أكاديمية كاوست
KAUST ACADEMY

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

How to Regularize Autoencoders for Real-World Engineering Data?
A Systematic Study Across Data Regimes and Latent Bottlenecks

Bandar Mansour BinHazza, Raed Sulieman Aljohany, Eyad Sulaiman Alayed,
Mentor: Dr. Jad Mounayer

Abstract

Engineering companies often operate with small, localized datasets because high-fidelity simulations, experiments, and field measurements are expensive and typically cover only a narrow operating region. In such low-data settings, an autoencoder may reconstruct observed states accurately while still producing latent interpolations that leave the valid physical solution manifold. We therefore compare ten autoencoder regularization strategies across four synthetic datasets designed to mimic localized engineering behavior while retaining known physical generators. Generator-based relative (L_2) errors are used as the primary interpolation-quality metric instead of Fréchet Inception Distance (FID). For the Gaussian, Thin Plate, and Shallow Water datasets, interpolation quality is evaluated by the distance to the nearest valid physical state, whereas Reaction-Diffusion is assessed using a stricter physical-parameter-path comparison. The best-performing regularizer is dataset-dependent: Contractive leads in the single available Gaussian run, while Denoising, IRMAE, and L1 lead the five-seed Thin Plate, Shallow Water, and Reaction-Diffusion experiments, respectively. Reconstruction rankings differ substantially from physical interpolation rankings; Sparse achieves the lowest reconstruction error on three datasets while remaining comparatively weak in physical validity. These results show that reconstruction accuracy is necessary but insufficient for evaluating autoencoders in low-data engineering settings, that no regularization strategy performs best universally, and that explicit physics-based validation is essential when assessing the quality of learned latent spaces.

Date: 23 August 2026

Contents

1 Introduction **2**

2 Related Work **2**

3 Methodology **3**

3.1 Shared autoencoder backbone and training . . . 3

3.2 Regularization strategies . . . 3

3.3 Datasets and low-data protocols . . . 3

3.4 Interpolation metrics . . . 5

3.5 Why FID became secondary . . . 5

4 Results **5**

4.1 Gaussian . . . 6

4.2 Thin Plate . . . 7

4.3 Shallow Water . . . 8

4.4 Reaction-Diffusion . . . 9

5 Discussion **10**

5.1 No universal winner . . . 10

5.2 Reconstruction does not certify interpolation . . . 10

5.3 Local smoothness and learned rank . . . 10

5.4 Hard rank must be reported as implemented . . . 10

5.5 Physical metrics improve scientific relevance . . . 11

5.6 Limitations . . . 11

6 Conclusion **11**

A Complete Numerical Results **11**

A.1 Gaussian . . . 12

A.2 Thin Plate . . . 13

A.3 Shallow Water . . . 14

A.4 Reaction-Diffusion . . . 15

References **15**

1 Introduction

An autoencoder learns an encoder E and decoder D so that $D(E(x))$ approximates an observed field x. In reduced-order modeling, design exploration, and scientific generation, endpoint reconstruction is only part of the requirement. The latent space must also behave sensibly between observations. For endpoints x_1 and x_2, the standard test is

$$z(\alpha) = (1-\alpha)E(x_1) + \alpha E(x_2), \qquad \hat{x}(\alpha) = D(z(\alpha)). \tag{1}$$

A model can reconstruct x_1 and x_2 well and still generate cross-fades, disconnected structures, or fields that no valid physical parameter setting can produce.

In many companies, engineering data are scarce and localized. High-fidelity simulations, laboratory experiments, and sensor campaigns are costly, so the available observations are often concentrated around a narrow operating envelope instead of covering the full physical state space. This industrial constraint motivates our study. We designed four synthetic datasets that mimic localized engineering behavior in a controlled low-data regime while retaining known generators. The generators let us test whether a model learns a coherent local physical manifold from limited observations rather than only memorizing the measured endpoints.

We compare published autoencoder regularizers using a common convolutional backbone, reconstruction objective, and training schedule, while retaining method-specific latent and rank configurations required by the respective regularization mechanisms. The four datasets span a translated Gaussian bump, a localized Thin Plate response, a coupled Shallow-Water wave field, and a localized Gray–Scott Reaction-Diffusion pattern. Their known physical parameters permit generator-based validation of latent interpolation.

The study asks three questions: which regularizers reduce reconstruction error under limited data; which regularizers keep decoded latent lines close to valid physical states; and how tight bottlenecks, overcomplete representations, and rank-control mechanisms behave when effective dimension and implemented rank budget differ.

Takeaway

For scarce, localized engineering data, reconstruction and physical interpolation are different axes. The best physical method changes across all four datasets, and the model with the best reconstruction is never the model with the best interpolation in our experiments.

2 Related Work

The compared regularizers originate from distinct primary mechanisms. An L1 penalty encourages coefficient sparsity [1], whereas quadratic weight decay limits neural-network weight magnitude [2]. Contractive autoencoders penalize the encoder Jacobian [3], and denoising autoencoders reconstruct clean inputs from corrupted observations [4]. Our Sparse implementation instead applies a Kullback–Leibler penalty toward a target mean latent activation; related peer-reviewed sparse-autoencoder work enforces sparsity by retaining only the k largest hidden activations [5]. These are related sparsity principles but not identical objectives. Variational autoencoders impose a probabilistic latent model and prior through variational inference [6].

The rank-oriented methods also have separate, directly matched sources. IRMAE inserts trainable linear layers whose optimization implicitly favors a low-effective-rank code [7]. LoRAE explicitly penalizes the nuclear norm of a learned latent projection [8]. RRAE uses a prescribed-rank singular-value decomposition, with strong and weak formulations described by Mounayer

et al. [9]. The broader taxonomy was cross-checked against the peer-reviewed review by Chen and Guo [10].

FID was introduced as a distribution-level feature statistic for generated images [11]. Its natural-image feature space and high-dimensional covariance estimates are poorly matched to the small scientific datasets in this project. The final evaluation therefore treats FID as a secondary diagnostic and prioritizes simulator-based relative L2 errors.

3 Methodology

3.1 Shared autoencoder backbone and training

We implemented every experiment with the same convolutional encoder-decoder. Two stride-2 convolutions reduce 64×64 fields to 16×16 feature maps before a fully connected latent mapping. The decoder mirrors this path with a fully connected expansion and transposed convolutions. Gaussian and Thin Plate have one channel, Shallow Water has three channels, and Reaction-Diffusion has two. Outputs are linear because the scientific fields are continuous and may be signed.

We trained every model with Adam at learning rate 10^{-3}, batch size 32, and 1000 fixed epochs. No early stopping or best-checkpoint restoration is used. The reported reconstruction value is the final-epoch validation relative L2 error,

$$\mathcal{L}_{\text{rec}} = 100 \frac{\|x - \hat{x}\|_2}{\|x\|_2}. \tag{2}$$

3.2 Regularization strategies

Table 1 Regularization mechanisms evaluated in our experiments.

Technique	Latent family	Mechanism
Vanilla	Tight	No additional regularization
L1	Tight	L1 penalty on trainable parameters
L2	Tight	Squared L2 penalty on trainable parameters
Contractive	Tight	Encoder Jacobian penalty
Sparse	Overcomplete	KL sparsity penalty on sigmoid latent activations
Denoising	Tight	Reconstruction from corrupted input
VAE	Tight	Gaussian latent distribution and KL prior
IRMAE	Overcomplete	Three inserted trainable linear maps
LoRAE	Overcomplete	Nuclear norm of an inserted linear map
RRAE	Overcomplete	SVD-based hard rank projection

The method-specific settings were held fixed across datasets and are listed explicitly in Table 2. Because the reconstruction term is a percentage-scale relative L_2 loss, these coefficients should be interpreted on that scale rather than as coefficients for an unscaled mean-squared error.

For RRAE, both the minibatch rank projection used during training and the finalized inference basis are computed through each notebook's stable-SVD helper. In the updated Gaussian notebook, this helper uses `torch.linalg.svd`, requests the `gesvd` driver on CUDA, and falls back to CPU LAPACK if needed.

3.3 Datasets and low-data protocols

The Gaussian fields are fixed-width bumps with $\sigma = 0.2$. Thin Plate samples load center, positive magnitude, and width, but per-field normalization cancels magnitude exactly; its normalized

Table 2 Method-specific implementation settings used in every dataset.

Technique	Implemented setting
L1	Parameter penalty weight $\lambda_{L1} = 10^{-3}$
L2	Squared-parameter penalty weight $\lambda_{L2} = 10^{-2}$
Contractive	Encoder-Jacobian penalty weight $\lambda_c = 10$
Sparse	Target mean activation $\rho = 0.05$; KL sparsity weight 1.0
Denoising	Independent masking corruption with probability 0.25
VAE	KL weight $\beta = 10^{-1}$
IRMAE	Three bias-free linear maps in the 200-dimensional latent space
LoRAE	Nuclear-norm penalty weight $\lambda_{\mathrm{LoRAE}} = 1.0$
RRAE	Hard ranks 2, 16, 9, and 3 for Gaussian, Thin Plate, Shallow Water, and Reaction-Diffusion, respectively

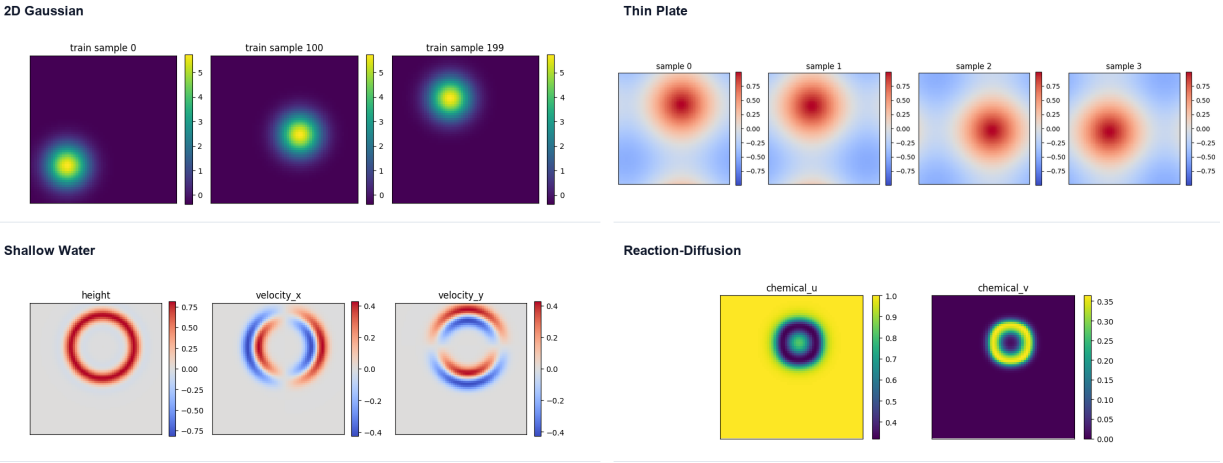


Figure 1 Representative fields from the four current synthetic datasets. Shallow Water uses three coupled channels and Reaction-Diffusion uses the complete two-chemical state.

Table 3 Low-data protocol settings used in our experiments.

Dataset	Channels	Split	Varying or effective parameters	Tight/large
Gaussian	1	200/100	2: center x, y	2/200
Thin Plate	1	85/15	3 effective: load x, y, width	4/200
Shallow Water	3	638/112	4: center x, y, time, amplitude	9/200
Reaction-Diffusion	2	170/30	3: seed x, y, radius	3/200

image manifold therefore has three effective parameters. Shallow Water varies center, elapsed time, and amplitude across three coupled channels. Reaction-Diffusion varies seed center and seed radius while fixing `simulation_steps`=380, so time is not a fourth parameter.

3.4 Interpolation metrics

Each dataset uses 50 fixed endpoint pairs and nine interior coefficients $\alpha = 0.1, \ldots, 0.9$, giving 450 decoded cases per model. Gaussian, Thin Plate, and Shallow Water fit each decoded interpolation to the nearest state produced by the known generator $G(p)$,

$$p^* = \arg\min_p \|\hat{x} - G(p)\|_2, \qquad e_{\text{manifold}} = 100\,\frac{\|\hat{x} - G(p^*)\|_2}{\|G(p^*)\|_2 + \varepsilon_d}. \tag{3}$$

The metric stabilizer is implemented as $\varepsilon_d = 10^{-12}$ for Gaussian and 10^{-8} for Thin Plate, Shallow Water, and Reaction-Diffusion. It is used in the physical-error denominator but not in the shared reconstruction loss. Reaction-Diffusion uses the stricter complete-state comparison against the simulator state generated by linearly interpolated physical parameters. These percentages answer different scientific questions and are not directly comparable across metric definitions.

As an implementation sanity check, the updated Gaussian notebook evaluates the nearest-Gaussian metric on 20 normalized training fields and reports a mean and standard deviation of 0.0000% at its printed precision, effectively zero.

3.5 Why FID became secondary

Our Thin Plate experiment exposes the central numerical problem. Inception-v3 produces 2048-dimensional features from only 85 real training fields, so the covariance rank is at most 84. SciPy's matrix square root reported a singular-matrix warning. The direct 2048×2048 calculation took 12,503.42 ms; a low-rank factored calculation took 42.42 ms, a 295-fold speedup, with relative numerical difference 8.63×10^{-10}.

Table 4 FID computation diagnosis from our Thin Plate experiment.

Diagnostic	Observed value
Real feature bank	85×2048
Maximum covariance rank	84
Direct matrix square root	12,503.42 ms; singular warning
Factored calculation	42.42 ms
Speedup	295×
Relative numerical difference	8.63×10^{-10}
Final role	Secondary, not a primary ranking metric

The factorization fixes speed and numerical stability, but not the statistical limitation: an 85-sample covariance in 2048 dimensions is heavily rank deficient. Inception features are also trained on natural images, FID is distribution-level rather than path-wise, and FID values depend strongly on the selected feature extractor and reference population. Physics-aware generator errors are therefore the primary evaluation whenever the simulator is available.

4 Results

Table 5 Physical and reconstruction leaders. The winner changes with dataset and objective.

Dataset	Physical leader	Reconstruction leader	Metric scope
Gaussian	Contractive: 3.195%	Sparse: 3.11%	Nearest manifold; one trained run
Thin Plate	Denoising: 9.881%	Sparse: 4.969%	Nearest manifold; five seeds
Shallow Water	IRMAE: 30.247%	RRAE: 9.852%	Nearest manifold; five seeds
Reaction-Diffusion	L1: 3.873%	Sparse: 1.890%	Strict path; five seeds

4.1 Gaussian

Contractive is the best-performing method in the available Gaussian run, with the lowest nearest-Gaussian error at 3.195%. Vanilla follows at 4.112%, then L1 at 4.841% and RRAE at 5.292%. Sparse has the lowest final reconstruction error at 3.11%, followed by LoRAE at 5.20% and Contractive at 5.42%; Vanilla and L2 are near 17.2%, RRAE reaches 18.78%, and Denoising reaches 22.48%.

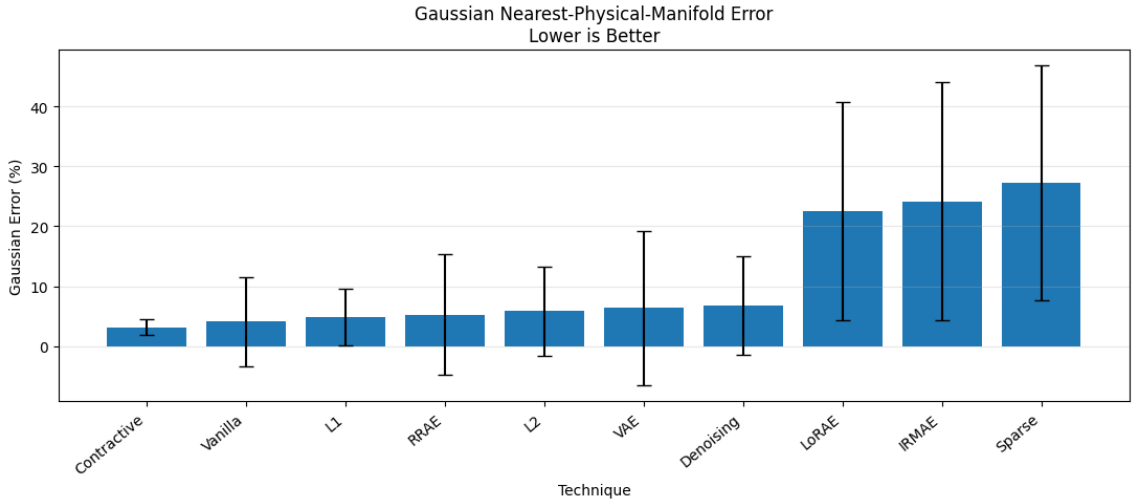


Figure 2 Gaussian nearest-manifold interpolation error. Lower is better. Error bars show standard deviation across 450 interpolation cases from one trained model, not training seeds.

The result is consistent with a locally smooth two-parameter translation manifold: the Contractive penalty directly suppresses encoder sensitivity, whereas an overcomplete code must discover the low-dimensional geometry. The updated Gaussian notebook instantiates RRAE with rank 2, equal to `SMALL_LATENT` and the two Gaussian center parameters; its physical error is 5.292%.

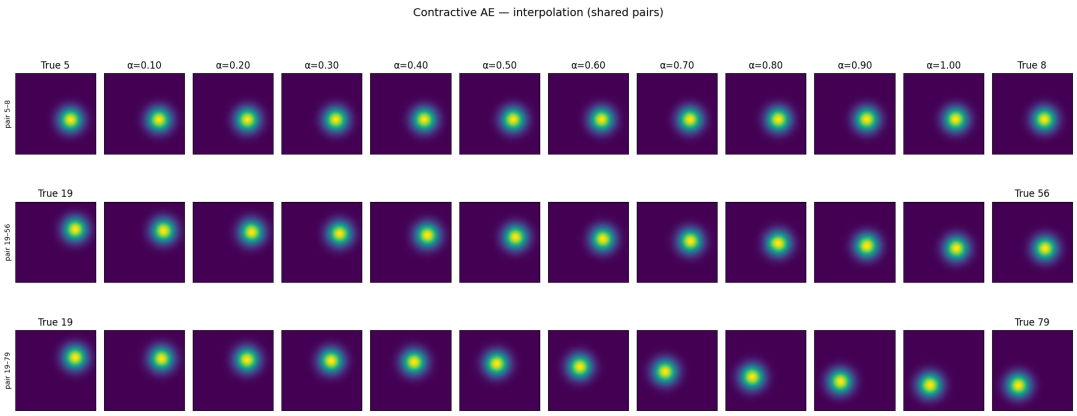


Figure 3 Representative Contractive Gaussian interpolation from the updated notebook. The numerical ranking uses all 450 evaluated cases.

4.2 Thin Plate

Denoising has the lowest five-seed nearest-manifold error at $9.881 \pm 1.780\%$, with Vanilla and Contractive nearly tied at 10.751% and 10.775%. Sparse again has the best reconstruction error ($4.969 \pm 0.551\%$) but a much weaker physical score ($18.259 \pm 0.701\%$). The normalized field has three effective parameters, while our tight models use latent dimension 4 and RRAE uses rank 16.

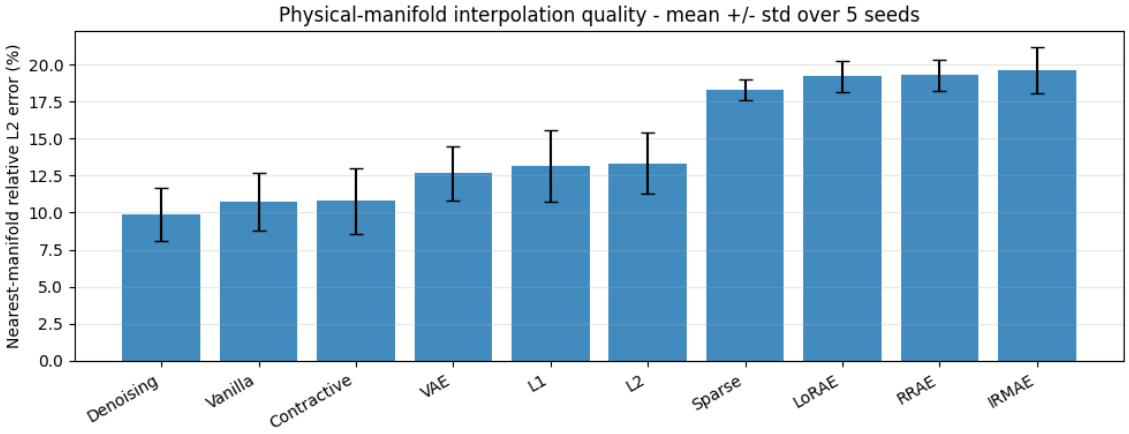


Figure 4 Thin Plate nearest-manifold interpolation error. Bars and error bars are the mean and sample standard deviation across five training seeds.

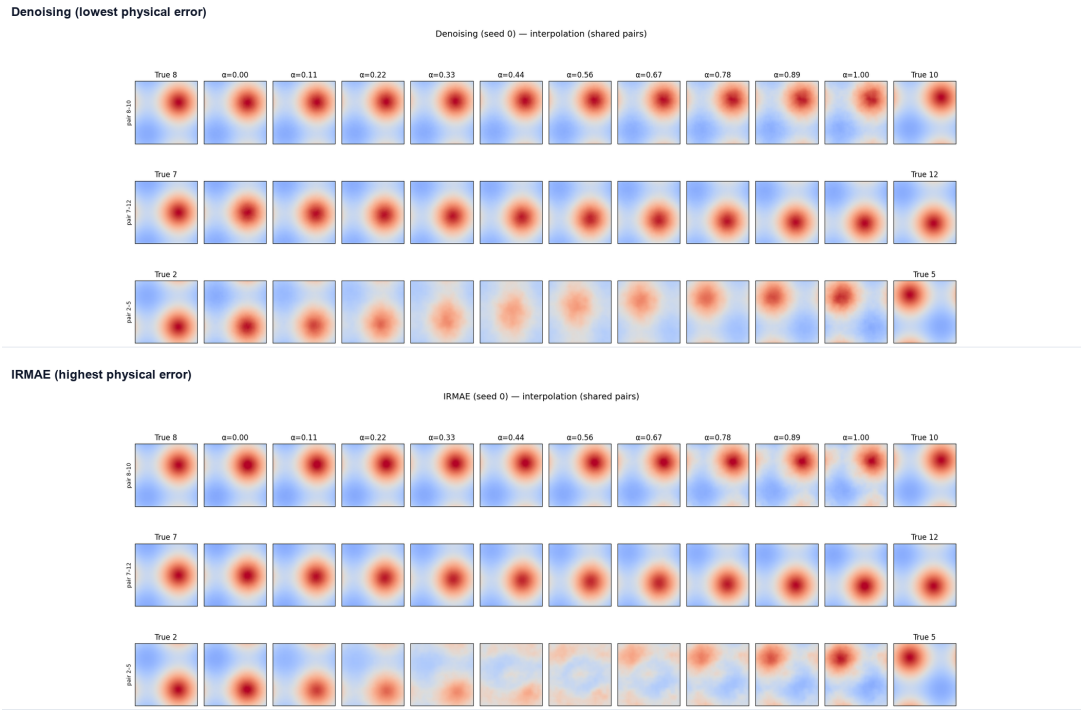


Figure 5 Representative Thin Plate interpolation comparison. The quantitative ranking uses all 450 cases per seed.

4.3 Shallow Water

The learned rank-reduction methods lead the nearest-manifold metric: IRMAE $30.247 \pm 0.719\%$, LoRAE $31.656 \pm 0.541\%$, and RRAE $32.703 \pm 0.646\%$. RRAE has the lowest reconstruction error at $9.852 \pm 0.430\%$, while IRMAE and LoRAE reconstruct less accurately but score better physically.

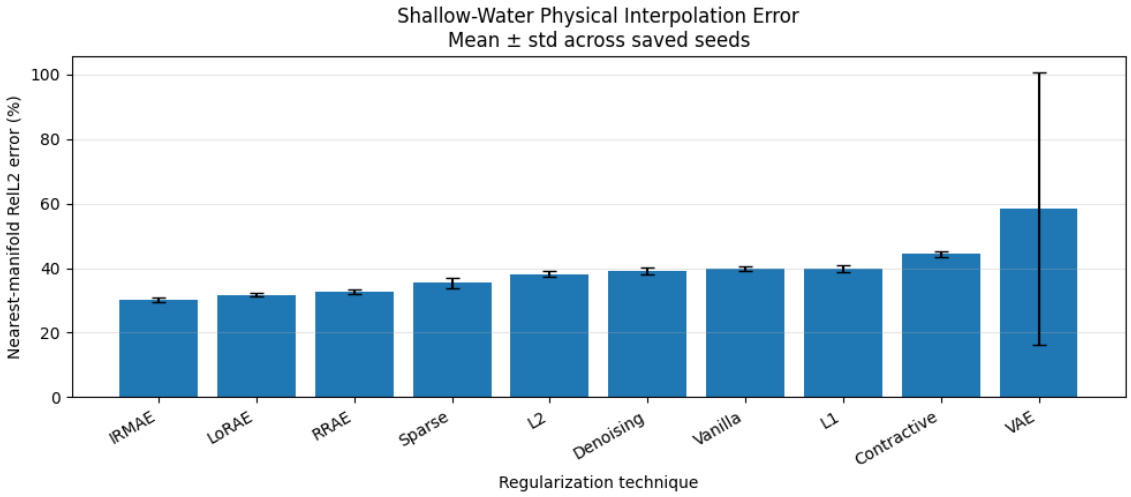


Figure 6 Shallow-Water nearest-manifold interpolation error across five training seeds. The VAE seed-0 outlier is retained.

VAE has mean physical error 58.458% with standard deviation 42.175%. Its seed-0 error is 133.886%, whereas seeds 1–4 are about 38–41%. This outlier is retained rather than removed. Our qualitative figure is a strict-path view for seed 2; the primary bar chart uses nearest-manifold

fitting across all five seeds.



Figure 7 Representative Shallow-Water strict-path comparison for seed 2. This visual is supplementary to the nearest-manifold ranking above.

4.4 Reaction-Diffusion

Under the strict physical-parameter-path complete-state metric, L1 leads at $3.873 \pm 0.188\%$, followed by L2 at $4.444 \pm 0.273\%$ and Contractive at $4.718 \pm 0.389\%$. Sparse is last physically at $7.479 \pm 0.132\%$ but has the lowest reconstruction error at $1.890 \pm 0.281\%$.

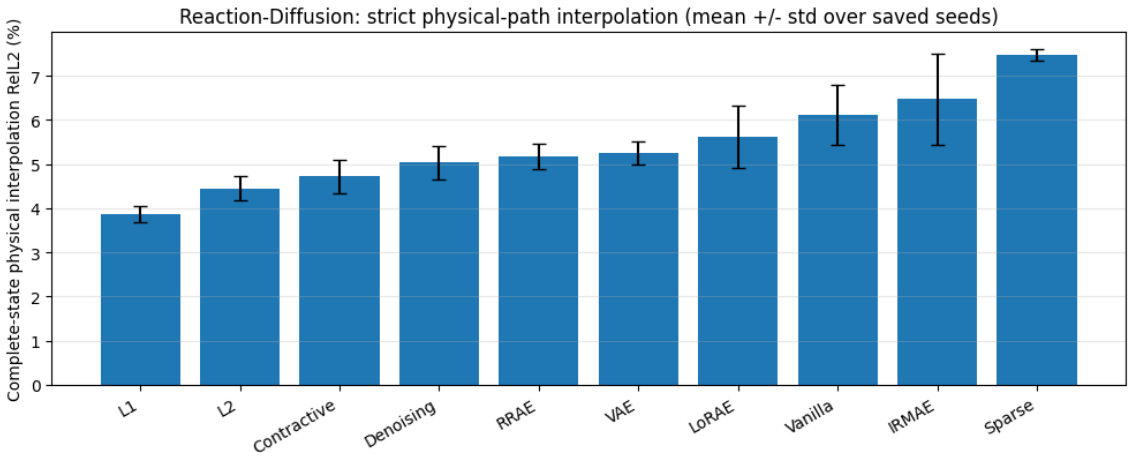


Figure 8 Reaction-Diffusion strict physical-path complete-state interpolation error across five training seeds.

Reaction-Diffusion has exactly three varying parameters and a fixed simulation horizon, while RRAE uses rank 3. Its ranking therefore reflects the interaction between regularization and this dataset rather than a shortage in the prescribed rank.

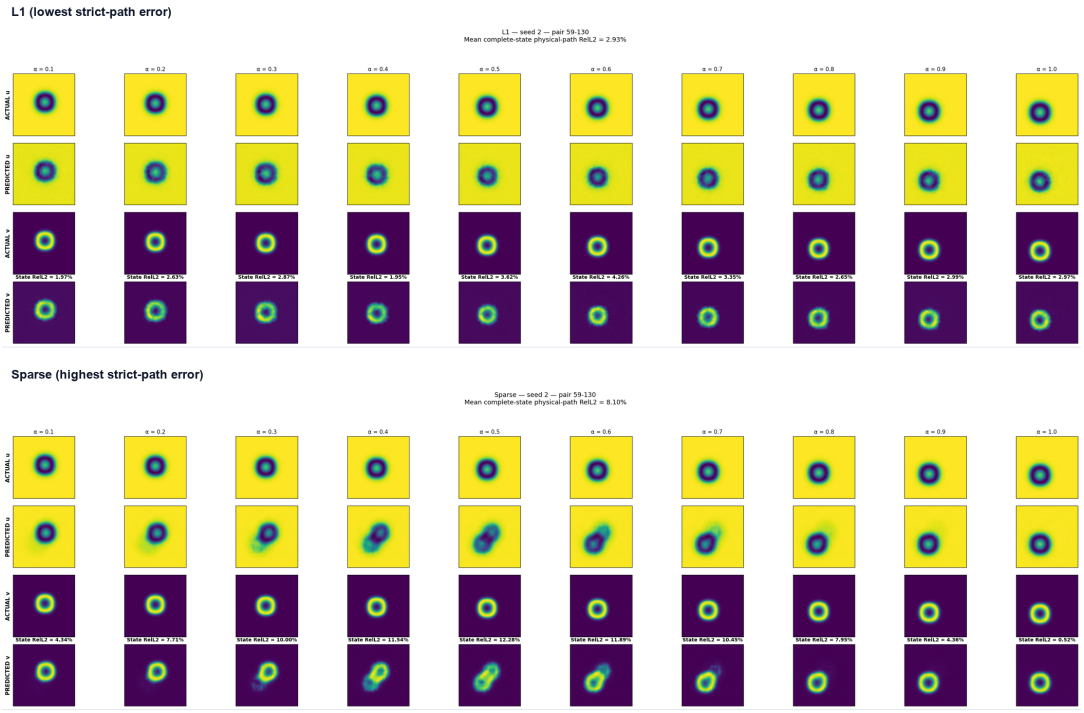


Figure 9 Representative complete-state Reaction-Diffusion interpolation comparison from our evaluation.

5 Discussion

5.1 No universal winner

The available Gaussian run and the three five-seed datasets produce four physical-interpolation leaders: Contractive, Denoising, IRMAE, and L1. Regularization interacts with manifold geometry, channel coupling, effective dimension, latent budget, and optimization stability. The evidence does not support a universal best method.

5.2 Reconstruction does not certify interpolation

Sparse has the lowest reconstruction error on Gaussian, Thin Plate, and Reaction-Diffusion, yet it is physically weak on all three and last under the Reaction-Diffusion strict path. Conversely, Denoising leads Thin Plate physically without leading reconstruction, and IRMAE leads Shallow Water physically despite a higher reconstruction error than RRAE. Endpoint accuracy can coexist with cross-fading or off-manifold intermediate states.

5.3 Local smoothness and learned rank

Contractive regularization is well aligned with the smooth translation manifold of the Gaussian data, where small physical movements produce small input changes. The coupled Shallow-Water system instead favors learned rank-reduction mechanisms in the nearest-manifold metric. These are plausible mechanism-level interpretations rather than universal causal proofs because penalty strengths and latent budgets differ by protocol.

5.4 Hard rank must be reported as implemented

We configured RRAE with dataset-specific rank budgets. The updated Gaussian notebook uses rank 2, Thin Plate uses rank 16, Shallow Water uses rank 9, and Reaction-Diffusion uses rank

3. Gaussian and Reaction-Diffusion match their two and three varying physical parameters, respectively; Thin Plate and Shallow Water do not. Controlled rank sweeps are required to isolate the effect of rank selection.

5.5 Physical metrics improve scientific relevance

Generator-based errors answer a question FID cannot: whether a decoded interpolation corresponds to a valid state under the known simulator. Nearest-manifold fitting tolerates alternative valid paths, while strict-path comparison tests a specific parameter trajectory. The metric definition must accompany every reported percentage.

5.6 Limitations

Four constraints qualify the rankings. First, Gaussian uses one trained run, so its dispersion across 450 interpolation cases describes within-run case variation rather than training-seed uncertainty. Second, nearest-manifold error for Gaussian, Thin Plate, and Shallow Water and strict-path error for Reaction-Diffusion answer different questions, so their absolute percentages are not directly comparable. Third, RRAE rank does not always match physical dimension: Gaussian uses rank 2 and Reaction-Diffusion rank 3, matching their varying-parameter counts, while Thin Plate uses rank 16 and Shallow Water rank 9. Finally, the methods retain unequal latent budgets and fixed method-specific penalty strengths, so the comparison isolates the evaluated protocols rather than regularizer identity alone. The retained Shallow-Water VAE outlier and the synthetic-data scope provide additional reasons to avoid universal causal claims.

Priority follow-up work is to conduct hyperparameter ablation studies on all four datasets, sweeping each technique's relevant penalty strengths and structural settings to identify the best-performing value or configuration for that technique under each dataset-specific physical metric; diagnose the Shallow-Water VAE outlier; and evaluate scarce measured engineering data with approximate physics-based validators.

6 Conclusion

This study addresses a practical industrial constraint: engineering companies often have only a small number of observations concentrated around localized operating conditions. In that regime, a model that reconstructs the available samples can still behave unphysically between them. Our controlled synthetic datasets reproduce this low-data, localized structure while providing known generators for quantitative physical validation.

Across the available Gaussian run and the three five-seed datasets, the physical leader changes from Contractive to Denoising to IRMAE to L1, and the reconstruction leader differs from the physical leader every time. Autoencoder regularization must therefore be evaluated beyond endpoint reconstruction.

Our practical recommendation is procedural rather than algorithmic: use a matched backbone; report actual latent and rank settings; evaluate multiple training seeds; and include a generator-based physical interpolation metric whenever a simulator is known. FID may remain a secondary distributional diagnostic, but high-dimensional small-sample covariance estimation and natural-image feature mismatch make it a weak primary criterion for localized engineering fields.

A Complete Numerical Results

The following tables reproduce all numerical results used in the report. Gaussian is a single trained run. Thin Plate, Shallow Water, and Reaction-Diffusion report mean and sample standard

deviation across five training seeds.

A.1 Gaussian

Table 6 Gaussian Physical RelL2 results. The Gaussian physical standard deviation is across 450 interpolation cases from one trained run.

Technique	Mean (%)	Std. (%)
Contractive	3.195	1.270
Vanilla	4.112	7.452
L1	4.841	4.686
RRAE	5.292	9.983
L2	5.851	7.469
VAE	6.392	12.882
Denoising	6.869	8.213
LoRAE	22.563	18.168
IRMAE	24.143	19.880
Sparse	27.251	19.521

Table 7 Gaussian Reconstruction RelL2 results. Only one Gaussian training run was available, so seed standard deviations are not reported.

Technique	Mean (%)	Std. (%)
Sparse	3.110	–
LoRAE	5.200	–
Contractive	5.420	–
IRMAE	7.200	–
L1	9.700	–
VAE	10.500	–
Vanilla	17.220	–
L2	17.280	–
RRAE	18.780	–
Denoising	22.480	–

A.2 Thin Plate

Table 8 Thin Plate Physical RelL2 results. Mean and sample standard deviation are across five training seeds.

Technique	Mean (%)	Std. (%)
Denoising	9.881	1.780
Vanilla	10.751	1.941
Contractive	10.775	2.255
VAE	12.662	1.821
L1	13.158	2.428
L2	13.335	2.060
Sparse	18.259	0.701
LoRAE	19.203	1.034
RRAE	19.277	1.045
IRMAE	19.604	1.574

Table 9 Thin Plate Reconstruction RelL2 results. Mean and sample standard deviation are across five training seeds.

Technique	Mean (%)	Std. (%)
Sparse	4.969	0.551
L1	5.689	1.099
L2	5.730	0.385
Vanilla	5.973	0.477
Contractive	6.040	0.485
RRAE	6.063	0.418
VAE	6.207	0.673
LoRAE	6.333	0.287
Denoising	6.527	0.511
IRMAE	8.143	1.102

A.3 Shallow Water

Table 10 Shallow Water Physical RelL2 results. Mean and sample standard deviation are across five training seeds.

Technique	Mean (%)	Std. (%)
IRMAE	30.247	0.719
LoRAE	31.656	0.541
RRAE	32.703	0.646
Sparse	35.428	1.671
L2	38.195	0.971
Denoising	39.185	0.990
Vanilla	39.745	0.772
L1	39.833	1.035
Contractive	44.379	0.832
VAE	58.458	42.175

Table 11 Shallow Water Reconstruction RelL2 results. Mean and sample standard deviation are across five training seeds.

Technique	Mean (%)	Std. (%)
RRAE	9.852	0.430
L1	11.794	0.391
L2	12.102	0.424
Vanilla	12.363	0.320
Denoising	12.842	0.676
Sparse	12.886	0.631
Contractive	13.705	0.401
IRMAE	14.465	1.312
LoRAE	15.854	0.967
VAE	30.051	35.258

A.4 Reaction-Diffusion

Table 12 Reaction-Diffusion Physical RelL2 results. Mean and sample standard deviation are across five training seeds.

Technique	Mean (%)	Std. (%)
L1	3.873	0.188
L2	4.444	0.273
Contractive	4.718	0.389
Denoising	5.040	0.379
RRAE	5.172	0.287
VAE	5.255	0.254
LoRAE	5.615	0.698
Vanilla	6.122	0.674
IRMAE	6.475	1.033
Sparse	7.479	0.132

Table 13 Reaction-Diffusion Reconstruction RelL2 results. Mean and sample standard deviation are across five training seeds.

Technique	Mean (%)	Std. (%)
Sparse	1.890	0.281
IRMAE	2.604	0.328
LoRAE	2.971	0.384
Contractive	3.679	0.902
VAE	3.702	0.946
L1	4.495	0.408
L2	4.629	0.088
RRAE	4.902	0.243
Denoising	5.125	0.196
Vanilla	5.199	0.166

References

[1] Robert Tibshirani. "Regression Shrinkage and Selection via the Lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1 (1996), pp. 267–288. DOI: 10.1111/j.2517-6161.1996.tb02080.x.

[2] Anders Krogh and John A. Hertz. "A Simple Weight Decay Can Improve Generalization". In: *Advances in Neural Information Processing Systems*. Vol. 4. 1991, pp. 950–957.

[3] Salah Rifai et al. "Contractive Auto-Encoders: Explicit Invariance During Feature Extraction". In: *Proceedings of the 28th International Conference on Machine Learning*. 2011, pp. 833–840.

[4] Pascal Vincent et al. "Extracting and Composing Robust Features with Denoising Autoencoders". In: *Proceedings of the 25th International Conference on Machine Learning*. 2008, pp. 1096–1103. DOI: 10.1145/1390156.1390294.

[5] Alireza Makhzani and Brendan J. Frey. "k-Sparse Autoencoders". In: *International Conference on Learning Representations*. 2014. arXiv: 1312.5663 [cs.LG].

[6] Diederik P. Kingma and Max Welling. "Auto-Encoding Variational Bayes". In: *International Conference on Learning Representations*. 2014. arXiv: 1312.6114 [stat.ML].

[7] Li Jing, Jure Zbontar, and Yann LeCun. "Implicit Rank-Minimizing Autoencoder". In: *Advances in Neural Information Processing Systems*. Vol. 33. 2020, pp. 14736–14746.

[8] Alokendu Mazumder et al. "Learning Low-Rank Latent Spaces with Simple Deterministic Autoencoder: Theoretical and Empirical Insights". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2024, pp. 2851–2860.

[9] Jad Mounayer et al. *Rank Reduction Autoencoders—Enhancing Interpolation on Nonlinear Manifolds*. 2024. arXiv: 2405.13980 [cs.LG].

[10] Shuangshuang Chen and Wei Guo. "Auto-Encoders in Deep Learning—A Review with New Perspectives". In: *Mathematics* 11.8, 1777 (2023). DOI: 10.3390/math11081777.

[11] Martin Heusel et al. "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium". In: *Advances in Neural Information Processing Systems*. Vol. 30. 2017.