamin Nachman.

Learning particle physics by example: Location-aware generative adversarial networks for physics synthesis. Computing and Software for Big Science, 1(1):4, 2017. [30] Anja Butter, Sascha Diefenbacher, Gregor Kasieczka, Benjamin Nachman, and Tilman Plehn. GANplifying event samples. SciPost Phys., 10:139, 2021. [31] Diederik P Kingma and Max Welling. Auto-encoding variational bayes, 2022. [32] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. Advances in neural information processing systems, 27, 2014. [33] Ivan Kobyzev, Simon J.D. Prince, and Marcus A. Brubaker.

Normalizing flows: An introduction and review of current methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(11):3964–3979, 2021. [34] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models.

In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, Advances in Neural Information Processing Systems, volume 33, pages 6840–

### **6851. Curran Associates, Inc., 2020.**

[35] Su Yeon Chang, Steven Herbert, Sofia Vallecorsa, Elias F Combarro, and Ross Duncan. Dual-parameterized quantum circuit gan model in high energy physics. In EPJ Web of Conferences, volume 251, page 03050. EDP Sciences, 2021. [36] Fedor Ratnikov, Alexander Rogachev, Sergey Mokhnenko, Artem Maevskiy, Denis Derkach, Adam Davis, Nikita Kazeev, Lucio Anderlini, Matteo Barbetti, and Benedetto Gianluca Siddi. A full detector description using neural network driven simulation. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1046:167591, 2023. [37] Lucio Anderlini, Matteo Barbetti, Simone Capelli, Gloria Corti, Adam Davis, Denis Derkach, Nikita Kazeev, Artem Maevskiy, Maurizio Martinelli, Sergei Mokhnenko, Benedetto Gianluca

Siddi, and Zehua Xu. The LHCb ultra-fast simulation option, Lamarr: design and validation. EPJ Web of Conferences, 295:03040, May 2024. Proceedings of the 26th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2023). [38] Konstantin T Matchev, Alexander Roman, and Prasanth Shyamsundar. Uncertainties associated with gan-generated datasets in high energy physics. SciPost Physics, 12(3):104, [39] Su Yeon Chang, Sofia Vallecorsa, Elías F Combarro, and Federico Carminati. Quantum generative adversarial networks in a continuous-variable architecture to simulate high energy physics detectors. arXiv preprint arXiv:2101.11132, 2021. [40] Adrian Alan Pol, Victor Berger, Cecile Germain, Gianluca Cerminara, and Maurizio Pierini. Anomaly detection with In 2019 18th IEEE international variational autoencoders. International conference on machine learning and applications (ICMLA), pages 1651-1657. IEEE, 2019. [41] Giulia Lavizzari, Giacomo Boldrini, Simone Gennai, and Pietro Govoni. A variational autoencoder for model independent searches of new physics at lhc. International Conference on Image Analysis and Processing, pages 159-169.

Springer, 2023. [42] Markus Stoye, Johann Brehmer, Gilles Louppe, Juan Pavez, and Kyle Cranmer. Likelihood-free inference with an improved arXiv preprint arXiv:1808.00973, cross-entropy estimator. [43] Tigran Ramazyan, Mikhail Hushchyn, and Denis Derkach.

Global optimisation of black-box functions with generative models in the wasserstein space. In ECAI 2024, pages 2394-

## 2401. IOS Press, 2024.

[44] Risma Febrianti, Yekti Widyaningsih, and Saskya Soemartojo.

The parameter estimation of logistic regression with maximum likelihood method and score function modification. In Journal of physics: Conference series, volume 1725, page 012014. IOP Publishing, 2021. [45] Anders Andreassen, Shih-Chieh Hsu, Benjamin Nachman, Natchanon Suaysom, and Adi Suresh. Parameter estimation using neural networks in the presence of detector effects. Physical Review D, 103(3):036001, 2021. [46] ATLAS Collaboration. Alignment of the ATLAS Inner Detector Tracking System with 2010 LHC proton-proton collisions  $\sqrt{s} = 7$  TeV. Technical report, CERN, Geneva, 2011. [47] Peter Brückman, A Hicheur, and S J Haywood. Global chi2 approach to the Alignment of the ATLAS Silicon Tracking Detectors. Technical report, CERN, Geneva, 2005. [48] A Bocci and W Hulsbergen. TRT Alignment For SR1 Cosmics and Beyond. Technical report, CERN, Geneva, 2007. [49] Zhiwen Zuo, Lei Zhao, Ailin Li, Zhizhong Wang, Haibo Chen, Wi Xing, and Dongming Lu. Dual distribution matching gan.

Neurocomputing, 478:37-48, 2022. [50] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks.

Yee Whye Teh, editors, Proceedings of the 34th International Conference on Machine Learning, volume 70 of Proceedings of Machine Learning Research, pages 214-223. PMLR, 06-11 Aug 2017. [51] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. Improved training of Wasserstein GANs. Advances in neural information processing systems, 30, 2017. [52] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. arXiv preprint arXiv:1802.05957, 2018. [53] ATLAS Collaboration. Study of alignment-related systematic effects on the ATLAS Inner Detector tracking. Technical report, CERN, Geneva, 2012. [54] S. Agostinelli et al. GEANT4—a simulation toolkit. Nucl. Instrum. Meth. A, 506:250–303, 2003.

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv – Machine translation. Verify with original.*