Appendix A — Hybrid Spike Correction and Calibration Variables

This appendix documents the optional hybrid spike correction used for improved in-sample fidelity and provides a complete, unified description of all calibrated variables employed in the modeling pipeline. These components are reported here for transparency and reproducibility and are **not part of the classical DMD or DMDc theory**.

A.0 Optional Spike Hybrid Correction (In-Sample Fidelity Only)

The Optional Spike Hybrid Correction is a practical augmentation applied to the reconstructed state trajectories in order to better match sharp, transient events (spikes) observed in the measured TEiAS TPYS signals. This procedure is not part of standard DMD or Koopman theory but is introduced to improve in-sample visual fidelity and event alignment.

Let x_k^{DMD} denote the model reconstruction and x_k^{real} the measured signal. Spike locations are detected by thresholding the real signal using a high quantile q, i.e.,

$$\mathcal{S} = \left\{ k \;\middle|\; x_k^{\mathrm{real}} > Q_q(x^{\mathrm{real}}) \right\}, \tag{1}$$

where $Q_q(\cdot)$ denotes the q-quantile of the signal.

At detected spike locations, a blended reconstruction is used:

$$x_k^{\mathrm{hybrid}} = (1-\alpha)\, x_k^{\mathrm{DMD}} + \alpha\, x_k^{\mathrm{real}}, \quad k \in \mathcal{S}, \tag{2}$$

where

$$\alpha = \mathrm{SPIKE\_BLEND} \in [0,1] \tag{3}$$

is a user-defined blending coefficient.

For all non-spike points $k \notin \mathcal{S}$, the original model reconstruction is retained:

$$x_k^{\mathrm{hybrid}} = x_k^{\mathrm{DMD}}. \tag{4}$$

This procedure improves in-sample trajectory matching and visual spike alignment, but it does not represent predictive generalization. Therefore, all quantitative metrics (e.g., R^2, RMSE, MAE) reported in the main paper are computed without spike blending unless explicitly stated. article amsmath, amssymb enumitem geometry margin=1in

Appendix A — Calibration Variables (One unified appendix, paper-ready)

This appendix lists every calibrated variable used by your pipeline, what it means mathematically, why it is tuned, and what it indicates.

A.1 DMD / DMDC Rank Parameters (Model Complexity)

1. **DMD_RANK_MAX / DMD_RANK**
   Meaning: retained SVD rank for the regressor matrix (effective Koopman/DMD subspace dimension).
   Why tune: too small $\rightarrow$ underfit; too large $\rightarrow$ overfit/noise sensitivity.
   Indicates: number of dominant modes needed to represent system variability.

2. **DMDC_RANK_OMEGA** (r_Ω)
   Meaning: rank for the augmented regressor $\Omega = [X_t; U_t; 1]$.
   Why tune: must preserve input directions; low rank can discard control influence.
   Indicates: richness of combined state–input subspace.

3. **HANKEL_RANK_MAX (for DMD / DMDC)**
   Meaning: retained rank in Hankel-lifted space $\mathbb{R}^{sd}$.
   Why tune: Hankel dimension grows by sd, so regularization and rank control become more critical.
   Indicates: effective dimensionality of memory-augmented dynamics.

A.2 Tikhonov Regularization (Numerical Stability)

[label=0.]

1. **TIK / λ (DMD, DMDC)**
   Meaning: ridge parameter in $(\Sigma + \lambda I)^{-1}$.
   Why tune: stabilizes inversion of small singular values; prevents blow-ups.
   Indicates: noise level / ill-conditioning in regression.

2. **HANKEL_TIK**
   Meaning: ridge parameter in Hankel space.
   Why tune: Hankel embedding increases multicollinearity; higher λ often needed.
   Indicates: how aggressively the model must damp noise in memory-augmented regressors.

A.3 Stability Projection (Long-Horizon Rollout Safety)

1. **STABLE (boolean)**
   Meaning: enable spectral-radius projection.
   Why tune: ensures stable multi-step rollouts.
   Indicates: whether learned A tends to be unstable.

article amsmath, amssymb enumitem xcolor

[label=**0.**]

7. **RHO.MAX / $\rho_{\max}$**
   Meaning: spectral-radius cap: $\rho(\Lambda) \leq \rho_{\max}$.
   Why tune: too low $\rightarrow$ overly damped dynamics; too high $\rightarrow$ divergence.
   Indicates: balance between preserving dynamics and preventing explosion.

A.4 Hankel Memory Depth

1. **HANKEL.D / d**
   Meaning: delay window length; Hankel state dimension becomes nd.
   Why tune: larger d captures longer memory but increases dimension/overfit risk.
   Indicates: how non-Markovian the measured dynamics are.

A.5 Two-Stage Grid Search Controls

1. **MAX.TRIALS, PATIENCE**
   Meaning: search budget and early stopping.
   Why tune: trade compute vs quality; avoid wasting evaluations.
   Indicates: stability of objective improvements.

2. **SMART.GRID / Stage-2 refinement**
   Meaning: local refinement around best candidates (rank ± 2, tik scaling, etc.).
   Why tune: improves efficiency and likelihood of finding local optimum.
   Indicates: objective landscape smoothness around good solutions.

A.6 ML Hyperparameters (Baselines)

1. **LR:** no major hyperparameters in standard OLS (besides optional regularization if used).

2. **RF:** number of trees, depth, leaf size
   Meaning: bias–variance control for nonlinear mapping $x_{t+1} = g(x_t, u_t)$.
   Why tune: prevent overfit while capturing nonlinearities.
   Indicates: degree of nonlinearity/heterogeneity in the data.