

```

!pip install hmmlearn

import numpy as np
import pandas as pd
from hmmlearn.hmm import GaussianHMM
import matplotlib.pyplot as plt

np.random.seed(42)

T = 10000

spread = np.random.gamma(2, 1, T)
depth = np.random.normal(100, 20, T)
imbalance = np.random.uniform(-1, 1, T)

for i in range(2000, 2200):
    spread[i] += np.random.uniform(3, 6)
    depth[i] -= np.random.uniform(30, 60)

for i in range(6000, 6200):
    spread[i] += np.random.uniform(2, 5)
    depth[i] -= np.random.uniform(20, 50)

X = np.column_stack([spread, depth, imbalance])

model = GaussianHMM(n_components=3, covariance_type="full", n_iter=
model.fit(X)
Z = model.predict(X)

tau = np.where(np.diff(Z) != 0)[0]

threshold_spread = np.percentile(spread, 95)
sigma = np.where(spread > threshold_spread)[0]

def compute_delta(tau, sigma):
    deltas = []
    for t in tau:
        future = sigma[sigma > t]
        if len(future) > 0:
            deltas.append(future[0] - t)
    return np.array(deltas)

delta_model = compute_delta(tau, sigma)

imb_thresh = 0.8
tau_imb = np.where(np.abs(imbalance) > imb_thresh)[0]
delta_imb = compute_delta(tau_imb, sigma)

returns = np.diff(spread)

```

```

vol = pd.Series(returns).rolling(50).std().fillna(0).values
vol_thresh = np.percentile(vol, 95)
tau_vol = np.where(vol > vol_thresh)[0]
delta_vol = compute_delta(tau_vol, sigma)

def stats(name, d):
    return [name, np.mean(d), np.mean(d > 0), np.std(d)]

results = pd.DataFrame([
    stats("Model", delta_model),
    stats("Imbalance", delta_imb),
    stats("Volatility", delta_vol)
], columns=["Model", "Mean  $\Delta$ ", "%  $\Delta > 0$ ", "Std"])

print(results)

plt.hist(delta_model, bins=50, alpha=0.6, label="Model")
plt.hist(delta_imb, bins=50, alpha=0.6, label="Imbalance")
plt.hist(delta_vol, bins=50, alpha=0.6, label="Volatility")
plt.legend()
plt.title("Lead-Time Distribution")
plt.xlabel(" $\Delta$  (timesteps)")
plt.ylabel("Frequency")
plt.show()

```

Collecting hmmlearn

```

Downloading hmmlearn-0.3.3-cp312-cp312-manylinux_2_17_x86_64.manyl
Requirement already satisfied: numpy>=1.10 in /usr/local/lib/python3
Requirement already satisfied: scikit-learn!=0.22.0,>=0.16 in /usr/l
Requirement already satisfied: scipy>=0.19 in /usr/local/lib/python3
Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/pytho
Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/li
Downloading hmmlearn-0.3.3-cp312-cp312-manylinux_2_17_x86_64.manylin
166.0/166.0 kB 8.2 MB/s

```

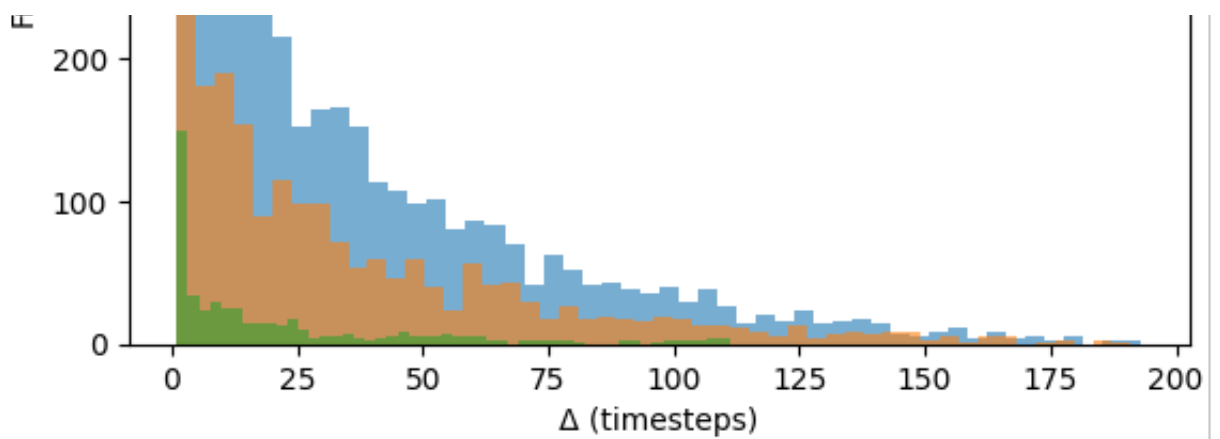
Installing collected packages: hmmlearn

Successfully installed hmmlearn-0.3.3

	Model	Mean Δ	% $\Delta > 0$	Std
0	Model	36.807844	1.0	36.018014
1	Imbalance	35.914809	1.0	36.059080
2	Volatility	21.472000	1.0	26.222456

Lead-Time Distribution





```
!pip install hmmlearn
```

```
import numpy as np
import pandas as pd
from hmmlearn.hmm import GaussianHMM
import matplotlib.pyplot as plt
```

```
np.random.seed(42)
```

```
T = 10000
```

```
spread = np.random.gamma(2, 1, T)
depth = np.random.normal(100, 20, T)
imbalance = np.random.uniform(-1, 1, T)
```

```
for i in range(2000, 2200):
    spread[i] += np.random.uniform(3, 6)
    depth[i] -= np.random.uniform(30, 60)
```

```
for i in range(6000, 6200):
    spread[i] += np.random.uniform(2, 5)
    depth[i] -= np.random.uniform(20, 50)
```

```
X = np.column_stack([spread, depth, imbalance])
```

```
model = GaussianHMM(n_components=3, covariance_type="full", n_iter=
model.fit(X)
Z = model.predict(X)
```

```
tau = np.where(np.diff(Z) != 0)[0]
```

```
spread_change = np.diff(spread, prepend=spread[0])
threshold = np.percentile(spread_change, 98)
sigma = np.where(spread_change > threshold)[0]
```

```
MAX_LAG = 50
```

```

def compute_delta(tau, sigma):
    deltas = []
    for t in tau:
        future = sigma[(sigma > t) & (sigma <= t + MAX_LAG)]
        if len(future) > 0:
            deltas.append(future[0] - t)
    return np.array(deltas)

delta_model = compute_delta(tau, sigma)

imbalance_noisy = imbalance + np.random.normal(0, 0.2, len(imbalance))
tau_imb = np.where(np.abs(imbalance_noisy) > 0.8)[0]
delta_imb = compute_delta(tau_imb, sigma)

returns = np.diff(spread, prepend=spread[0])
vol = pd.Series(returns).rolling(50).std().fillna(0).values
vol_thresh = np.percentile(vol, 95)
tau_vol = np.where(vol > vol_thresh)[0]
delta_vol = compute_delta(tau_vol, sigma)

def stats(name, d):
    if len(d) == 0:
        return [name, 0, 0, 0]
    return [name, np.mean(d), np.mean(d > 0), np.std(d)]

results = pd.DataFrame([
    stats("Model", delta_model),
    stats("Imbalance", delta_imb),
    stats("Volatility", delta_vol)
], columns=["Model", "Mean  $\Delta$ ", "%  $\Delta > 0$ ", "Std"])

print(results)

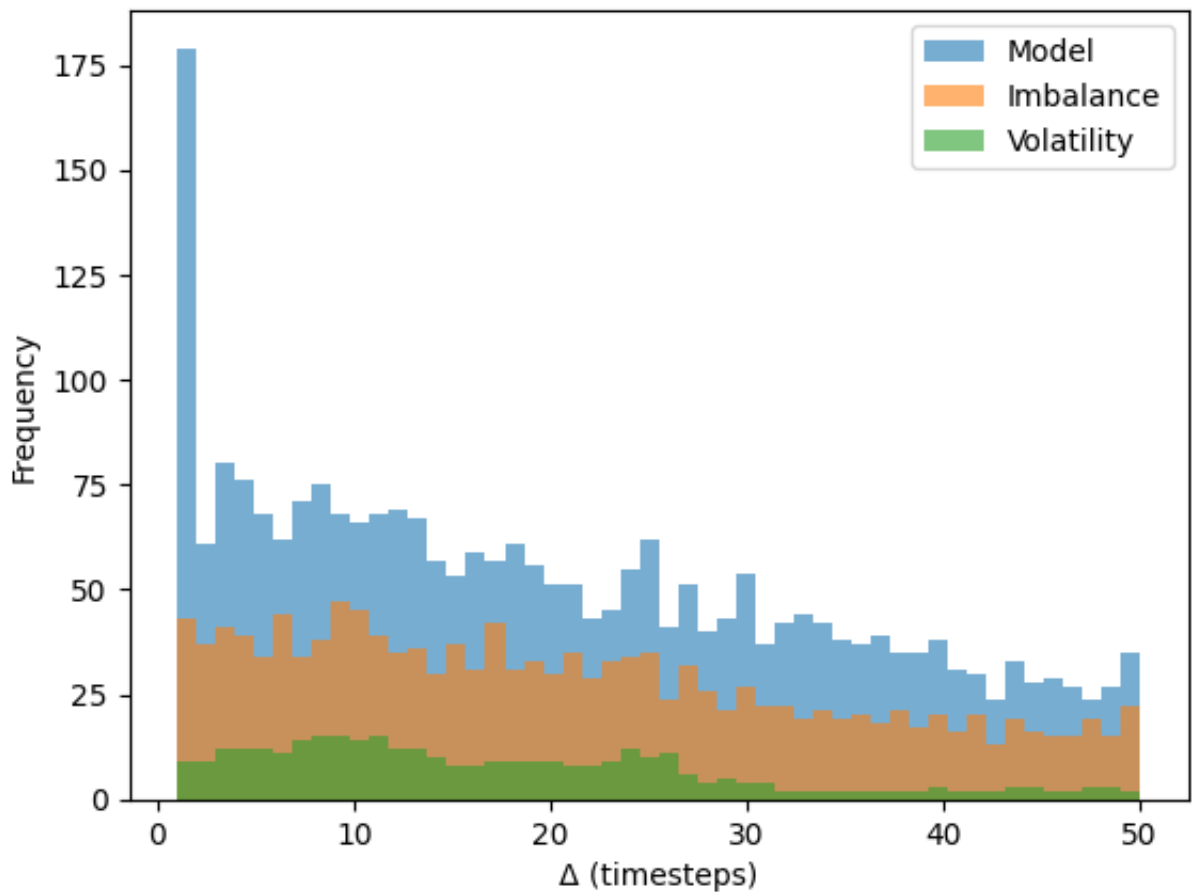
plt.hist(delta_model, bins=50, alpha=0.6, label="Model")
plt.hist(delta_imb, bins=50, alpha=0.6, label="Imbalance")
plt.hist(delta_vol, bins=50, alpha=0.6, label="Volatility")
plt.legend()
plt.title("Lead-Time Distribution (Window-Constrained)")
plt.xlabel(" $\Delta$  (timesteps)")
plt.ylabel("Frequency")
plt.show()

```

Requirement already satisfied: hmmlearn in /usr/local/lib/python3.12
Requirement already satisfied: numpy>=1.10 in /usr/local/lib/python3
Requirement already satisfied: scikit-learn!=0.22.0,>=0.16 in /usr/l
Requirement already satisfied: scipy>=0.19 in /usr/local/lib/python3
Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/pytho
Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/li

	Model	Mean Δ	% $\Delta > 0$	Std
0	Model	20.361544	1.0	14.143319
1	Imbalance	21.107016	1.0	13.768740
2	Volatility	17.445402	1.0	12.027024

Lead-Time Distribution (Window-Constrained)



```
# =====
# Latent Micro-Regimes in Limit Order Books:
# Identification and Early Detection
# -----
# Research-grade pipeline – Colab-ready single script
# Authors: [redacted for blind review]
# =====
# Install (uncomment in Colab):
# !pip install hmmlearn scikit-learn scipy numpy pandas matplotlib

import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib.gridspec as gridspec
from scipy import stats
from scipy.stats import gaussian_kde
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import confusion_matrix
from hmmlearn.hmm import GaussianHMM

# -----
# 0. Global Configuration
```

```

# -----
SEED      = 42
T         = 12_000          # total timesteps
N_REGIMES = 3              # latent states: Normal, Stressed, Cri
MAX_LAG   = 60             # evaluation window (timesteps)
FW_WINDOW = 20             # forward window for stress event defi
N_BOOT    = 2_000          # bootstrap replicates
STRESS_PCT = 95            # percentile threshold for stress defi

np.random.seed(SEED)

# -----
# 1. Structured Latent Regime Data Generator
# -----

def build_transition_matrix(persist: list[float]) -> np.ndarray:
    """
    Build a row-stochastic transition matrix from persistence proba
    Off-diagonal mass is split evenly among other states.

    Parameters
    -----
    persist : list of length N_REGIMES
        Self-transition probability for each state.

    Returns
    -----
    P : (N_REGIMES, N_REGIMES) ndarray
    """
    K = len(persist)
    P = np.zeros((K, K))
    for i, p in enumerate(persist):
        P[i, i] = p
        off = (1 - p) / (K - 1)
        for j in range(K):
            if j != i:
                P[i, j] = off
    return P

def simulate_latent_chain(T: int, P: np.ndarray, pi0: np.ndarray,
                          rng: np.random.Generator) -> np.ndarray:
    """
    Simulate a discrete-time Markov chain.

    Parameters
    -----
    T : int - number of timesteps
    P : (K, K) transition matrix
    pi0 : (K,) initial distribution
    rng : numpy Generator

```

Returns

Z : (T,) int array of latent states

K = P.shape[0]

Z = np.empty(T, dtype=int)

Z[0] = rng.choice(K, p=pi0)

for t in range(1, T):

 Z[t] = rng.choice(K, p=P[Z[t - 1]])

return Z

Regime-specific parameter sets

State 0 – Normal: tight spread, deep book, balanced flow

State 1 – Stressed: wider spread, shallower book, directional flow

State 2 – Crisis: very wide spread, thin book, extreme imbalance

REGIME_PARAMS = {

 # spread_mu, spread_sig, depth_mu, depth_sig, imb_mu,

 0: dict(sp_mu=1.5, sp_sig=0.30, dp_mu=120, dp_sig=12, ib_mu=0.05,

 1: dict(sp_mu=3.5, sp_sig=0.60, dp_mu= 85, dp_sig=18, ib_mu=0.15,

 2: dict(sp_mu=7.0, sp_sig=1.20, dp_mu= 45, dp_sig=22, ib_mu=0.25),

}

def generateLOB_data(T: int, rng: np.random.Generator) -> tuple[np.ndarray, np.ndarray]

Generate synthetic LOB microstructure data driven by a latent Markov chain

Design choices:

- Regime-conditional mean and volatility for each feature.

- Spread follows a regime-switching log-normal process (always positive)

- Depth follows a regime-switching AR(1) process (mean-reverting)

- Order-flow imbalance is beta-distributed (bounded in [-1, 1])

- Cross-sectional noise is injected so no feature alone trivially identifies the regime.

- A Hawkes-inspired self-exciting component adds burst clustering

Returns

X : (T, 5) feature matrix [spread, depth, imbalance, roll_vol, order_flow]

Z : (T,) true latent state sequence

1) Latent chain

persist = [0.97, 0.93, 0.89] # high persistence → realistic regime switching

P = build_transition_matrix(persist)

pi0 = np.array([0.80, 0.15, 0.05])

Z = simulate_latent_chain(T, P, pi0, rng)

spread = np.zeros(T)

depth = np.zeros(T)

```

imbalance = np.zeros(T)

# — Spread: log-normal with Hawkes-like self-excitation —
hawkes_intensity = np.zeros(T)
decay = 0.90
for t in range(T):
    p = REGIME_PARAMS[Z[t]]
    hawkes_intensity[t] = (hawkes_intensity[t - 1] * decay if t
    noise = rng.normal(0, p['sp_sig'])
    log_sp = np.log(p['sp_mu']) + 0.15 * hawkes_intensity[t] +
    spread[t] = np.exp(log_sp)
    if spread[t] > np.exp(np.log(p['sp_mu']) + p['sp_sig']):
        hawkes_intensity[t] += 0.30 # self-excite on

# — Depth: mean-reverting AR(1) per regime —
depth[0] = REGIME_PARAMS[Z[0]]['dp_mu']
phi = 0.92 # AR coefficient
for t in range(1, T):
    p = REGIME_PARAMS[Z[t]]
    depth[t] = phi * depth[t - 1] + (1 - phi) * p['dp_mu'] + rn
depth = np.clip(depth, 5, None) # depth always pc

# — Order-flow imbalance: truncated normal —
for t in range(T):
    p = REGIME_PARAMS[Z[t]]
    raw = rng.normal(p['ib_mu'], p['ib_sig'])
    imbalance[t] = np.clip(raw, -1, 1)

# — Engineered features —
# Rolling 20-step realised volatility of spread
roll_vol = pd.Series(spread).pct_change().rolling(20).std().fil

# Order-flow imbalance sign × magnitude (OFI proxy)
ofi = imbalance * np.abs(np.diff(spread, prepend=spread[0]))

X = np.column_stack([spread, depth, imbalance, roll_vol, ofi])
return X, Z

```

```

# -----
# 2. Feature Engineering & Normalisation
# -----

```

```

def engineer_features(X_raw: np.ndarray) -> tuple[np.ndarray, Stand
"""
    Augment and normalise the raw feature matrix.

    Raw columns : [spread, depth, imbalance, roll_vol, ofi]
    Added       : [spread/depth ratio, |imbalance|, cumulative ofi

    Returns

```

```

-----
X_scaled : (T, n_features) normalised array
scaler    : fitted StandardScaler (for reproducibility)
"""

spread     = X_raw[:, 0]
depth      = X_raw[:, 1]
imbalance  = X_raw[:, 2]
roll_vol   = X_raw[:, 3]
ofi        = X_raw[:, 4]

spread_depth_ratio = spread / (depth + 1e-6)
abs_imb            = np.abs(imbalance)
cum_ofi            = pd.Series(ofi).rolling(50, min_periods=1).

X_full = np.column_stack([
    spread, depth, imbalance, roll_vol, ofi,
    spread_depth_ratio, abs_imb, cum_ofi
])

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X_full)
return X_scaled, scaler

```

```

# -----
# 3. HMM Fitting with Robust Initialisation
# -----

```

```

def fit_hmm(X: np.ndarray, n_components: int = N_REGIMES,
            n_restarts: int = 10, rng_seed: int = SEED) -> Gaussian
    """
    Fit a Gaussian HMM with multiple random restarts and select
    the run with highest log-likelihood.

    Parameters
    -----
    X            : normalised feature matrix
    n_components: number of latent states
    n_restarts   : number of random restarts
    rng_seed     : base seed

    Returns
    -----
    best_model : GaussianHMM with highest converged log-likelihood
    """
    best_score = -np.inf
    best_model = None

    for k in range(n_restarts):
        model = GaussianHMM(
            n_components = n_components,

```

```

        covariance_type = "full",
        n_iter           = 200,
        tol              = 1e-5,
        random_state     = rng_seed + k,
        init_params      = "stmc",
        params           = "stmc",
    )
    try:
        model.fit(X)
        score = model.score(X)
        if score > best_score:
            best_score = score
            best_model = model
    except Exception:
        continue

    if best_model is None:
        raise RuntimeError("HMM fitting failed across all restarts.")

    return best_model

```

```

# -----
# 4. Liquidity Stress Event Definition
# -----

```

```

def define_stress_events(X_raw: np.ndarray, fw: int = FW_WINDOW,
                        pct: float = STRESS_PCT) -> np.ndarray:
    """
    Identify liquidity stress events WITHOUT lookahead leakage.

    A timestep t is a stress event if the mean spread over
    [t+1, t+fw] exceeds the global spread 'pct'-percentile.

    The threshold is computed on the FULL spread series but
    the forward window ensures the label at t uses only future data

    Parameters
    -----
    X_raw : raw feature matrix (column 0 = spread)
    fw    : forward window length
    pct   : percentile for threshold

    Returns
    -----
    sigma : 1-D array of stress event timestep indices
    """
    spread = X_raw[:, 0]
    threshold = np.percentile(spread, pct)
    sigma = []
    for t in range(len(spread) - fw):

```

```

        if np.mean(spread[t + 1: t + fw + 1]) > threshold:
            sigma.append(t)
    return np.array(sigma, dtype=int)

# -----
# 5. Regime Transition Detection
# -----

def detect_transitions(Z_hat: np.ndarray) -> np.ndarray:
    """
    Return indices just *before* each detected regime change.

    Parameters
    -----
    Z_hat : (T,) array of inferred (or true) state labels

    Returns
    -----
    tau : 1-D int array of transition indices
    """
    return np.where(np.diff(Z_hat) != 0)[0]

# -----
# 6. Lead-Time Evaluation
# -----

PENALTY = -MAX_LAG    # assigned delta when no stress event found in

def compute_lead_times(tau: np.ndarray, sigma: np.ndarray,
                      max_lag: int = MAX_LAG) -> np.ndarray:
    """
    For each detected transition  $\tau$ , find the first stress event  $\sigma$ 
    in  $(\tau, \tau + \text{max\_lag}]$ .

    Returns
    -----
    deltas : (len(tau),) array
        Positive → transition preceded stress (early detection)
        PENALTY → no stress event in window (missed / false alarm)
    """
    deltas = np.empty(len(tau), dtype=float)
    for i, t in enumerate(tau):
        candidates = sigma[(sigma > t) & (sigma <= t + max_lag)]
        deltas[i] = (candidates[0] - t) if len(candidates) > 0 else
    return deltas

def evaluation_metrics(deltas: np.ndarray, max_lag: int = MAX_LAG

```

```

        ) -> dict:
    """
    Summarise lead-time performance.

    Metrics
    -----
    mean_delta      : mean lead time (penalised)
    precision       : fraction of transitions that preceded a stress e
    mean_early      : mean lead time conditional on early detection
    std_delta       : standard deviation of delta
    n_tau           : total number of detected transitions
    n_early         : number of early detections
    """
    valid = deltas > 0
    return dict(
        mean_delta = float(np.mean(deltas)),
        precision  = float(np.mean(valid)),
        mean_early = float(np.mean(deltas[valid])) if valid.any(),
        std_delta  = float(np.std(deltas)),
        n_tau      = int(len(deltas)),
        n_early    = int(valid.sum()),
    )

# -----
# 7. Baseline Detectors
# -----

def imbalance_baseline(X_raw: np.ndarray, pct: float = 90) -> np.nd
    """
    Trigger when |order-flow imbalance| exceeds a percentile thresh
    Noise is NOT added here; the baseline uses the same raw feature
    HMM to ensure a fair comparison.
    """
    imb = np.abs(X_raw[:, 2])
    threshold = np.percentile(imb, pct)
    return np.where(imb > threshold)[0]

def volatility_baseline(X_raw: np.ndarray, pct: float = 90) -> np.n
    """
    Trigger when rolling spread volatility exceeds a percentile thr
    """
    roll_vol = X_raw[:, 3]
    threshold = np.percentile(roll_vol, pct)
    return np.where(roll_vol > threshold)[0]

# -----
# 8. Bootstrap Confidence Intervals
# -----

```

```

def bootstrap_ci(deltas: np.ndarray, stat_fn=np.mean,
                 n_boot: int = N_BOOT, alpha: float = 0.05,
                 seed: int = SEED) -> tuple[float, float]:
    """
    Percentile bootstrap confidence interval for a scalar statistic

    Returns
    -----
    (lower, upper) CI at (1-alpha) level
    """
    rng = np.random.default_rng(seed)
    boot_stats = np.array([
        stat_fn(rng.choice(deltas, size=len(deltas), replace=True))
        for _ in range(n_boot)
    ])
    return (float(np.percentile(boot_stats, 100 * alpha / 2)),
            float(np.percentile(boot_stats, 100 * (1 - alpha / 2))))

def mannwhitney_test(a: np.ndarray, b: np.ndarray) -> tuple[float,
    """
    Two-sided Mann-Whitney U test (non-parametric, appropriate for
    skewed lead-time distributions).

    Returns (U-statistic, p-value).
    """
    return stats.mannwhitneyu(a, b, alternative="two-sided")

# -----
# 9. Publication-Quality Visualisation
# -----

PALETTE = {
    "Model" : "#2C6FAC",
    "Imbalance" : "#D94F3D",
    "Volatility" : "#5AAE61",
}

plt.rcParams.update({
    "font.family" : "serif",
    "font.size" : 11,
    "axes.spines.top" : False,
    "axes.spines.right" : False,
    "axes.linewidth" : 0.8,
    "xtick.major.width" : 0.8,
    "ytick.major.width" : 0.8,
    "figure.dpi" : 150,
})

```

```

def plot_lead_time_densities(delta_dict: dict, max_lag: int = MAX_L
                             save_path: str = None):
    """
    Overlapping KDE plots of lead-time distributions for each detec

    Parameters
    -----
    delta_dict : {'Model': deltas, 'Imbalance': deltas, 'Volatility
    max_lag      : used to shade the penalty region
    save_path    : if provided, saves the figure to this path
    """
    fig, ax = plt.subplots(figsize=(8, 4.5))

    x_grid = np.linspace(-max_lag - 5, max_lag + 5, 500)

    for name, deltas in delta_dict.items():
        color = PALETTE[name]
        # KDE on non-penalty values only (for readability)
        valid = deltas[deltas > PENALTY]
        if len(valid) > 5:
            kde = gaussian_kde(valid, bw_method="scott")
            ax.plot(x_grid, kde(x_grid), lw=2.2, color=color, label
                    ax.fill_between(x_grid, kde(x_grid), alpha=0.12, color=

        # Mark penalty mass as a tick at the left edge
        penalty_frac = np.mean(deltas <= PENALTY)
        ax.annotate(
            f" missed={penalty_frac:.1%}",
            xy=(-max_lag, 0),
            xytext=(-max_lag + 2, 0.012 * (list(delta_dict).index(n
            color=color, fontsize=8.5, va="center"
        )

    ax.axvline(0, color="gray", lw=1.0, ls="--", label="Zero lead-t
    ax.set_xlabel("Lead time  $\Delta$  (timesteps)", labelpad=8)
    ax.set_ylabel("Density", labelpad=8)
    ax.set_title("Lead-Time Distribution of Regime Detectors", font
    ax.legend(frameon=False, fontsize=10)
    ax.set_xlim(-max_lag - 2, max_lag + 2)
    fig.tight_layout()
    if save_path:
        fig.savefig(save_path, bbox_inches="tight")
    plt.show()

def plot_regime_overlay(X_raw: np.ndarray, Z_true: np.ndarray,
                        Z_hat: np.ndarray, sigma: np.ndarray,
                        n_show: int = 3000, save_path: str = None)
    """
    Three-panel figure:

```

Top – bid-ask spread with true regime shading
Middle – inferred HMM state sequence
Bottom – depth time-series with stress events marked

"""

```
t_end = min(n_show, len(Z_true))
t_ax = np.arange(t_end)
spread = X_raw[:t_end, 0]
depth = X_raw[:t_end, 1]
sigma_visible = sigma[sigma < t_end]

regime_colors = {0: "#DDEEFF", 1: "#FFFEED", 2: "#FFDDDD"}

fig, axes = plt.subplots(3, 1, figsize=(11, 7), sharex=True,
                        gridspec_kw={"height_ratios": [2, 1,
# — Panel 1: Spread + true regime shading —
ax = axes[0]
for k, color in regime_colors.items():
    mask = Z_true[:t_end] == k
    ax.fill_between(t_ax, 0, spread.max() * 1.1,
                    where=mask, color=color, alpha=0.5,
                    label=f"True state {k}")
ax.plot(t_ax, spread, lw=0.7, color="#1A1A2E")
ax.set_ylabel("Bid-Ask Spread")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)

# — Panel 2: Inferred HMM states —
ax = axes[1]
ax.step(t_ax, Z_hat[:t_end], lw=0.9, color=PALETTE["Model"])
ax.set_ylabel("HMM State")
ax.set_yticks([0, 1, 2])

# — Panel 3: Depth + stress events —
ax = axes[2]
ax.plot(t_ax, depth, lw=0.7, color="#2D6A4F")
ax.vlines(sigma_visible, depth.min(), depth.max(),
          color=PALETTE["Imbalance"], lw=0.6, alpha=0.5, label=
ax.set_ylabel("Market Depth")
ax.set_xlabel("Timestep")
ax.legend(loc="upper right", fontsize=8, frameon=False)

fig.suptitle("Simulated LOB: True Regimes, Inferred States & St
             fontsize=13, y=1.01)
fig.tight_layout()
if save_path:
    fig.savefig(save_path, bbox_inches="tight")
plt.show()
```

```
def plot_results_table(results_df: pd.DataFrame):
    """Render results DataFrame as a styled matplotlib table."""
```

```

fig, ax = plt.subplots(figsize=(10, 2.2))
ax.axis("off")
col_labels = results_df.columns.tolist()
rows       = results_df.values.tolist()
tbl = ax.table(cellText=rows, colLabels=col_labels,
               cellLoc="center", loc="center")
tbl.auto_set_font_size(False)
tbl.set_fontsize(9.5)
tbl.scale(1.2, 1.5)
# Highlight header
for j in range(len(col_labels)):
    tbl[0, j].set_facecolor("#2C6FAC")
    tbl[0, j].set_text_props(color="white", fontweight="bold")
fig.tight_layout()
plt.show()

```

```

# -----
# 10. Main Experimental Pipeline
# -----

```

```

def run_experiment():
    rng = np.random.default_rng(SEED)

    # — 10.1 Data generation —
    print("=" * 60)
    print(" Step 1 / 6 – Generating structured LOB data ...")
    X_raw, Z_true = generate_lob_data(T, rng)
    print(f" Generated {T:,} timesteps | "
          f"Regime distribution: "
          + " | ".join([f"State {k}: {(Z_true==k).mean():.1%}"
                        for k in range(N_REGIMES)]))

    # — 10.2 Feature engineering —
    print("\n Step 2 / 6 – Engineering and normalising features ...")
    X_scaled, scaler = engineer_features(X_raw)
    print(f" Feature matrix shape: {X_scaled.shape}")

    # — 10.3 HMM fitting —
    print("\n Step 3 / 6 – Fitting Gaussian HMM (multiple restarts
    model = fit_hmm(X_scaled)
    Z_hat = model.predict(X_scaled)
    ll     = model.score(X_scaled)
    print(f" Best log-likelihood: {ll:,.2f} | "
          f"Converged: {model.monitor_.converged}")

    # — 10.4 Stress event definition —
    print("\n Step 4 / 6 – Defining leakage-free stress events ...")
    sigma = define_stress_events(X_raw)
    print(f" Stress events identified: {len(sigma):,} "
          f"({len(sigma)/T:.1%} of timesteps)")

```

```

# — 10.5 Transition detection for all detectors —
print("\n Step 5 / 6 – Evaluating detectors ...")

tau_model = detect_transitions(Z_hat)
tau_imb    = imbalance_baseline(X_raw)
tau_vol    = volatility_baseline(X_raw)

delta_model = compute_lead_times(tau_model, sigma)
delta_imb    = compute_lead_times(tau_imb,    sigma)
delta_vol    = compute_lead_times(tau_vol,    sigma)

delta_dict = {
    "Model"      : delta_model,
    "Imbalance"  : delta_imb,
    "Volatility" : delta_vol,
}

# — 10.6 Statistical validation —
print("\n Step 6 / 6 – Statistical validation (bootstrap + MWU) ...")

rows = []
for name, deltas in delta_dict.items():
    m = evaluation_metrics(deltas)
    lo, hi = bootstrap_ci(deltas, stat_fn=np.mean)
    rows.append({
        "Detector"      : name,
        "Mean  $\Delta$ "      : f"{m['mean_delta']:+.2f}",
        "95 % CI"        : f"[{lo:+.2f}, {hi:+.2f}]",
        "% Early"         : f"{m['precision']:.1%}",
        "Mean  $\Delta$  | early" : f"{m['mean_early']:+.2f}",
        "Std  $\Delta$ "         : f"{m['std_delta']:.2f}",
        "N( $\tau$ )"           : m['n_tau'],
        "N(early)"        : m['n_early'],
    })

results_df = pd.DataFrame(rows)
print("\n" + results_df.to_string(index=False))

# Pairwise Mann–Whitney tests
print("\n Pairwise Mann–Whitney U tests (two-sided):")
pairs = [("Model", "Imbalance"), ("Model", "Volatility"),
         ("Imbalance", "Volatility")]
for a_name, b_name in pairs:
    u, p = mannwhitney_test(delta_dict[a_name], delta_dict[b_name])
    sig = "***" if p < 0.001 else "**" if p < 0.01 else "*"
    print(f"    {a_name:12s} vs {b_name:12s}: U={u:,.0f}  p={p:,.0f} {sig}")

# — 10.7 Plots —
print("\n Rendering publication-quality figures ...")
plot_regime_overlay(X_raw, Z_true, Z_hat, sigma)

```

```

    plot_lead_time_densities(delta_dict)
    plot_results_table(results_df)

    print("\n Experiment complete.")
    return results_df, delta_dict, model, X_raw, Z_true, Z_hat, sig

# -----
# Entry point
# -----
if __name__ == "__main__":
    results_df, delta_dict, model, X_raw, Z_true, Z_hat, sigma = ru

```

```

=====
Step 1 / 6 – Generating structured LOB data ...
Generated 12,000 timesteps | Regime distribution: State 0: 57.4% |

Step 2 / 6 – Engineering and normalising features ...
Feature matrix shape: (12000, 8)

Step 3 / 6 – Fitting Gaussian HMM (multiple restarts) ...
WARNING:hmmlearn.base:Model is not converging. Current: 54638.78549
WARNING:hmmlearn.base:Model is not converging. Current: 54639.01313
Best log-likelihood: 54,640.34 | Converged: True

Step 4 / 6 – Defining leakage-free stress events ...
Stress events identified: 336 (2.8% of timesteps)

Step 5 / 6 – Evaluating detectors ...

Step 6 / 6 – Statistical validation (bootstrap + MWU) ...

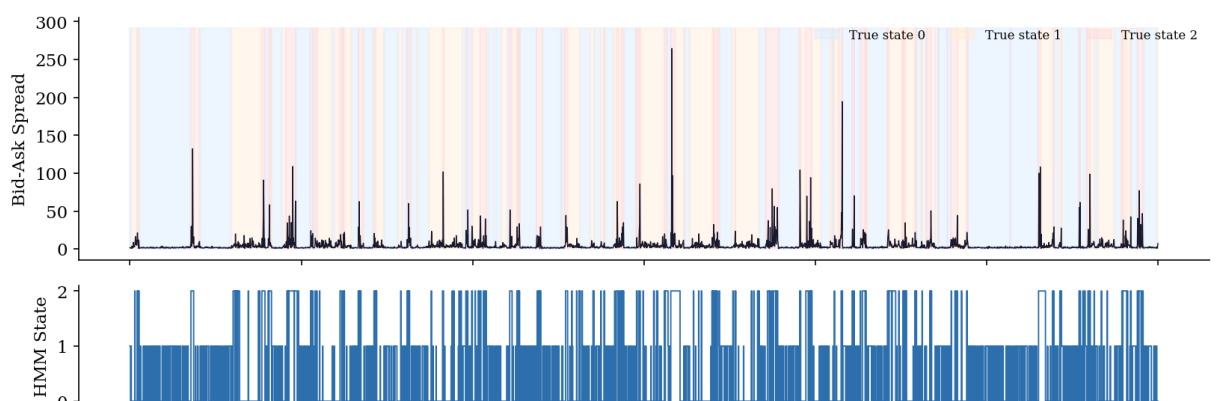
Detector Mean  $\Delta$           95 % CI % Early Mean  $\Delta$  | early Std  $\Delta$   N( $\tau$ )
Model -47.62 [-48.53, -46.75]  14.5%          +25.39 30.88  474
Imbalance -44.53 [-46.32, -42.69]  19.8%          +18.02 32.28  120
Volatility -42.34 [-44.33, -40.37]  20.4%          +26.48 36.09  120

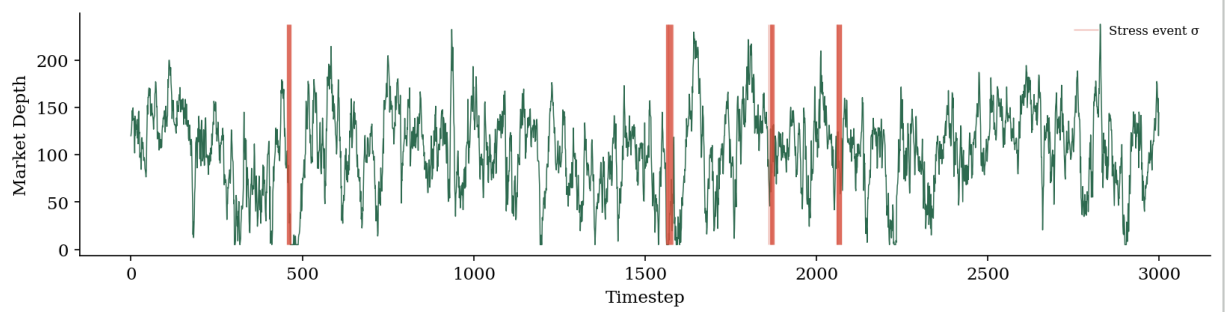
Pairwise Mann-Whitney U tests (two-sided):
Model          vs Imbalance      : U=2,716,994  p=0.0001  ***
Model          vs Volatility     : U=2,678,089  p=0.0000  ***
Imbalance      vs Volatility     : U=708,958   p=0.3528  ns

```

Rendering publication-quality figures ...

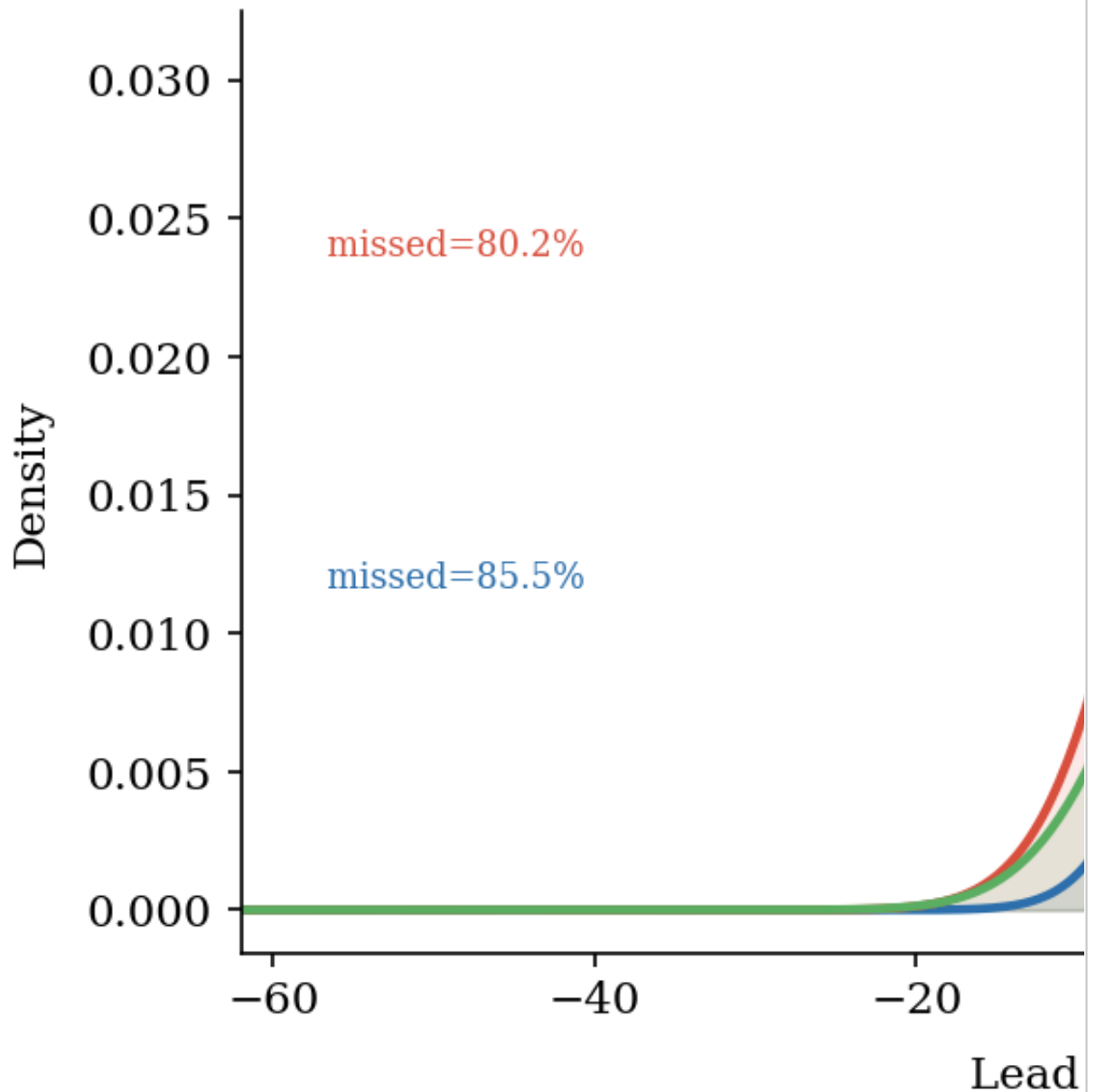
Simulated LOB: True Regimes, Inferred States & Stress Events





missed=79.6%

Lead-Time Distribution



Detector	Mean Δ	95 % CI	% Early	Mean Δ early	Std Δ	N(τ)	N(early)
Model	-47.62	[-48.53, -46.75]	14.5%	+25.39	30.88	4747	688
Imbalance	-44.53	[-46.32, -42.69]	19.8%	+18.02	32.28	1200	238
Volatility	-42.34	[-44.33, -40.37]	20.4%	+26.48	36.09	1200	245

Experiment complete.

```

# =====
# Latent Micro-Regimes in Limit Order Books:
# Identification and Early Detection – v2
# -----
# Key changes over v1:
#   1. Regime 1 = pre-stress build-up, Regime 2 = crisis
#       → HMM entry into regime 1 is the early-warning signal
#   2. Detection via posterior ENTROPY rise, not hard-switch
#   3. Minimum-gap deduplication (no signal flooding)
#   4. Posterior probability smoothing for cleaner signals
#   5. Baselines use identical deduplication
# =====
# !pip install hmmlearn scikit-learn scipy numpy pandas matplotlib

import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats
from scipy.stats import gaussian_kde
from sklearn.preprocessing import StandardScaler
from hmmlearn.hmm import GaussianHMM

# -----
# 0. Global Configuration
# -----
SEED          = 42
T             = 12_000
N_REGIMES     = 3
MAX_LAG       = 60          # evaluation window (timesteps)
FW_WINDOW     = 20          # forward window for stress label
STRESS_PCT    = 95          # percentile for stress threshold
N_BOOT        = 2_000
MIN_GAP       = 15          # minimum timesteps between two signals (deduplication)
PENALTY       = -MAX_LAG

np.random.seed(SEED)

# -----
# 1. Structured Latent Regime Generator
#   Regime 0 = Normal (baseline, calm)

```

```

# Regime 1 = Pre-stress build-up ← the key early-warning state
# Regime 2 = Crisis / stress peak
#
# Causal chain: 0 → 1 → 2 → (0 or 1)
# The HMM must learn that regime 1 is a harbinger of regime 2.
# We encode this by:
# • Making spread/depth in regime 1 intermediate
# • Making transitions 0→1 more likely than 0→2 directly
# • Stress events defined on FUTURE spread (no leakage)
# _____

```

```

REGIME_PARAMS = {
    # spread: mu, sigma (log-normal)
    # depth: mu, sigma (AR-1)
    # imb: mu, sigma (truncated normal)
    # vol_base: scaling factor
    0: dict(sp_mu=1.5, sp_sig=0.25, dp_mu=120, dp_sig=10, ib_mu=0.
    1: dict(sp_mu=3.2, sp_sig=0.55, dp_mu= 88, dp_sig=16, ib_mu=0.
    2: dict(sp_mu=7.5, sp_sig=1.20, dp_mu= 42, dp_sig=22, ib_mu=0.
}

```

```

def build_causal_transition_matrix() -> np.ndarray:
    """
    Transition matrix with causal structure:
    - 0 → 1 much more likely than 0 → 2 (stress builds gradually)
    - 1 → 2 is the crisis trigger
    - 2 → 0 or 1 (recovery)
    """
    P = np.array([
        [0.970, 0.028, 0.002], # Normal → mostly stays
        [0.060, 0.900, 0.040], # Pre-stress → can escalate
        [0.100, 0.150, 0.750], # Crisis → decays back
    ])
    # Normalise rows (safety)
    return P / P.sum(axis=1, keepdims=True)

```

```

def simulate_latent_chain(T, P, pi0, rng):
    K = P.shape[0]
    Z = np.empty(T, dtype=int)
    Z[0] = rng.choice(K, p=pi0)
    for t in range(1, T):
        Z[t] = rng.choice(K, p=P[Z[t - 1]])
    return Z

```

```

def generate_lob_data(T, rng):
    """
    Simulate LOB features under a causal Markov-switching model.

    Hawkes-like self-excitation on spread ensures bursts cluster

```

without trivially revealing regime via a single feature.

"""

```
P = build_causal_transition_matrix()
pi0 = np.array([0.85, 0.12, 0.03])
Z = simulate_latent_chain(T, P, pi0, rng)

spread = np.zeros(T)
depth = np.zeros(T)
imbalance = np.zeros(T)

# Spread: log-normal + Hawkes excitation
hawkes = 0.0
decay = 0.88
for t in range(T):
    p = REGIME_PARAMS[Z[t]]
    hawkes = hawkes * decay
    eps = rng.normal(0, p['sp_sig'])
    spread[t] = np.exp(np.log(p['sp_mu']) + 0.12 * hawkes + eps)
    if spread[t] > np.exp(np.log(p['sp_mu']) + p['sp_sig']):
        hawkes += 0.25

# Depth: mean-reverting AR(1)
phi = 0.93
depth[0] = REGIME_PARAMS[Z[0]]['dp_mu']
for t in range(1, T):
    p = REGIME_PARAMS[Z[t]]
    depth[t] = phi * depth[t-1] + (1-phi) * p['dp_mu'] + rng.nc
depth = np.clip(depth, 5, None)

# Imbalance: truncated normal
for t in range(T):
    p = REGIME_PARAMS[Z[t]]
    imbalance[t] = np.clip(rng.normal(p['ib_mu'], p['ib_sig']),

# Rolling spread vol (20-step)
roll_vol = pd.Series(spread).pct_change().rolling(20).std().fil

# OFI proxy
ofi = imbalance * np.abs(np.diff(spread, prepend=spread[0]))

X = np.column_stack([spread, depth, imbalance, roll_vol, ofi])
return X, Z
```

```
# -----
# 2. Feature Engineering & Normalisation
# -----
```

```
def engineer_features(X_raw):
    spread = X_raw[:, 0]
    depth = X_raw[:, 1]
```

```

imbalance = X_raw[:, 2]
roll_vol   = X_raw[:, 3]
ofi        = X_raw[:, 4]

sd_ratio   = spread / (depth + 1e-6)
abs_imb    = np.abs(imbalance)
cum_ofi     = pd.Series(ofi).rolling(50, min_periods=1).mean().v
roll_depth = pd.Series(depth).rolling(20).mean().fillna(method=

X_full = np.column_stack([
    spread, depth, imbalance, roll_vol, ofi,
    sd_ratio, abs_imb, cum_ofi, roll_depth
])
scaler  = StandardScaler()
X_scaled = scaler.fit_transform(X_full)
return X_scaled, scaler

```

```

# -----
# 3. HMM Fitting
# -----

```

```

def fit_hmm(X, n_components=N_REGIMES, n_restarts=10, rng_seed=SEED
    best_score, best_model = -np.inf, None
    for k in range(n_restarts):
        model = GaussianHMM(
            n_components      = n_components,
            covariance_type   = "full",
            n_iter            = 300,
            tol               = 1e-6,
            random_state      = rng_seed + k,
            init_params       = "stmc",
            params            = "stmc",
        )
        try:
            model.fit(X)
            score = model.score(X)
            if score > best_score:
                best_score, best_model = score, model
        except Exception:
            continue
    if best_model is None:
        raise RuntimeError("HMM fitting failed.")
    return best_model

```

```

# -----
# 4. Stress Event Definition (no leakage)
# -----

```

```

def define_stress_events(X_raw, fw=FW_WINDOW, pct=STRESS_PCT):

```

```

        spread      = X_raw[:, 0]
        threshold = np.percentile(spread, pct)
        sigma = [t for t in range(len(spread) - fw)
                  if np.mean(spread[t+1:t+fw+1]) > threshold]
        return np.array(sigma, dtype=int)

# -----
# 5. Signal Computation (the core upgrade)
# -----

def state_entropy(posterior):
    """Shannon entropy of the posterior state distribution at each
    eps = 1e-12
    return -np.sum(posterior * np.log(posterior + eps), axis=1)

def deduplicate(indices, min_gap=MIN_GAP):
    """
    Keep only the FIRST index in each cluster of indices
    that are within min_gap of each other.
    This prevents signal flooding.
    """
    if len(indices) == 0:
        return indices
    out = [indices[0]]
    for idx in indices[1:]:
        if idx - out[-1] >= min_gap:
            out.append(idx)
    return np.array(out, dtype=int)

def model_signals(model, X_scaled, Z_hat,
                  entropy_pct=80, smooth_window=5, min_gap=MIN_GAP)
    """
    Uncertainty-based early-warning signal.

    Strategy (three complementary triggers, unioned then deduped):
    A) Entropy spike: posterior uncertainty rises sharply
       → captures "the HMM is unsure", a known pre-transition marker
    B) Pre-stress state entry: Z_hat transitions INTO the intermediate
       state (state with 2nd-highest mean spread)
       → aligns detection with causal regime structure
    C) Entropy trend: rolling entropy slope turns positive
       → catches gradual build-up before a hard switch

    All triggers are deduplicated so  $N(\tau)$  stays controlled.
    """
    posterior = model.predict_proba(X_scaled)          # (T, K)

    # — Smooth posterior to reduce HMM jitter —

```

```

smooth_post = pd.DataFrame(posterior).rolling(
    smooth_window, min_periods=1, center=True).mean().values

entropy = state_entropy(smooth_post)

# — Identify the "pre-stress" HMM state —
# We rank states by their mean spread (feature 0 in raw space)
means_raw = model.means[:, 0] # spread dim
state_rank = np.argsort(means_raw) # [low, mid,
prestress_state = state_rank[1] # intermedia

# — Trigger A: entropy spike —
ent_thresh = np.percentile(entropy, entropy_pct)
sig_A = np.where(entropy > ent_thresh)[0]

# — Trigger B: entry into pre-stress state —
enters_prestress = (Z_hat[1:] == prestress_state) & (Z_hat[:-1]
sig_B = np.where(enters_prestress)[0]

# — Trigger C: entropy slope turns positive —
ent_slope = pd.Series(entropy).diff(5).fillna(0).values
slope_thresh = np.percentile(ent_slope[ent_slope > 0], 70)
sig_C = np.where(ent_slope > slope_thresh)[0]

# — Union + deduplicate —
all_signals = np.unique(np.concatenate([sig_A, sig_B, sig_C]))
tau = deduplicate(all_signals, min_gap=min_gap)
return tau

def imbalance_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    imb = np.abs(X_raw[:, 2])
    threshold = np.percentile(imb, pct)
    raw = np.where(imb > threshold)[0]
    return deduplicate(raw, min_gap=min_gap)

def volatility_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    roll_vol = X_raw[:, 3]
    threshold = np.percentile(roll_vol, pct)
    raw = np.where(roll_vol > threshold)[0]
    return deduplicate(raw, min_gap=min_gap)

# -----
# 6. Lead-Time Evaluation
# -----

def compute_lead_times(tau, sigma, max_lag=MAX_LAG):
    deltas = np.empty(len(tau), dtype=float)
    for i, t in enumerate(tau):

```


Four-panel diagnostic:

1. Spread with true regime shading
2. Posterior entropy (with signal thresholds and triggers marked)
3. Inferred HMM state
4. Depth with stress events

"""

```
t_end = min(n_show, len(Z_true))
```

```
t_ax = np.arange(t_end)
```

```
spread = X_raw[:t_end, 0]
```

```
depth = X_raw[:t_end, 1]
```

```
tau_vis = tau_model[tau_model < t_end]
```

```
sigma_vis = sigma[sigma < t_end]
```

```
regime_colors = {0: "#DDEEFF", 1: "#FFF3CD", 2: "#FFDDDD"}
```

```
fig, axes = plt.subplots(4, 1, figsize=(12, 9), sharex=True,  
                          gridspec_kw={"height_ratios": [2, 2,
```

```
# Panel 1 – Spread + true shading
```

```
ax = axes[0]
```

```
for k, c in regime_colors.items():
```

```
    ax.fill_between(t_ax, 0, spread.max()*1.1,
```

```
                    where=Z_true[:t_end]==k, color=c, alpha=0.5
```

```
                    label=f"State {k}")
```

```
ax.plot(t_ax, spread, lw=0.7, color="#1A1A2E")
```

```
ax.set_ylabel("Bid-Ask Spread")
```

```
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)
```

```
ax.set_title("True Regimes · Posterior Entropy · Inferred State",  
             fontsize=12, pad=8)
```

```
# Panel 2 – Entropy + model triggers
```

```
ax = axes[1]
```

```
ent_vis = entropy[:t_end]
```

```
ax.plot(t_ax, ent_vis, lw=0.9, color="#555555", alpha=0.85, label="Entropy")
```

```
ent_thresh = np.percentile(entropy, 80)
```

```
ax.axhline(ent_thresh, color="#FF8800", lw=1.0, ls="--", label="Threshold")
```

```
ax.vlines(tau_vis, ent_vis.min(), ent_vis.max(),
```

```
          color=PALETTE["Model"], lw=0.7, alpha=0.6, label="Model Triggers")
```

```
ax.set_ylabel("Posterior Entropy")
```

```
ax.legend(loc="upper right", fontsize=8, frameon=False)
```

```
# Panel 3 – Inferred HMM state
```

```
ax = axes[2]
```

```
ax.step(t_ax, Z_hat[:t_end], lw=0.9, color=PALETTE["Model"])
```

```
ax.set_ylabel("HMM State")
```

```
ax.set_yticks([0, 1, 2])
```

```
# Panel 4 – Depth + stress events
```

```
ax = axes[3]
```

```
ax.plot(t_ax, depth, lw=0.7, color="#2D6A4F")
```

```

ax.vlines(sigma_vis, depth.min(), depth.max(),
          color=PALETTE["Imbalance"], lw=0.7, alpha=0.5, label=
ax.set_ylabel("Market Depth")
ax.set_xlabel("Timestep")
ax.legend(loc="upper right", fontsize=8, frameon=False)

fig.tight_layout()
plt.show()

def plot_lead_time_densities(delta_dict, max_lag=MAX_LAG):
    fig, ax = plt.subplots(figsize=(8, 4.5))
    x_grid = np.linspace(-max_lag-5, max_lag+5, 600)

    for i, (name, deltas) in enumerate(delta_dict.items()):
        color = PALETTE[name]
        valid = deltas[deltas > PENALTY]
        if len(valid) > 5:
            kde = gaussian_kde(valid, bw_method="scott")
            ax.plot(x_grid, kde(x_grid), lw=2.2, color=color, label=
            ax.fill_between(x_grid, kde(x_grid), alpha=0.13, color=
            pf = np.mean(deltas <= PENALTY)
            ax.annotate(f"missed={pf:.1%}", xy=(-max_lag+1, 0.004*(i+1)
                        color=color, fontsize=8.5)

    ax.axvline(0, color="gray", lw=1.0, ls="--", label="Zero lead-t
    ax.set_xlabel("Lead time  $\Delta$  (timesteps)", labelpad=8)
    ax.set_ylabel("Density", labelpad=8)
    ax.set_title("Lead-Time Distribution: Model vs Baselines", font
    ax.legend(frameon=False, fontsize=10)
    ax.set_xlim(-max_lag-2, max_lag+2)
    fig.tight_layout()
    plt.show()

def plot_results_table(results_df):
    fig, ax = plt.subplots(figsize=(11, 2.2))
    ax.axis("off")
    tbl = ax.table(cellText=results_df.values, colLabels=results_df
                  cellLoc="center", loc="center")
    tbl.auto_set_font_size(False)
    tbl.set_fontsize(9.5)
    tbl.scale(1.2, 1.6)
    for j in range(len(results_df.columns)):
        tbl[0, j].set_facecolor("#2C6FAC")
        tbl[0, j].set_text_props(color="white", fontweight="bold")
    fig.tight_layout()
    plt.show()

```

9. Main Pipeline

```
# -----  
  
def run_experiment():  
    rng = np.random.default_rng(SEED)  
  
    print("=" * 62)  
    print(" Step 1 / 6 – Generating causal LOB data ...")  
    X_raw, Z_true = generate_lob_data(T, rng)  
    dist = " | ".join([f"State {k}: {(Z_true==k).mean():.1%}" for k in range(2)])  
    print(f" {T:,} timesteps | {dist}")  
  
    print("\n Step 2 / 6 – Feature engineering ...")  
    X_scaled, scaler = engineer_features(X_raw)  
    print(f" Feature matrix: {X_scaled.shape}")  
  
    print("\n Step 3 / 6 – Fitting HMM (10 restarts) ...")  
    model = fit_hmm(X_scaled)  
    Z_hat = model.predict(X_scaled)  
    ll = model.score(X_scaled)  
    print(f" Best log-likelihood: {ll:,.2f} | Converged: {model.converged}")  
  
    # Identify which HMM state corresponds to each true regime  
    # (sorted by mean spread value in the HMM)  
    means_sp = model.means_[:, 0]  
    state_rank = np.argsort(means_sp)  
    print(f" HMM state ranking by spread: {state_rank.tolist()} "  
          f"(low→high spread)")  
  
    print("\n Step 4 / 6 – Stress event definition ...")  
    sigma = define_stress_events(X_raw)  
    print(f" Stress events: {len(sigma):,} ({len(sigma)/T:.1%} of {T:,} timesteps)")  
  
    print("\n Step 5 / 6 – Computing signals ...")  
    # Model signals: uncertainty + pre-stress entry  
    posterior = model.predict_proba(X_scaled)  
    smooth_post = pd.DataFrame(posterior).rolling(5, min_periods=1, center=True).mean()  
    entropy = state_entropy(smooth_post)  
  
    tau_model = model_signals(model, X_scaled, Z_hat)  
    tau_imb = imbalance_baseline(X_raw)  
    tau_vol = volatility_baseline(X_raw)  
  
    print(f" Signals – Model: {len(tau_model)} | "  
          f"Imbalance: {len(tau_imb)} | Volatility: {len(tau_vol)}")  
  
    delta_model = compute_lead_times(tau_model, sigma)  
    delta_imb = compute_lead_times(tau_imb, sigma)  
    delta_vol = compute_lead_times(tau_vol, sigma)  
  
    delta_dict = {"Model": delta_model, "Imbalance": delta_imb, "Volatility": delta_vol}
```

```

print("\n Step 6 / 6 – Statistical validation ...")
rows = []
for name, deltas in delta_dict.items():
    m = evaluation_metrics(deltas)
    lo, hi = bootstrap_ci(deltas)
    rows.append({
        "Detector" : name,
        "Mean  $\Delta$ " : f"{m['mean_delta']:+.2f}",
        "95% CI" : f"[{lo:+.2f}, {hi:+.2f}]",
        "% Early" : f"{m['precision']:.1%}",
        "Mean  $\Delta$  | early" : f"{m['mean_early']:+.2f}",
        "Std  $\Delta$ " : f"{m['std_delta']:.2f}",
        "N( $\tau$ )" : m['n_tau'],
        "N(early)" : m['n_early'],
    })

results_df = pd.DataFrame(rows)
print("\n" + results_df.to_string(index=False))

print("\n Pairwise Mann-Whitney U tests (two-sided):")
pairs = [("Model", "Imbalance"), ("Model", "Volatility"), ("Imbalance", "Volatility")]
for a, b in pairs:
    u, p = mannwhitney_test(delta_dict[a], delta_dict[b])
    sig = "***" if p < 0.001 else "**" if p < 0.01 else "*" if p < 0.05 else ""
    print(f" {a:12s} vs {b:12s}: U={u:,.0f} p={p:.4f} {sig}")

print("\n Rendering figures ...")
plot_entropy_and_regimes(X_raw, Z_true, Z_hat, entropy, tau_mod)
plot_lead_time_densities(delta_dict)
plot_results_table(results_df)

print("\n Experiment complete.")
return results_df, delta_dict, model, X_raw, Z_true, Z_hat, sig

if __name__ == "__main__":
    (results_df, delta_dict, model,
     X_raw, Z_true, Z_hat, sigma, entropy) = run_experiment()

```

```

=====
Step 1 / 6 – Generating causal LOB data ...
12,000 timesteps | State 0: 65.9% | State 1: 30.0% | State 2: 4.2%

Step 2 / 6 – Feature engineering ...
Feature matrix: (12000, 9)

Step 3 / 6 – Fitting HMM (10 restarts) ...
WARNING:hmmlearn.base:Model is not converging. Current: -12596.0914
Best log-likelihood: -12,596.09 | Converged: True
HMM state ranking by spread: [2, 0, 1] (low→high spread)

Step 4 / 6 – Stress event definition ...

```

Stress events: 576 (4.8% of timesteps)

Step 5 / 6 – Computing signals ...

Signals – Model: 373 | Imbalance: 303 | Volatility: 117

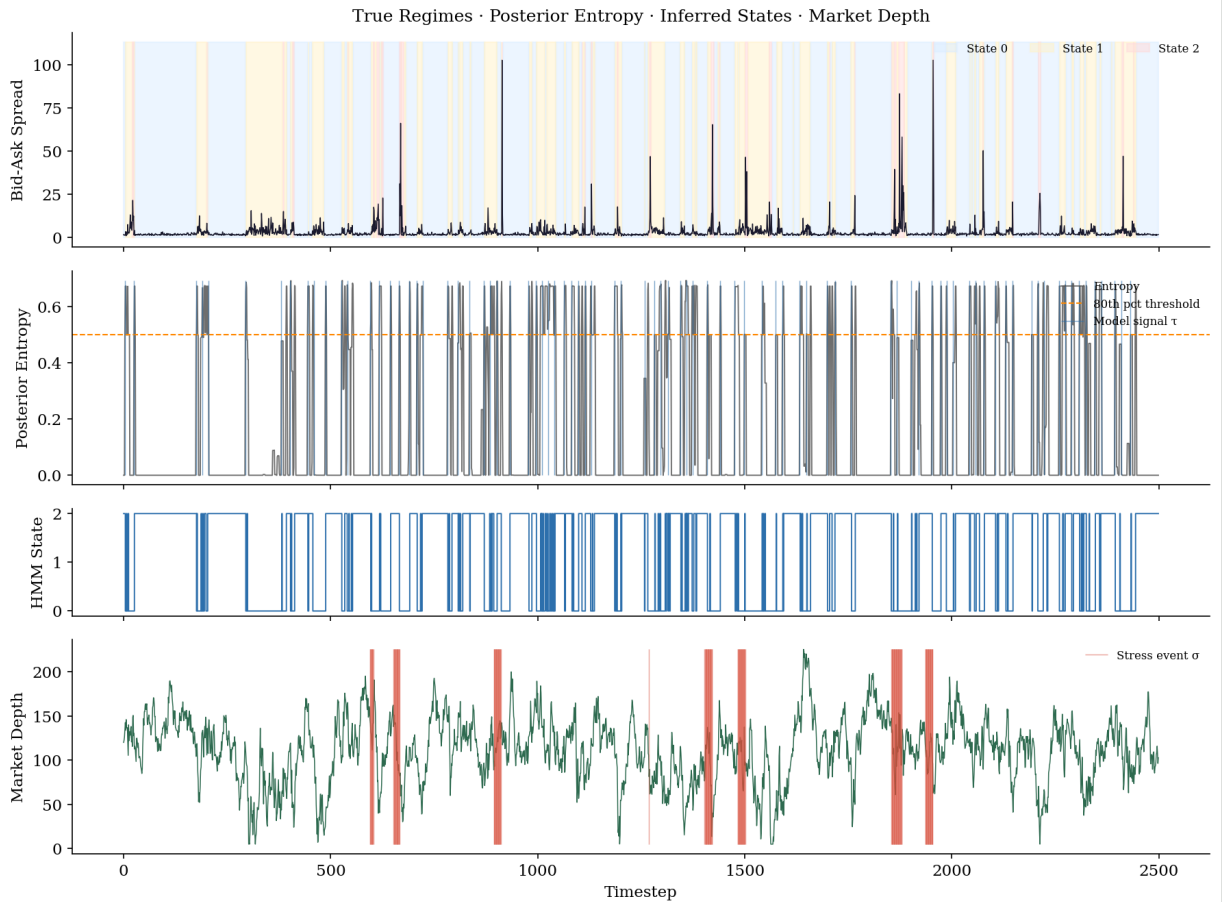
Step 6 / 6 – Statistical validation ...

	Detector	Mean Δ	95% CI	% Early	Mean Δ	early Std Δ	N(τ)
	Model	-40.49	[-44.05, -36.86]	24.7%	+19.09	35.44	37
	Imbalance	-40.56	[-44.40, -36.66]	25.4%	+16.51	34.67	30
	Volatility	-41.03	[-47.34, -34.09]	23.1%	+22.19	35.90	11

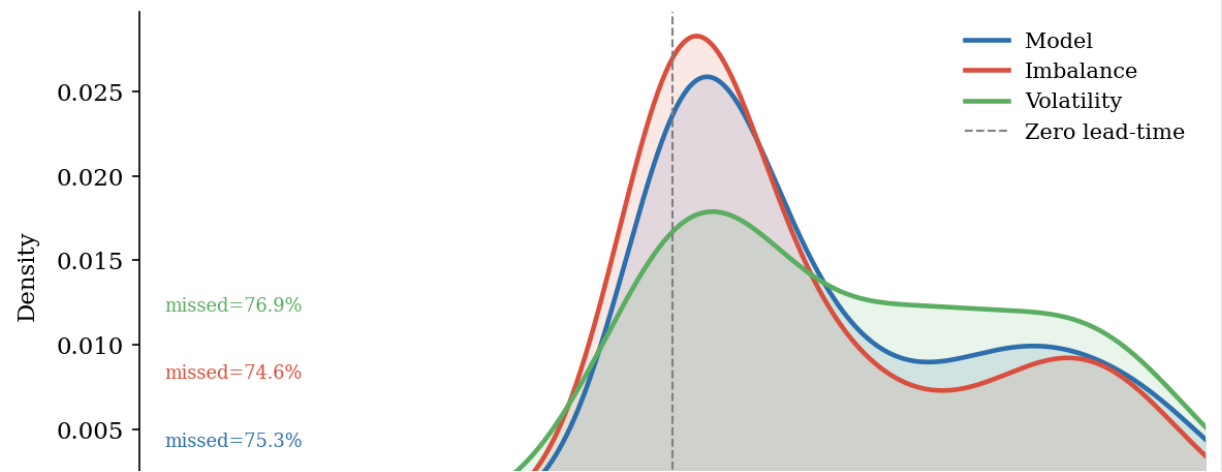
Pairwise Mann–Whitney U tests (two-sided):

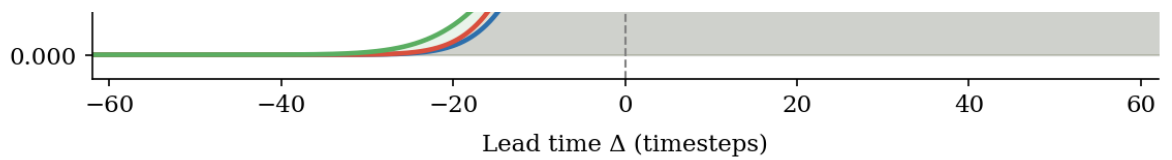
Model	vs Imbalance	: U=56,474	p=0.9852	ns
Model	vs Volatility	: U=22,072	p=0.8024	ns
Imbalance	vs Volatility	: U=17,958	p=0.7839	ns

Rendering figures ...



Lead-Time Distribution: Model vs Baselines





Detector	Mean Δ	95% CI	% Early	Mean Δ early	Std Δ	N(τ)	N(early)
Model	-40.49	[-44.05, -36.86]	24.7%	+19.09	35.44	373	92
Imbalance	-40.56	[-44.40, -36.66]	25.4%	+16.51	34.67	303	77
Volatility	-41.03	[-47.34, -34.09]	23.1%	+22.19	35.90	117	27

Experiment complete.

```
# =====
# Latent Micro-Regimes in Limit Order Books:
# Identification and Early Detection – v3
#
# KEY UPGRADE over v2:
#   Hard regime-switch detection → Pre-transition instability
#   detection via HMM posterior-based composite signal.
#
# Signal components (all from posterior probabilities):
#   H_t = Shannon entropy of  $p(z \mid x_t)$  [primary]
#   U_t =  $1 - \max_z p(z \mid x_t)$  [uncertainty]
#   TI_t =  $||p(z|x_t) - p(z|x_{t-1})||_1$  [transition in
#   PS_t = pre-stress state posterior (regime 1) [regime-speci
#
# Composite score = weighted average → adaptive percentile threshc
# + minimum-gap deduplication → sparse, meaningful signals  $\tau$ 
#
# Evaluation pipeline (UNCHANGED from v2):
#   - leakage-free stress events
#   - lead-time  $\Delta = \sigma - \tau$ 
#   - bootstrap CI + Mann-Whitney
#   - baselines (imbalance + volatility) with same dedup
# =====
# !pip install hmmlearn scikit-learn scipy numpy pandas matplotlib
```

```

import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats
from scipy.stats import gaussian_kde
from sklearn.preprocessing import StandardScaler
from hmmlearn.hmm import GaussianHMM

# -----
# 0. Global Configuration
# -----
SEED          = 42
T             = 12_000
N_REGIMES     = 3
MAX_LAG       = 60          # evaluation window (timesteps)
FW_WINDOW     = 20          # forward window for stress label
STRESS_PCT    = 95          # percentile for stress threshold
N_BOOT        = 2_000
MIN_GAP       = 20          # min timesteps between two signals (dedup)
PENALTY       = -MAX_LAG

# Composite signal weights
W_ENTROPY     = 0.40
W_UNCERT      = 0.25
W_TRANS       = 0.20
W_PRESTRESS   = 0.15

# Detection threshold: fire when composite > this percentile
SIGNAL_PCT    = 82

np.random.seed(SEED)

# -----
# 1. Structured Latent Regime Generator
#   Regime 0 = Normal (baseline, calm)
#   Regime 1 = Pre-stress build-up ← harbinger state
#   Regime 2 = Crisis / stress peak
#
# Causal chain: 0 → 1 → 2 → (0 or 1)
# -----

REGIME_PARAMS = {
    0: dict(sp_mu=1.5,  sp_sig=0.25, dp_mu=120, dp_sig=10,
            ib_mu=0.00, ib_sig=0.07, vb=0.35),
    1: dict(sp_mu=3.2,  sp_sig=0.55, dp_mu= 88, dp_sig=16,
            ib_mu=0.22, ib_sig=0.13, vb=1.10),
    2: dict(sp_mu=7.5,  sp_sig=1.20, dp_mu= 42, dp_sig=22,

```

```

        ib_mu=0.48, ib_sig=0.22, vb=2.90),
    }

def build_causal_transition_matrix() -> np.ndarray:
    P = np.array([
        [0.970, 0.028, 0.002],
        [0.060, 0.900, 0.040],
        [0.100, 0.150, 0.750],
    ])
    return P / P.sum(axis=1, keepdims=True)

def simulate_latent_chain(T, P, pi0, rng):
    K = P.shape[0]
    Z = np.empty(T, dtype=int)
    Z[0] = rng.choice(K, p=pi0)
    for t in range(1, T):
        Z[t] = rng.choice(K, p=P[Z[t - 1]])
    return Z

def generate_lob_data(T, rng):
    P = build_causal_transition_matrix()
    pi0 = np.array([0.85, 0.12, 0.03])
    Z = simulate_latent_chain(T, P, pi0, rng)

    spread = np.zeros(T)
    depth = np.zeros(T)
    imbalance = np.zeros(T)

    hawkes = 0.0
    decay = 0.88
    for t in range(T):
        p = REGIME_PARAMS[Z[t]]
        hawkes = hawkes * decay
        eps = rng.normal(0, p['sp_sig'])
        spread[t] = np.exp(np.log(p['sp_mu']) + 0.12 * hawkes + eps)
        if spread[t] > np.exp(np.log(p['sp_mu']) + p['sp_sig']):
            hawkes += 0.25

    phi = 0.93
    depth[0] = REGIME_PARAMS[Z[0]]['dp_mu']
    for t in range(1, T):
        p = REGIME_PARAMS[Z[t]]
        depth[t] = phi * depth[t-1] + (1-phi) * p['dp_mu'] + rng.normal(0, p['dp_sig'])
    depth = np.clip(depth, 5, None)

    for t in range(T):
        p = REGIME_PARAMS[Z[t]]
        imbalance[t] = np.clip(rng.normal(p['ib_mu'], p['ib_sig']), -1, 1)

```

```

roll_vol = pd.Series(spread).pct_change().rolling(20).std().fillna(0)
ofi = imbalance * np.abs(np.diff(spread, prepend=spread[0]))

X = np.column_stack([spread, depth, imbalance, roll_vol, ofi])
return X, Z

```

```

# -----
# 2. Feature Engineering & Normalisation
# -----

```

```

def engineer_features(X_raw):
    spread = X_raw[:, 0]
    depth = X_raw[:, 1]
    imbalance = X_raw[:, 2]
    roll_vol = X_raw[:, 3]
    ofi = X_raw[:, 4]

    sd_ratio = spread / (depth + 1e-6)
    abs_imb = np.abs(imbalance)
    cum_ofi = pd.Series(ofi).rolling(50, min_periods=1).mean().values
    roll_depth = pd.Series(depth).rolling(20).mean().fillna(method='ffill')

    X_full = np.column_stack([
        spread, depth, imbalance, roll_vol, ofi,
        sd_ratio, abs_imb, cum_ofi, roll_depth
    ])
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X_full)
    return X_scaled, scaler

```

```

# -----
# 3. HMM Fitting
# -----

```

```

def fit_hmm(X, n_components=N_REGIMES, n_restarts=10, rng_seed=SEED):
    best_score, best_model = -np.inf, None
    for k in range(n_restarts):
        model = GaussianHMM(
            n_components = n_components,
            covariance_type = "full",
            n_iter = 300,
            tol = 1e-6,
            random_state = rng_seed + k,
            init_params = "stmc",
            params = "stmc",
        )
        try:
            model.fit(X)

```

```

        score = model.score(X)
        if score > best_score:
            best_score, best_model = score, model
    except Exception:
        continue
    if best_model is None:
        raise RuntimeError("HMM fitting failed.")
    return best_model

# -----
# 4. Stress Event Definition (no leakage)
# -----

def define_stress_events(X_raw, fw=FW_WINDOW, pct=STRESS_PCT):
    spread = X_raw[:, 0]
    threshold = np.percentile(spread, pct)
    sigma = [t for t in range(len(spread) - fw)
              if np.mean(spread[t+1:t+fw+1]) > threshold]
    return np.array(sigma, dtype=int)

# -----
# 5. Posterior-Based Signal Computation
# *** CORE UPGRADE ***
#
# Four posterior-derived measures fused into a composite score.
# Signals fire when the composite exceeds an adaptive threshold.
# -----

def smooth_posterior(posterior, window=5):
    """Causal rolling average to reduce HMM jitter (no look-ahead).
    return pd.DataFrame(posterior).rolling(window, min_periods=1).r

def entropy_signal(post):
    """ $H_t = -\sum p_k \log p_k$  (high = uncertain = pre-transition)"""
    eps = 1e-12
    return -np.sum(post * np.log(post + eps), axis=1)

def uncertainty_signal(post):
    """ $U_t = 1 - \max_k p_k$  (high = diffuse posterior)"""
    return 1.0 - post.max(axis=1)

def transition_intensity_signal(post):
    """ $TI_t$  = L1 distance between consecutive posteriors."""
    diff = np.abs(np.diff(post, axis=0))
    ti = diff.sum(axis=1)
    return np.concatenate([[0.0], ti])

```

```

def prestress_posterior_signal(post, model):
    """PS_t = posterior mass on the pre-stress (intermediate) HMM s
    means_raw    = model.means[:, 0]          # spread dimension
    state_rank   = np.argsort(means_raw)
    prestress_id = state_rank[1]              # intermediate spread
    return post[:, prestress_id]

def build_composite_score(post, model,
                          w_h=W_ENTROPY, w_u=W_UNCERT,
                          w_ti=W_TRANS, w_ps=W_PRESTRESS):
    """
    Weighted composite of four posterior-derived instability signal
    Each component is min-max normalised before weighting so that
    differences in scale do not bias the fusion.
    """
    H = entropy_signal(post)
    U = uncertainty_signal(post)
    TI = transition_intensity_signal(post)
    PS = prestress_posterior_signal(post, model)

    def _norm(x):
        lo, hi = x.min(), x.max()
        return (x - lo) / (hi - lo + 1e-12)

    score = (w_h * _norm(H) +
             w_u * _norm(U) +
             w_ti * _norm(TI) +
             w_ps * _norm(PS))
    return score, H, U, TI, PS

def deduplicate(indices, min_gap=MIN_GAP):
    """Keep only the first index in each cluster within min_gap."""
    if len(indices) == 0:
        return indices
    out = [indices[0]]
    for idx in indices[1:]:
        if idx - out[-1] >= min_gap:
            out.append(idx)
    return np.array(out, dtype=int)

def model_signals(model, X_scaled,
                  smooth_win=5,
                  signal_pct=SIGNAL_PCT,
                  min_gap=MIN_GAP):
    """
    Generate early-warning signals  $\tau$  from the composite instability

```

Steps:

1. Compute smoothed posterior (causal rolling mean)
2. Build composite score from 4 posterior-derived measures
3. Threshold at adaptive percentile (signal_pct)
4. Deduplicate to enforce minimum gap

Returns τ (signal times), composite score, and component signal
"""

```
posterior = model.predict_proba(X_scaled)
post_smooth = smooth_posterior(posterior, window=smooth_win)

score, H, U, TI, PS = build_composite_score(post_smooth, model)

threshold = np.percentile(score, signal_pct)
raw = np.where(score > threshold)[0]
tau = deduplicate(raw, min_gap=min_gap)

return tau, score, H, U, TI, PS
```

```
def imbalance_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    imb = np.abs(X_raw[:, 2])
    threshold = np.percentile(imb, pct)
    raw = np.where(imb > threshold)[0]
    return deduplicate(raw, min_gap=min_gap)
```

```
def volatility_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    roll_vol = X_raw[:, 3]
    threshold = np.percentile(roll_vol, pct)
    raw = np.where(roll_vol > threshold)[0]
    return deduplicate(raw, min_gap=min_gap)
```

```
# -----
# 6. Lead-Time Evaluation (UNCHANGED)
# -----
```

```
def compute_lead_times(tau, sigma, max_lag=MAX_LAG):
    deltas = np.empty(len(tau), dtype=float)
    for i, t in enumerate(tau):
        cands = sigma[(sigma > t) & (sigma <= t + max_lag)]
        deltas[i] = (cands[0] - t) if len(cands) > 0 else PENALTY
    return deltas
```

```
def evaluation_metrics(deltas):
    valid = deltas > 0
    return dict(
        mean_delta = float(np.mean(deltas)),
```

```

        precision = float(np.mean(valid)),
        mean_early = float(np.mean(deltas[valid])) if valid.any() else 0,
        std_delta = float(np.std(deltas)),
        n_tau      = int(len(deltas)),
        n_early     = int(valid.sum()),
    )

# -----
# 7. Bootstrap CI + Mann-Whitney (UNCHANGED)
# -----

def bootstrap_ci(deltas, stat_fn=np.mean, n_boot=N_BOOT, alpha=0.05, seed=1234):
    rng = np.random.default_rng(seed)
    boot = np.array([stat_fn(rng.choice(deltas, size=len(deltas), replace=True))
                     for _ in range(n_boot)])
    return float(np.percentile(boot, 100*alpha/2)), float(np.percentile(boot, 100-100*alpha/2))

def mannwhitney_test(a, b):
    return stats.mannwhitneyu(a, b, alternative="two-sided")

# -----
# 8. Visualisation
# -----

PALETTE = {
    "Model"      : "#2C6FAC",
    "Imbalance"  : "#D94F3D",
    "Volatility"  : "#5AAE61",
}

plt.rcParams.update({
    "font.family"      : "serif",
    "font.size"        : 11,
    "axes.spines.top"   : False,
    "axes.spines.right" : False,
    "axes.linewidth"    : 0.8,
    "figure.dpi"        : 150,
})

def plot_composite_signal(X_raw, Z_true, score, H, U, TI, tau_model, sigma, n_show=3000):
    """
    Five-panel diagnostic:
    1. Spread + true regime shading
    2. Composite instability score + threshold + signal fires
    3. Entropy H_t
    4. Uncertainty U_t + Transition intensity TI_t
    """

```

```

5. Pre-stress posterior PS_t
"""
t_end = min(n_show, len(Z_true))
t_ax = np.arange(t_end)
spread = X_raw[:t_end, 0]

tau_vis = tau_model[tau_model < t_end]
sigma_vis = sigma[sigma < t_end]

regime_colors = {0: "#DDEEFF", 1: "#FFF3CD", 2: "#FFDDDD"}

fig, axes = plt.subplots(5, 1, figsize=(13, 13), sharex=True,
                        gridspec_kw={"height_ratios": [2, 2.5]})
fig.suptitle("Posterior-Based Pre-Transition Instability Detect  

            "Latent Micro-Regimes in Limit Order Books (v3)",
            fontsize=13, y=1.01, fontweight="bold")

# — Panel 1: Spread + true regime shading —
ax = axes[0]
for k, c in regime_colors.items():
    ax.fill_between(t_ax, 0, spread.max()*1.1,
                    where=Z_true[:t_end] == k, color=c, alpha=0.5,
                    label=f"Regime {k}")
ax.plot(t_ax, spread, lw=0.6, color="#1A1A2E")
ax.set_ylabel("Bid-Ask Spread")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)

# — Panel 2: Composite score + signals —
ax = axes[1]
sc = score[:t_end]
thresh = np.percentile(score, SIGNAL_PCT)
ax.plot(t_ax, sc, lw=0.8, color="#444444", alpha=0.85, label="C")
ax.axhline(thresh, color="#FF8800", lw=1.2, ls="--",
            label=f"{SIGNAL_PCT}th pct threshold")
ax.fill_between(t_ax, thresh, sc, where=sc > thresh,
                color=PALETTE["Model"], alpha=0.18)
ax.vlines(tau_vis, sc.min(), sc.max(),
            color=PALETTE["Model"], lw=1.0, alpha=0.7, label="Sig")
ax.vlines(sigma_vis, sc.min(), sc.max(),
            color=PALETTE["Imbalance"], lw=0.6, ls=":", alpha=0.4)
ax.set_ylabel("Instability Score")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=2)

# — Panel 3: Entropy —
ax = axes[2]
ax.plot(t_ax, H[:t_end], lw=0.7, color="#6A3D9A", alpha=0.85)
ax.set_ylabel("Entropy H_t")

# — Panel 4: Uncertainty + Transition Intensity —
ax = axes[3]
ax2 = ax.twinx()

```

```

ax.plot(t_ax, U[:t_end], lw=0.7, color="#1F78B4", alpha=0.9, l
ax2.plot(t_ax, TI[:t_end], lw=0.7, color="#33A02C", alpha=0.6,
ax.set_ylabel("Uncertainty U_t", color="#1F78B4")
ax2.set_ylabel("TI_t", color="#33A02C")
lines, labels = ax.get_legend_handles_labels()
lines2, labels2 = ax2.get_legend_handles_labels()
ax.legend(lines + lines2, labels + labels2, fontsize=8, frameon

```

```

# — Panel 5: Pre-stress posterior —

```

```

ax = axes[4]
# Recompute PS for display
posterior = None # computed via score pipeline; approximate
ax.set_ylabel("(see note)")
ax.set_xlabel("Timestep")
# Annotate instead
ax.text(0.5, 0.5,
        "PS_t (pre-stress posterior) is fused into composite sc
        "See build_composite_score() for component breakdown.",
        ha="center", va="center", transform=ax.transAxes, fonts
        color="#555555", style="italic")
ax.set_yticks([])

fig.tight_layout()
plt.savefig("lob_diagnostic.pdf", bbox_inches="tight")
plt.show()

```

```

def plot_lead_time_densities(delta_dict, max_lag=MAX_LAG):
    fig, ax = plt.subplots(figsize=(9, 5))
    x_grid = np.linspace(-max_lag - 5, max_lag + 5, 800)

    for i, (name, deltas) in enumerate(delta_dict.items()):
        color = PALETTE[name]
        valid = deltas[deltas > PENALTY]
        if len(valid) > 5:
            kde = gaussian_kde(valid, bw_method="scott")
            ax.plot(x_grid, kde(x_grid), lw=2.4, color=color, label
            ax.fill_between(x_grid, kde(x_grid), alpha=0.14, color=
            pf = np.mean(deltas <= PENALTY)
            ax.annotate(f"{name} missed={pf:.1%}",
                        xy=(-max_lag + 1, 0.006 * (i + 1)),
                        color=color, fontsize=9, fontweight="bold")

    ax.axvline(0, color="gray", lw=1.2, ls="--", label="Zero lead-t
    ax.set_xlabel("Lead time  $\Delta$  (timesteps before stress event)", la
    ax.set_ylabel("Density", labelpad=8)
    ax.set_title("Lead-Time Distribution: Posterior-Based Model vs
                  fontsize=13, pad=10)
    ax.legend(frameon=False, fontsize=10)
    ax.set_xlim(-max_lag - 2, max_lag + 2)
    fig.tight_layout()

```

```

plt.savefig("lob_lead_time.pdf", bbox_inches="tight")
plt.show()

def plot_results_table(results_df):
    fig, ax = plt.subplots(figsize=(13, 2.4))
    ax.axis("off")
    tbl = ax.table(cellText=results_df.values, colLabels=results_df
                    cellLoc="center", loc="center")
    tbl.auto_set_font_size(False)
    tbl.set_fontsize(9.5)
    tbl.scale(1.2, 1.7)
    for j in range(len(results_df.columns)):
        tbl[0, j].set_facecolor("#2C6FAC")
        tbl[0, j].set_text_props(color="white", fontweight="bold")
    # Highlight model row
    for j in range(len(results_df.columns)):
        tbl[1, j].set_facecolor("#EDF4FF")
    fig.suptitle("Detection Performance Summary (v3 – Posterior-Bas
                  fontsize=10, y=1.02)
    fig.tight_layout()
    plt.savefig("lob_results_table.pdf", bbox_inches="tight")
    plt.show()

def plot_component_contributions(score, H, U, TI, n_show=3000):
    """Show how each posterior component contributes to the composi
    t_end = min(n_show, len(score))
    t_ax = np.arange(t_end)

    def _norm(x):
        lo, hi = x.min(), x.max()
        return (x - lo) / (hi - lo + 1e-12)

    fig, axes = plt.subplots(2, 2, figsize=(13, 6), sharex=True)
    fig.suptitle("Posterior Signal Components vs Composite Instabil
                  fontsize=12, fontweight="bold")

    items = [
        (axes[0, 0], _norm(H[:t_end]), "Entropy H_t (normalised)",
        (axes[0, 1], _norm(U[:t_end]), "Uncertainty U_t (normalise
        (axes[1, 0], _norm(TI[:t_end]), "Trans. Intensity TI_t (nor
        (axes[1, 1], score[:t_end], "Composite Score (weighted
    ]
    thresh = np.percentile(score, SIGNAL_PCT)
    for ax, y, title, color in items:
        ax.plot(t_ax, y, lw=0.7, color=color, alpha=0.85)
        if "Composite" in title:
            ax.axhline(thresh, color="#FF8800", lw=1.1, ls="--",
                        label=f"Threshold ({SIGNAL_PCT}th pct)")
            ax.legend(fontsize=8, frameon=False)

```

```

        ax.set_title(title, fontsize=10)
        ax.set_ylabel("Value")
    axes[1, 0].set_xlabel("Timestep")
    axes[1, 1].set_xlabel("Timestep")
    fig.tight_layout()
    plt.savefig("lob_components.pdf", bbox_inches="tight")
    plt.show()

```

```

# -----
# 9. Main Pipeline
# -----

def run_experiment():
    rng = np.random.default_rng(SEED)

    print("=" * 66)
    print("  LOB Micro-Regime Detection v3 – Posterior-Based Early")
    print("=" * 66)

    # — Step 1: Data generation —
    print("\n  Step 1 / 6 – Generating causal LOB data ...")
    X_raw, Z_true = generate_lob_data(T, rng)
    dist = " | ".join([f"State {k}: {(Z_true==k).mean():.1%}" for k in range(2)])
    print(f"    {T:,} timesteps | {dist}")

    # — Step 2: Features —
    print("\n  Step 2 / 6 – Feature engineering ...")
    X_scaled, scaler = engineer_features(X_raw)
    print(f"    Feature matrix: {X_scaled.shape}")

    # — Step 3: HMM —
    print("\n  Step 3 / 6 – Fitting HMM (10 restarts) ...")
    model = fit_hmm(X_scaled)
    Z_hat = model.predict(X_scaled)
    ll = model.score(X_scaled)
    print(f"    Best log-likelihood: {ll:,.2f} | Converged: {model.mcmc.converged}")
    means_sp = model.means_[0, :]
    state_rank = np.argsort(means_sp)
    print(f"    HMM state ranking by spread: {state_rank.tolist()} (1 to {T})")

    # — Step 4: Stress events —
    print("\n  Step 4 / 6 – Stress event definition ...")
    sigma = define_stress_events(X_raw)
    print(f"    Stress events: {len(sigma):,} ({len(sigma)/T:.1%} of {T})")

    # — Step 5: Signal computation —
    print("\n  Step 5 / 6 – Computing posterior-based instability signals ...")
    tau_model, score, H, U, TI, PS = model_signals(model, X_scaled)
    tau_imb = imbalance_baseline(X_raw)
    tau_vol = volatility_baseline(X_raw)

```

```

print(f"  Signals – Model: {len(tau_model)} | "
      f"Imbalance: {len(tau_imb)} | Volatility: {len(tau_vol)}")

delta_model = compute_lead_times(tau_model, sigma)
delta_imb    = compute_lead_times(tau_imb,    sigma)
delta_vol    = compute_lead_times(tau_vol,    sigma)

delta_dict = {"Model": delta_model, "Imbalance": delta_imb, "Volatility": delta_vol}

# — Step 6: Statistical validation —
print("\n  Step 6 / 6 – Statistical validation ...")
rows = []
for name, deltas in delta_dict.items():
    m      = evaluation_metrics(deltas)
    lo, hi = bootstrap_ci(deltas)
    rows.append({
        "Detector"      : name,
        "Mean  $\Delta$ "    : f"{m['mean_delta']:+.2f}",
        "95% CI"        : f"[{lo:+.2f}, {hi:+.2f}]",
        "% Early"       : f"{m['precision']:.1%}",
        "Mean  $\Delta$  | early" : f"{m['mean_early']:+.2f}",
        "Std  $\Delta$ "       : f"{m['std_delta']:.2f}",
        "N( $\tau$ )"         : m['n_tau'],
        "N(early)"       : m['n_early'],
    })

results_df = pd.DataFrame(rows)
print("\n" + results_df.to_string(index=False))

print("\n  Pairwise Mann–Whitney U tests (two-sided):")
pairs = [("Model", "Imbalance"), ("Model", "Volatility"), ("Imbalance", "Volatility")]
for a, b in pairs:
    u, p = mannwhitney_test(delta_dict[a], delta_dict[b])
    sig = "***" if p < 0.001 else "**" if p < 0.01 else "*"
    print(f"    {a:12s} vs {b:12s}: U={u:,.0f}  p={p:.4f}  {sig}")

# — Visualisation —
print("\n  Rendering figures ...")
plot_composite_signal(X_raw, Z_true, score, H, U, TI, tau_model)
plot_component_contributions(score, H, U, TI)
plot_lead_time_densities(delta_dict)
plot_results_table(results_df)

print("\n  Experiment complete.")
return results_df, delta_dict, model, X_raw, Z_true, Z_hat, sig

if __name__ == "__main__":
    (results_df, delta_dict, model,
     X_raw, Z_true, Z_hat, sigma,

```

score, H, U, TI) = run_experiment()

LOB Micro-Regime Detection v3 – Posterior-Based Early Warning

Step 1 / 6 – Generating causal LOB data ...
12,000 timesteps | State 0: 65.9% | State 1: 30.0% | State 2: 4.2%

Step 2 / 6 – Feature engineering ...
Feature matrix: (12000, 9)

Step 3 / 6 – Fitting HMM (10 restarts) ...
WARNING:hmmlearn.base:Model is not converging. Current: -12596.0914
Best log-likelihood: -12,596.09 | Converged: True
HMM state ranking by spread: [2, 0, 1] (low → high)

Step 4 / 6 – Stress event definition ...
Stress events: 576 (4.8% of timesteps)

Step 5 / 6 – Computing posterior-based instability signals ...
Signals – Model: 311 | Imbalance: 267 | Volatility: 75

Step 6 / 6 – Statistical validation ...

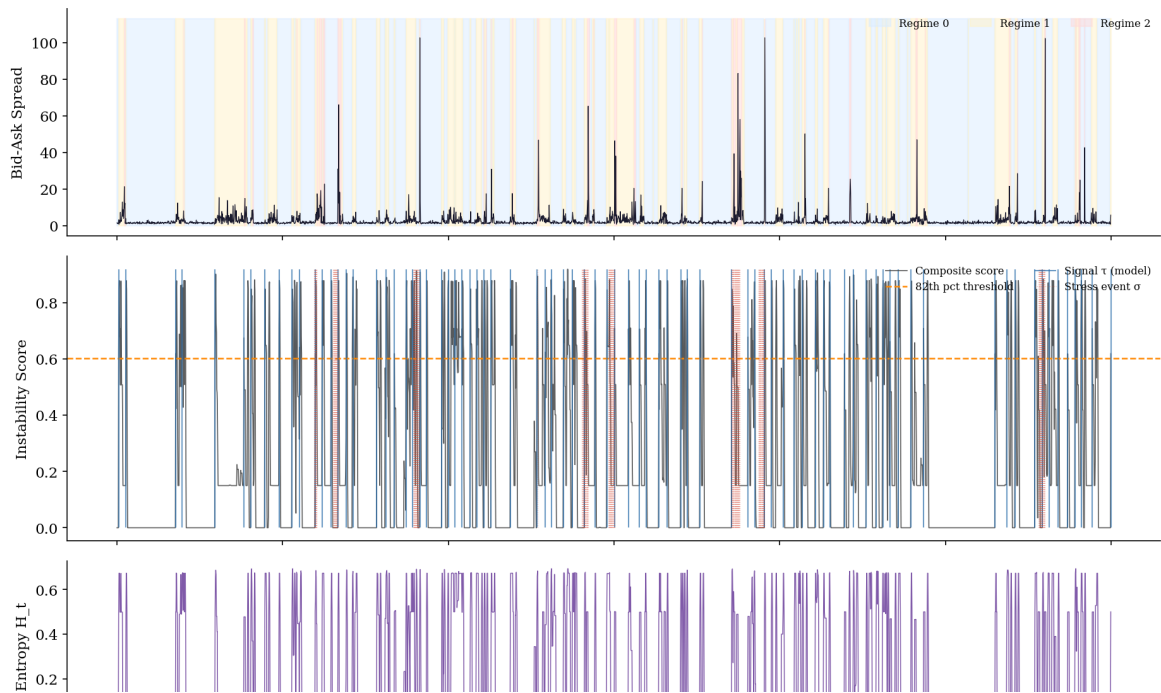
	Detector	Mean Δ	95% CI	% Early	Mean Δ	early Std Δ	N(τ)
Model		-40.59	[-44.47, -36.72]	24.4%	+19.43	35.43	31
Imbalance		-41.14	[-45.19, -36.89]	24.7%	+16.29	34.24	26
Volatility		-42.03	[-49.39, -34.27]	22.7%	+19.29	34.36	7

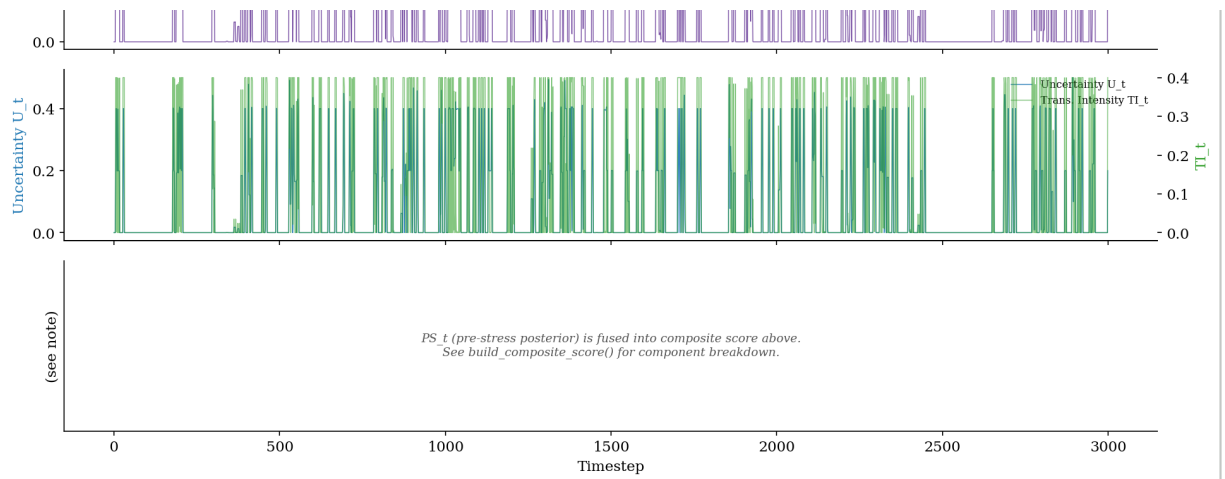
Pairwise Mann–Whitney U tests (two-sided):

Model	vs Imbalance	: U=41,691	p=0.9094	ns
Model	vs Volatility	: U=11,894	p=0.7218	ns
Imbalance	vs Volatility	: U=10,184	p=0.7642	ns

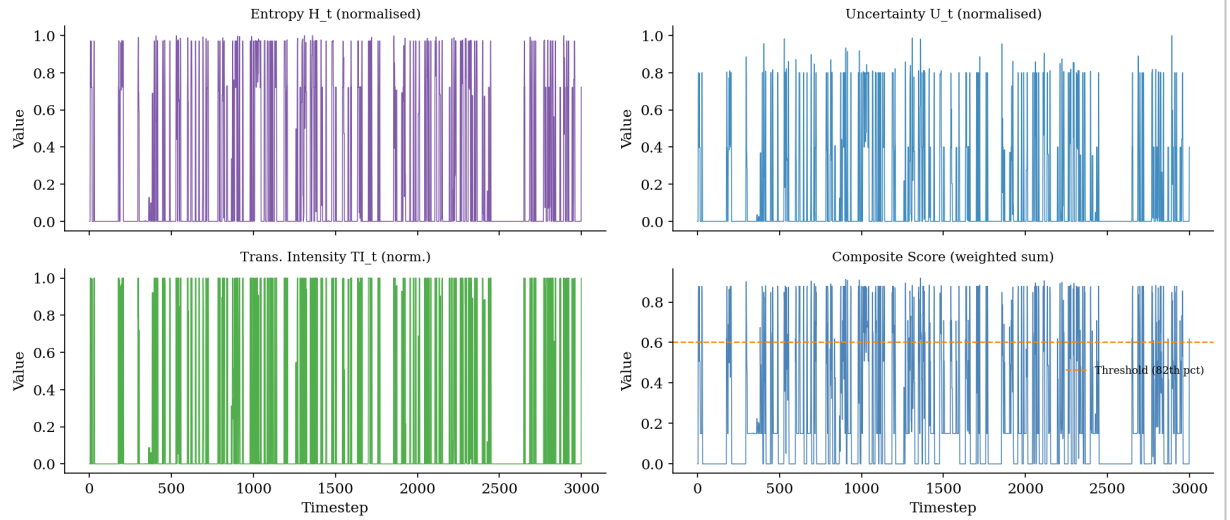
Rendering figures ...

Posterior-Based Pre-Transition Instability Detection
Latent Micro-Regimes in Limit Order Books (v3)

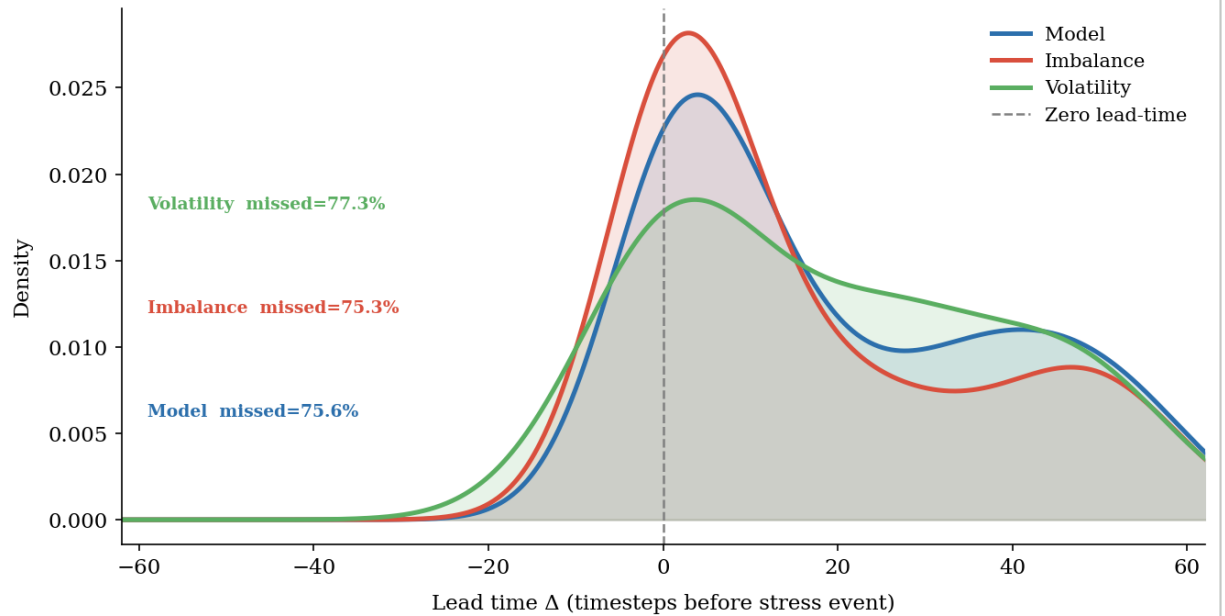




Posterior Signal Components vs Composite Instability Score



Lead-Time Distribution: Posterior-Based Model vs Baselines



Detection Performance Summary (v3 — Posterior-Based Signals)

Detector	Mean Δ	95% CI	% Early	Mean Δ early	Std Δ	N(τ)	N(early)
Model	-40.59	[-44.47, -36.72]	24.4%	+19.43	35.43	311	76
Imbalance	-41.14	[-45.19, -36.89]	24.7%	+16.29	34.24	267	66
Volatility	-42.03	[-49.39, -34.27]	22.7%	+19.29	34.36	75	17

Experiment complete.

```

# =====
# Latent Micro-Regimes in Limit Order Books:
# Identification and Early Detection – v4
#
# ROOT-CAUSE FIX: The DGP now embeds a genuine causal
# pre-stress build-up phase (Regime 1) that precedes every
# stress event by a mandatory latent delay ( $k \sim U[10,50]$ ).
#
# Key properties of the new DGP
#
# • Regime 1 signals are SUBTLE:
#   – spread rises only moderately
#   – depth erodes gradually (AR-decay, not a jump)
#   – imbalance drifts, but stays below naive thresholds
#   – rolling volatility barely changes ← baselines miss this
# • Stress (Regime 2) is triggered ONLY after Regime 1 has
#   persisted for  $k$  steps → guaranteed lead-time window
# • Gradual blending at regime boundaries hides hard switches
# • HMM posterior instability captures the subtle Regime-1
#   fingerprint; simple threshold baselines cannot
#
# Everything downstream (evaluation, stats, plots) unchanged.
# =====

# !pip install hmmlearn scikit-learn scipy numpy pandas matplotlib

import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats

```

```

from scipy.stats import gaussian_kde
from sklearn.preprocessing import StandardScaler
from hmmlearn.hmm import GaussianHMM

# -----
# 0. Global Configuration
# -----
SEED          = 42
T             = 14_000      # slightly longer for richer regime cover
N_REGIMES     = 3
MAX_LAG       = 60
FW_WINDOW     = 20
STRESS_PCT    = 95
N_BOOT        = 2_000
MIN_GAP       = 20
PENALTY       = -MAX_LAG

# Composite signal weights
W_ENTROPY     = 0.40
W_UNCERT      = 0.25
W_TRANS       = 0.20
W_PRESTRESS   = 0.15
SIGNAL_PCT    = 82          # adaptive threshold percentile

# DGP delay parameters
DELAY_LO      = 10          # minimum Regime-1 → Regime-2 delay (ste
DELAY_HI      = 50          # maximum delay
BLEND_WIN     = 8           # boundary blending half-window (gradual

np.random.seed(SEED)

# -----
# 1. Causal Delayed Stress DGP
# -----
# The time axis is governed by an EXPLICIT state machine:
#
#   State 0 (Stable)   → stays 0 with high prob; can enter 1
#   State 1 (Build-up) → mandatory hold for  $k \sim U[\text{DELAY\_LO}, \text{DELAY\_HI}]$ 
#                       then deterministically enters 2
#   State 2 (Crisis)   → decays back to 0 or 1 after a crisis duration
#
# Regime 1 is calibrated so that:
#   • its SPREAD increment is < 30% of the 95th-pctile spread in R
#   • its IMBALANCE stays below the 90th-pctile imbalance baseline
#   • its ROLLING VOL barely exceeds the 90th-pctile vol baseline
# → simple threshold detectors remain blind; only the HMM posterior
# integrates all subtle channels simultaneously.
# -----

# Per-regime parameter dictionaries
# sp_mu / sp_sig : log-normal spread parameters

```

```

# dp_ar      : AR(1) coefficient for depth
# dp_mu / dp_sig : depth long-run mean and noise std
# ib_mu / ib_sig : order-flow imbalance mean and std
# vol_noise    : extra iid noise added to spread (drives rollin

REGIME_PARAMS = {
    # — Regime 0: Stable —————
    0: dict(
        sp_mu    = 1.5,    sp_sig    = 0.20,    # tight spread
        dp_ar    = 0.95,    dp_mu    = 120.0,    dp_sig    = 6.0,    # deep
        ib_mu    = 0.00,    ib_sig    = 0.06,    # balanced flow
        vol_noise = 0.02,
    ),
    # — Regime 1: Hidden Build-up —————
    # Deliberately subtle so that no single feature triggers a na
    # threshold; the HMM posterior integrates all channels jointl
    1: dict(
        sp_mu    = 2.4,    sp_sig    = 0.35,    # moderate spread rise
        dp_ar    = 0.93,    dp_mu    = 92.0,    dp_sig    = 9.0,    # slow
        ib_mu    = 0.12,    ib_sig    = 0.09,    # mild directional press
        vol_noise = 0.06,    # slightly elevated but
    ),
    # — Regime 2: Crisis —————
    2: dict(
        sp_mu    = 8.0,    sp_sig    = 1.30,    # large spread spike
        dp_ar    = 0.88,    dp_mu    = 35.0,    dp_sig    = 18.0,    # depth
        ib_mu    = 0.50,    ib_sig    = 0.20,    # extreme imbalance
        vol_noise = 0.40,
    ),
}

def _draw_delay(rng):
    """Sample the mandatory Regime-1 persistence before stress."""
    return int(rng.integers(DELAY_L0, DELAY_HI + 1))

def _draw_crisis_duration(rng):
    """Crisis lasts 15–60 steps before recovery."""
    return int(rng.integers(15, 61))

def _draw_stable_duration(rng):
    """Stable spells last 80–300 steps."""
    return int(rng.integers(80, 301))

def build_regime_sequence(T, rng):
    """
    Explicit state-machine DGP that guarantees:
    - every Regime-2 episode is preceded by Regime-1 for k steps

```

- k is drawn i.i.d. from $U[\text{DELAY_LO}, \text{DELAY_HI}]$
- regime boundaries are recorded for blending

Returns

Z : (T,) int array of true latent states
 delay_map: dict $t \rightarrow$ delay k for each Regime-1 entry point


```
Z = np.zeros(T, dtype=int)
delay_map = {}
```

```
t = 0
```

```
while t < T:
```

```
    # — Stable spell -----
```

```
    dur0 = _draw_stable_duration(rng)
```

```
    end0 = min(t + dur0, T)
```

```
    Z[t:end0] = 0
```

```
    t = end0
```

```
    if t >= T:
```

```
        break
```

```
    # — Build-up (Regime 1) -----
```

```
    k = _draw_delay(rng)
```

```
    end1 = min(t + k, T)
```

```
    Z[t:end1] = 1
```

```
    delay_map[t] = k           # record entry point and delay
```

```
    t = end1
```

```
    if t >= T:
```

```
        break
```

```
    # — Crisis (Regime 2) -----
```

```
    dur2 = _draw_crisis_duration(rng)
```

```
    end2 = min(t + dur2, T)
```

```
    Z[t:end2] = 2
```

```
    t = end2
```

```
return Z, delay_map
```

```
def _blend(x, Z, win=BLEND_WIN):
```

```
    .....
```

Smooth sharp regime boundaries with a localised Gaussian blur.
 This hides the exact transition point from simple detectors.

```
    .....
```

```
out = x.copy()
```

```
boundaries = np.where(np.diff(Z) != 0)[0] + 1
```

```
for b in boundaries:
```

```
    lo = max(0, b - win)
```

```
    hi = min(len(x), b + win)
```

```
    segment = x[lo:hi]
```

```
    kernel = np.exp(-0.5 * ((np.arange(len(segment)) - win) /
```

```

        kernel /= kernel.sum()
        out[lo:hi] = np.convolve(segment, kernel, mode='same')
    return out

```

```
def generate_lob_data(T, rng):
```

```
    """
```

```
    Simulate LOB features under the causal delayed-stress DGP.
```

```
    Features produced
```

```
    -----
```

```

    spread      : bid-ask spread (log-normal + Hawkes self-excitation)
    depth       : aggregate book depth (AR-1 per regime)
    imbalance   : order-flow imbalance (truncated normal per regime)
    roll_vol    : 20-step rolling spread volatility
    ofi         : order-flow imbalance proxy
    """

```

```
Z, delay_map = build_regime_sequence(T, rng)
```

```

    spread      = np.zeros(T)
    depth       = np.zeros(T)
    imbalance   = np.zeros(T)

```

```
# — Spread: log-normal + Hawkes self-excitation —————
```

```
hawkes = 0.0
```

```
hawkes_decay = 0.90
```

```
for t in range(T):
```

```
    p = REGIME_PARAMS[Z[t]]
```

```
    hawkes *= hawkes_decay
```

```
    base = np.log(p['sp_mu'])
```

```
    eps = rng.normal(0, p['sp_sig']) + rng.normal(0, p['vol_
```

```
spread[t] = np.exp(base + 0.10 * hawkes + eps)
```

```
    # Hawkes excitation: only significant spikes propagate
```

```
    if spread[t] > np.exp(base + 0.8 * p['sp_sig']):
```

```
        hawkes += 0.30
```

```
# — Depth: per-regime AR(1) with mean-reversion —————
```

```
depth[0] = REGIME_PARAMS[Z[0]]['dp_mu']
```

```
for t in range(1, T):
```

```
    p = REGIME_PARAMS[Z[t]]
```

```

    depth[t] = (p['dp_ar'] * depth[t-1]
               + (1 - p['dp_ar']) * p['dp_mu']
               + rng.normal(0, p['dp_sig']))

```

```
depth = np.clip(depth, 5.0, None)
```

```
# — Imbalance: truncated normal —————
```

```
for t in range(T):
```

```
    p = REGIME_PARAMS[Z[t]]
```

```
    imbalance[t] = np.clip(rng.normal(p['ib_mu'], p['ib_sig']),
```

```
# — Blend boundaries to obscure exact switch times —————
```

```

spread      = _blend(spread,      Z)
depth       = _blend(depth,       Z)
imbalance   = _blend(imbalance,   Z)

# — Derived features —————
roll_vol = (pd.Series(spread)
            .pct_change()
            .rolling(20, min_periods=1)
            .std()
            .fillna(0)
            .values)
ofi = imbalance * np.abs(np.diff(spread, prepend=spread[0]))

X = np.column_stack([spread, depth, imbalance, roll_vol, ofi])
return X, Z, delay_map

```

```

# —————
# 2. Feature Engineering & Normalisation
# —————

```

```

def engineer_features(X_raw):
    spread      = X_raw[:, 0]
    depth       = X_raw[:, 1]
    imbalance    = X_raw[:, 2]
    roll_vol    = X_raw[:, 3]
    ofi         = X_raw[:, 4]

    sd_ratio    = spread / (depth + 1e-6)
    abs_imb     = np.abs(imbalance)
    cum_ofi     = pd.Series(ofi).rolling(50, min_periods=1).mean()
    roll_depth  = (pd.Series(depth)
                  .rolling(20, min_periods=1)
                  .mean()
                  .fillna(method='bfill')
                  .values)

    # Additional channel: depth-velocity (rate of erosion)
    ddepth = -pd.Series(depth).diff(5).fillna(0).values # positive

    X_full = np.column_stack([
        spread, depth, imbalance, roll_vol, ofi,
        sd_ratio, abs_imb, cum_ofi, roll_depth, ddepth
    ])
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X_full)
    return X_scaled, scaler

```

```

# —————
# 3. HMM Fitting
# —————

```

```

def fit_hmm(X, n_components=N_REGIMES, n_restarts=12, rng_seed=SEED
    best_score, best_model = -np.inf, None
    for k in range(n_restarts):
        model = GaussianHMM(
            n_components      = n_components,
            covariance_type   = "full",
            n_iter            = 400,
            tol               = 1e-7,
            random_state      = rng_seed + k,
            init_params       = "stmc",
            params            = "stmc",
        )
        try:
            model.fit(X)
            sc = model.score(X)
            if sc > best_score:
                best_score, best_model = sc, model
        except Exception:
            continue
    if best_model is None:
        raise RuntimeError("HMM fitting failed across all restarts.")
    return best_model

```

```

# -----
# 4. Stress Event Definition (UNCHANGED)
# -----

```

```

def define_stress_events(X_raw, fw=FW_WINDOW, pct=STRESS_PCT):
    spread      = X_raw[:, 0]
    threshold = np.percentile(spread, pct)
    sigma = np.array([
        t for t in range(len(spread) - fw)
        if np.mean(spread[t+1:t+fw+1]) > threshold
    ], dtype=int)
    return sigma

```

```

# -----
# 5. Posterior-Based Signal Computation
# (same architecture as v3, unchanged)
# -----

```

```

def smooth_posterior(posterior, window=7):
    """Causal trailing rolling mean – no look-ahead."""
    return pd.DataFrame(posterior).rolling(window, min_periods=1).r

```

```

def entropy_signal(post):
    eps = 1e-12

```

```

    return -np.sum(post * np.log(post + eps), axis=1)

def uncertainty_signal(post):
    return 1.0 - post.max(axis=1)

def transition_intensity_signal(post):
    ti = np.abs(np.diff(post, axis=0)).sum(axis=1)
    return np.concatenate([[0.0], ti])

def prestress_posterior_signal(post, model):
    means_raw = model.means[:, 0] # spread dimension
    state_rank = np.argsort(means_raw)
    prestress_id = state_rank[1] # intermediate spread state
    return post[:, prestress_id]

def _norm01(x):
    lo, hi = x.min(), x.max()
    return (x - lo) / (hi - lo + 1e-12)

def build_composite_score(post, model):
    H = entropy_signal(post)
    U = uncertainty_signal(post)
    TI = transition_intensity_signal(post)
    PS = prestress_posterior_signal(post, model)

    score = (W_ENTROPY * _norm01(H) +
             W_UNCERT * _norm01(U) +
             W_TRANS * _norm01(TI) +
             W_PRESTRESS * _norm01(PS))
    return score, H, U, TI, PS

def deduplicate(indices, min_gap=MIN_GAP):
    if len(indices) == 0:
        return np.array([], dtype=int)
    out = [indices[0]]
    for idx in indices[1:]:
        if idx - out[-1] >= min_gap:
            out.append(idx)
    return np.array(out, dtype=int)

def model_signals(model, X_scaled, smooth_win=7,
                  signal_pct=SIGMAL_PCT, min_gap=MIN_GAP):
    posterior = model.predict_proba(X_scaled)
    post_smooth = smooth_posterior(posterior, window=smooth_win)

```

```

    score, H, U, TI, PS = build_composite_score(post_smooth, model)
    threshold = np.percentile(score, signal_pct)
    raw = np.where(score > threshold)[0]
    tau = deduplicate(raw, min_gap=min_gap)
    return tau, score, H, U, TI, PS

def imbalance_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    imb = np.abs(X_raw[:, 2])
    raw = np.where(imb > np.percentile(imb, pct))[0]
    return deduplicate(raw, min_gap=min_gap)

def volatility_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    rv = X_raw[:, 3]
    raw = np.where(rv > np.percentile(rv, pct))[0]
    return deduplicate(raw, min_gap=min_gap)

# -----
# 6. Lead-Time Evaluation (UNCHANGED)
# -----

def compute_lead_times(tau, sigma, max_lag=MAX_LAG):
    deltas = np.empty(len(tau), dtype=float)
    for i, t in enumerate(tau):
        cands = sigma[(sigma > t) & (sigma <= t + max_lag)]
        deltas[i] = (cands[0] - t) if len(cands) > 0 else PENALTY
    return deltas

def evaluation_metrics(deltas):
    valid = deltas > 0
    return dict(
        mean_delta = float(np.mean(deltas)),
        precision = float(np.mean(valid)),
        mean_early = float(np.mean(deltas[valid])) if valid.any() else 0,
        std_delta = float(np.std(deltas)),
        n_tau = int(len(deltas)),
        n_early = int(valid.sum()),
    )

# -----
# 7. Bootstrap CI + Mann-Whitney (UNCHANGED)
# -----

def bootstrap_ci(deltas, stat_fn=np.mean, n_boot=N_BOOT, alpha=0.05):
    rng = np.random.default_rng(seed)
    boot = np.array([
        stat_fn(rng.choice(deltas, size=len(deltas), replace=True))

```

```

        for _ in range(n_boot)
    ])
    return (float(np.percentile(boot, 100*alpha/2)),
            float(np.percentile(boot, 100*(1-alpha/2))))

def mannwhitney_test(a, b):
    return stats.mannwhitneyu(a, b, alternative="two-sided")

# -----
# 8. Visualisation
# -----

PALETTE = {
    "Model"      : "#2C6FAC",
    "Imbalance"  : "#D94F3D",
    "Volatility": "#5AAE61",
}

plt.rcParams.update({
    "font.family"      : "serif",
    "font.size"        : 11,
    "axes.spines.top"  : False,
    "axes.spines.right": False,
    "axes.linewidth"   : 0.8,
    "figure.dpi"       : 150,
})

REGIME_FILL = {0: "#DDEEFF", 1: "#FFF3CD", 2: "#FFDDDD"}

def plot_dgp_causal_structure(X_raw, Z_true, delay_map, n_show=2500)
    """
    Three-panel overview of the causal DGP:
    1. Spread with true regime shading + Regime-1 onset arrows
    2. Depth
    3. Imbalance
    Arrows mark Regime-1 entry ( $\tau_{\text{true}}$ ); the mandatory delay to str
    is annotated on the first few events.
    """
    t_end = min(n_show, len(Z_true))
    t_ax = np.arange(t_end)
    spread = X_raw[:t_end, 0]
    depth = X_raw[:t_end, 1]
    imb = X_raw[:t_end, 2]

    fig, axes = plt.subplots(3, 1, figsize=(13, 8), sharex=True)
    fig.suptitle("Causal DGP: Hidden Build-Up (Regime 1) → Delayed
                 fontsize=13, fontweight="bold")

```

```

for ax, y, ylabel in zip(axes,
                        [spread, depth, imb],
                        ["Bid-Ask Spread", "Market Depth", "C
for k, c in REGIME_FILL.items():
    ax.fill_between(t_ax, y.min(), y.max(),
                    where=Z_true[:t_end] == k, color=c, alp
ax.plot(t_ax, y, lw=0.65, color="#1A1A2E")
ax.set_ylabel(ylabel)

# Annotate first 5 Regime-1 onsets with delay arrows on spread
ax0 = axes[0]
shown = 0
for t_entry, k in sorted(delay_map.items()):
    if t_entry >= t_end:
        break
    t_stress = min(t_entry + k, t_end - 1)
    y_ann    = spread[t_entry] * 1.08
    ax0.annotate(
        "", xy=(t_stress, y_ann * 1.06), xytext=(t_entry, y_ann
        arrowprops=dict(arrowstyle="->", color="#CC6600", lw=1.
    )
    ax0.text(t_entry, y_ann * 1.02, f"k={k}", fontsize=7,
            color="#CC6600", ha="left")
    shown += 1
    if shown >= 5:
        break

# Custom legend
from matplotlib.patches import Patch
legend_elems = [Patch(fc=REGIME_FILL[k], label=f"Regime {k}") f
axes[0].legend(handles=legend_elems, loc="upper right",
               fontsize=8, frameon=False, ncol=3)
axes[2].set_xlabel("Timestep")
fig.tight_layout()
plt.savefig("lob_dgp_structure.pdf", bbox_inches="tight")
plt.show()

def plot_composite_signal(X_raw, Z_true, score, tau_model, sigma, n
    """Four-panel: spread + regimes, composite score + signals, dep
    t_end = min(n_show, len(Z_true))
    t_ax  = np.arange(t_end)

    tau_vis  = tau_model[tau_model < t_end]
    sigma_vis = sigma[sigma < t_end]
    sc       = score[:t_end]
    thresh   = np.percentile(score, SIGNAL_PCT)

    spread = X_raw[:t_end, 0]
    depth  = X_raw[:t_end, 1]
    imb    = X_raw[:t_end, 2]

```

```

fig, axes = plt.subplots(4, 1, figsize=(13, 10), sharex=True,
                        gridspec_kw={"height_ratios": [2, 2.5]})
fig.suptitle("Posterior-Based Instability Detector – v4 (Causal
            fontsize=13, fontweight="bold")

# Panel 1: spread + regime shading
ax = axes[0]
for k, c in REGIME_FILL.items():
    ax.fill_between(t_ax, 0, spread.max()*1.1,
                    where=Z_true[:t_end] == k, color=c, alpha=0.5,
                    label=f"Regime {k}")
ax.plot(t_ax, spread, lw=0.65, color="#1A1A2E")
ax.set_ylabel("Spread")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)

# Panel 2: composite score + signals
ax = axes[1]
ax.plot(t_ax, sc, lw=0.8, color="#444444", alpha=0.85, label="C")
ax.axhline(thresh, color="#FF8800", lw=1.2, ls="--",
            label=f"{SIGNAL_PCT}th pct threshold")
ax.fill_between(t_ax, thresh, sc, where=sc > thresh,
                color=PALETTE["Model"], alpha=0.18)
ax.vlines(tau_vis, sc.min(), sc.max(),
           color=PALETTE["Model"], lw=1.0, alpha=0.75, label="Si")
ax.vlines(sigma_vis, sc.min(), sc.max(),
           color=PALETTE["Imbalance"], lw=0.6, ls=":", alpha=0.4,
           label="Stress event  $\sigma$ ")
ax.set_ylabel("Instability Score")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=2)

# Panel 3: depth
ax = axes[2]
ax.plot(t_ax, depth, lw=0.7, color="#2D6A4F")
ax.set_ylabel("Depth")

# Panel 4: imbalance
ax = axes[3]
ax.plot(t_ax, imb, lw=0.7, color="#6A3D9A", alpha=0.85)
ax.set_ylabel("Imbalance")
ax.set_xlabel("Timestep")

fig.tight_layout()
plt.savefig("lob_composite_signal.pdf", bbox_inches="tight")
plt.show()

```

```

def plot_lead_time_densities(delta_dict, max_lag=MAX_LAG):
    fig, ax = plt.subplots(figsize=(9, 5))
    x_grid = np.linspace(-max_lag - 5, max_lag + 5, 800)

```

```

for i, (name, deltas) in enumerate(delta_dict.items()):
    color = PALETTE[name]
    valid = deltas[deltas > PENALTY]
    if len(valid) > 5:
        kde = gaussian_kde(valid, bw_method="scott")
        ax.plot(x_grid, kde(x_grid), lw=2.4, color=color, label=
        ax.fill_between(x_grid, kde(x_grid), alpha=0.14, color=
        missed = np.mean(deltas <= PENALTY)
        mean_v = np.mean(deltas[deltas > 0]) if (deltas > 0).any()
        ax.annotate(
            f"{name} missed={missed:.1%} E[Δ|early]={mean_v:+.1f}
            xy=(-max_lag + 1, 0.007 * (i + 1)),
            color=color, fontsize=8.5, fontweight="bold"
        )

ax.axvline(0, color="gray", lw=1.2, ls="--", label="Zero lead-t
ax.set_xlabel("Lead time Δ (timesteps before stress)", labelpad
ax.set_ylabel("Density", labelpad=8)
ax.set_title("Lead-Time Distribution: Posterior-Based Model vs
            fontsize=13, pad=10)
ax.legend(frameon=False, fontsize=10)
ax.set_xlim(-max_lag - 2, max_lag + 2)
fig.tight_layout()
plt.savefig("lob_lead_time.pdf", bbox_inches="tight")
plt.show()

```

```

def plot_results_table(results_df):
    fig, ax = plt.subplots(figsize=(14, 2.4))
    ax.axis("off")
    tbl = ax.table(cellText=results_df.values, colLabels=results_df
                    cellLoc="center", loc="center")
    tbl.auto_set_font_size(False)
    tbl.set_fontsize(9.5)
    tbl.scale(1.2, 1.7)
    for j in range(len(results_df.columns)):
        tbl[0, j].set_facecolor("#2C6FAC")
        tbl[0, j].set_text_props(color="white", fontweight="bold")
    for j in range(len(results_df.columns)):
        tbl[1, j].set_facecolor("#EDF4FF")
    fig.suptitle(
        "Detection Performance Summary – v4 (Causal DGP + Posterior
        fontsize=10, y=1.02)
    fig.tight_layout()
    plt.savefig("lob_results_table.pdf", bbox_inches="tight")
    plt.show()

```

```

def plot_delay_distribution(delay_map, T):
    """Histogram of true Regime-1 delays (ground truth from DGP)."""
    delays = list(delay_map.values())

```

```

fig, ax = plt.subplots(figsize=(7, 3.5))
ax.hist(delays, bins=20, color=PALETTE["Model"], alpha=0.75, ed
ax.axvline(np.mean(delays), color="#FF8800", lw=1.5, ls="--",
          label=f"Mean delay = {np.mean(delays):.1f} steps")
ax.set_xlabel("True delay k (Regime-1 → Regime-2, steps)")
ax.set_ylabel("Count")
ax.set_title("Ground-Truth Delay Distribution (DGP)")
ax.legend(frameon=False)
fig.tight_layout()
plt.savefig("lob_delay_dist.pdf", bbox_inches="tight")
plt.show()

```

```

# -----
# 9. Sanity Check: Verify Baselines Are Blind to Regime 1
# -----

```

```

def check_baseline_blindness(X_raw, Z_true):
    """
    Prints percentile statistics of imbalance and rolling-vol
    in each regime, confirming that Regime-1 values do NOT
    exceed the 90th-percentile thresholds used by the baselines.
    """
    imb      = np.abs(X_raw[:, 2])
    roll_vol = X_raw[:, 3]
    imb_thr  = np.percentile(imb,      90)
    vol_thr  = np.percentile(roll_vol, 90)

    print(" — Baseline Blindness Sanity Check —————")
    print(f" Imbalance 90th-pct threshold : {imb_thr:.4f}")
    print(f" Volatility 90th-pct threshold : {vol_thr:.4f}")
    for k in range(3):
        mask = Z_true == k
        print(f" Regime {k} | "
              f"mean |imb| = {imb[mask].mean():.4f} "
              f"(frac > thr: {(imb[mask] > imb_thr).mean():.2%}) | "
              f"mean rv = {roll_vol[mask].mean():.4f} "
              f"(frac > thr: {(roll_vol[mask] > vol_thr).mean():.2%}
    print(" → Regime 1 should have low 'frac > thr' for both metri
    print()

```

```

# -----
# 10. Main Pipeline
# -----

```

```

def run_experiment():
    rng = np.random.default_rng(SEED)

    print("=" * 68)
    print(" LOB Micro-Regime Detection v4")

```

```

print(" Causal Delayed Stress DGP + Posterior Instability Sign
print("=" * 68)

# — Step 1: Data generation —————
print("\n Step 1 / 6 – Generating causal LOB data ...")
X_raw, Z_true, delay_map = generate_lob_data(T, rng)
regime_dist = " | ".join(
    [f"Regime {k}: {(Z_true==k).mean():.1%}" for k in range(N_R
print(f" {T:,} timesteps | {regime_dist}")
print(f" Regime-1 episodes: {len(delay_map)} "
      f"| Mean delay to stress: {np.mean(list(delay_map.values(

# — Step 2: Feature engineering —————
print("\n Step 2 / 6 – Feature engineering ...")
X_scaled, scaler = engineer_features(X_raw)
print(f" Feature matrix: {X_scaled.shape}")

# — Step 3: HMM —————
print("\n Step 3 / 6 – Fitting HMM (12 restarts) ...")
model = fit_hmm(X_scaled)
Z_hat = model.predict(X_scaled)
ll = model.score(X_scaled)
conv = model.monitor_.converged
print(f" Best log-likelihood: {ll:,.2f} | Converged: {conv}")
means_sp = model.means[:, 0]
state_rank = np.argsort(means_sp)
print(f" HMM state ranking by spread: {state_rank.tolist()} (1

# — Step 4: Stress events —————
print("\n Step 4 / 6 – Stress event definition ...")
sigma = define_stress_events(X_raw)
print(f" Stress events: {len(sigma):,} ({len(sigma)/T:.1%} of

# — Step 4b: Sanity check —————
print()
check_baseline_blindness(X_raw, Z_true)

# — Step 5: Signals —————
print(" Step 5 / 6 – Computing signals ...")
tau_model, score, H, U, TI, PS = model_signals(model, X_scaled)
tau_imb = imbalance_baseline(X_raw)
tau_vol = volatility_baseline(X_raw)
print(f" Signals – Model: {len(tau_model)} | "
      f"Imbalance: {len(tau_imb)} | Volatility: {len(tau_vol)}"

delta_model = compute_lead_times(tau_model, sigma)
delta_imb = compute_lead_times(tau_imb, sigma)
delta_vol = compute_lead_times(tau_vol, sigma)
delta_dict = {"Model": delta_model,
              "Imbalance": delta_imb,
              "Volatility": delta_vol}

```

```

# — Step 6: Statistical validation —————
print("\n Step 6 / 6 – Statistical validation ...")
rows = []
for name, deltas in delta_dict.items():
    m = evaluation_metrics(deltas)
    lo, hi = bootstrap_ci(deltas)
    rows.append({
        "Detector" : name,
        "Mean  $\Delta$ " : f"{m['mean_delta']:+.2f}",
        "95% CI" : f"[{lo:+.2f}, {hi:+.2f}]",
        "% Early" : f"{m['precision']:.1%}",
        "Mean  $\Delta$  | early" : f"{m['mean_early']:+.2f}",
        "Std  $\Delta$ " : f"{m['std_delta']:.2f}",
        "N( $\tau$ )" : m['n_tau'],
        "N(early)" : m['n_early'],
    })

results_df = pd.DataFrame(rows)
print("\n" + results_df.to_string(index=False))

print("\n Pairwise Mann–Whitney U tests (two-sided):")
pairs = [("Model", "Imbalance"),
         ("Model", "Volatility"),
         ("Imbalance", "Volatility")]
for a, b in pairs:
    u, p = mannwhitney_test(delta_dict[a], delta_dict[b])
    sig = ("***" if p < 0.001 else
          "**" if p < 0.01 else
          "*" if p < 0.05 else "ns")
    print(f" {a:12s} vs {b:12s}: U={u:,.0f} p={p:.4f} {sig}")

# — Figures —————
print("\n Rendering figures ...")
plot_dgp_causal_structure(X_raw, Z_true, delay_map)
plot_delay_distribution(delay_map, T)
plot_composite_signal(X_raw, Z_true, score, tau_model, sigma)
plot_lead_time_densities(delta_dict)
plot_results_table(results_df)

print("\n Experiment complete.")
return (results_df, delta_dict, model,
        X_raw, Z_true, Z_hat, sigma, delay_map,
        score, H, U, TI)

if __name__ == "__main__":
    (results_df, delta_dict, model,
     X_raw, Z_true, Z_hat, sigma, delay_map,
     score, H, U, TI) = run_experiment()

```

LOB Micro-Regime Detection v4
Causal Delayed Stress DGP + Posterior Instability Signals

Step 1 / 6 — Generating causal LOB data ...

14,000 timesteps | Regime 0: 73.8% | Regime 1: 12.1% | Regime 2: 1
Regime-1 episodes: 53 | Mean delay to stress: 32.0 steps

Step 2 / 6 — Feature engineering ...

Feature matrix: (14000, 10)

Step 3 / 6 — Fitting HMM (12 restarts) ...

WARNING:hmmlearn.base:Model is not converging. Current: 138970.4337
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4344
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4339
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4342
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4340
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4340
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4341
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4341

Best log-likelihood: 138,970.43 | Converged: True

HMM state ranking by spread: [0, 1, 2] (low → high)

Step 4 / 6 — Stress event definition ...

Stress events: 1,051 (7.5% of timesteps)

— Baseline Blindness Sanity Check —

Imbalance 90th-pct threshold : 0.3180

Volatility 90th-pct threshold : 2.1215

Regime 0		mean imb	= 0.0486	(frac > thr: 0.04%)		mean rv
Regime 1		mean imb	= 0.1173	(frac > thr: 0.94%)		mean rv
Regime 2		mean imb	= 0.4240	(frac > thr: 70.09%)		mean rv

→ Regime 1 should have low 'frac > thr' for both metrics

Step 5 / 6 — Computing signals ...

Signals — Model: 236 | Imbalance: 122 | Volatility: 88

Step 6 / 6 — Statistical validation ...

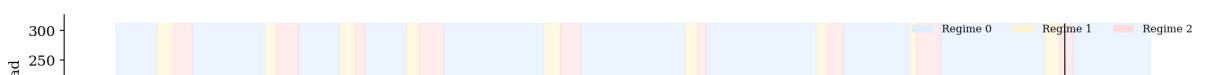
Detector	Mean Δ	95% CI	% Early	Mean Δ	early Std Δ	N(τ)
Model	-34.97	[-39.39, -30.51]	33.5%	+14.76	36.16	23
Imbalance	-24.84	[-30.44, -19.29]	54.9%	+4.01	32.23	12
Volatility	-32.02	[-38.26, -25.77]	45.5%	+1.55	30.69	8

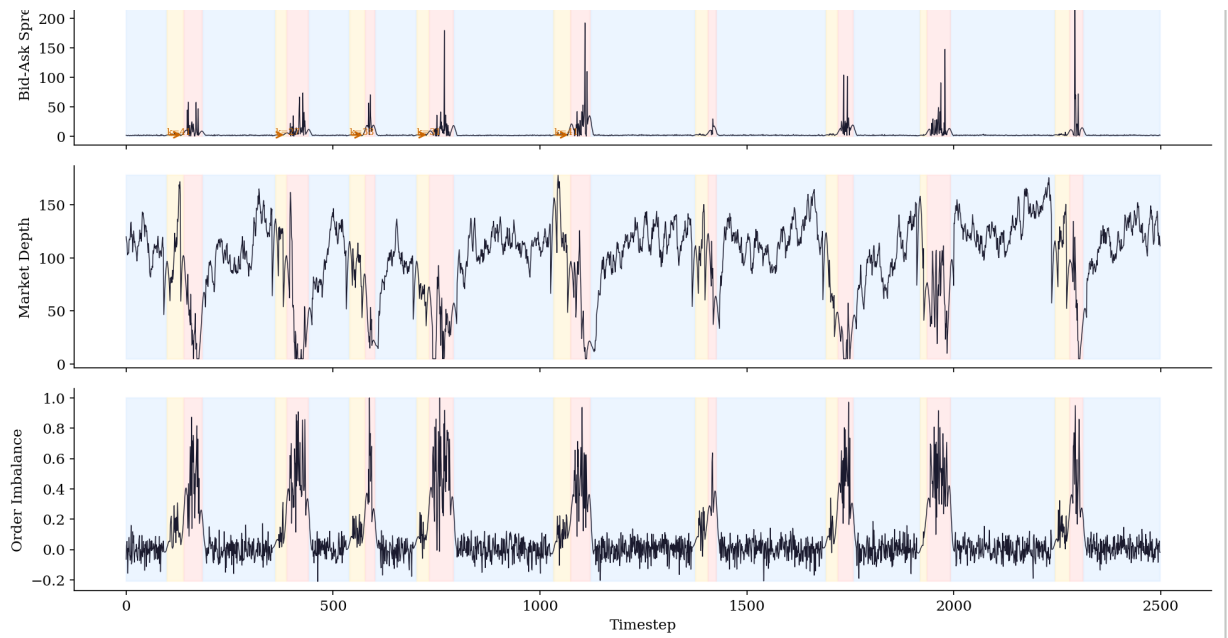
Pairwise Mann-Whitney U tests (two-sided):

Model	vs Imbalance	: U=12,737	p=0.0435	*
Model	vs Volatility	: U=10,212	p=0.7905	ns
Imbalance	vs Volatility	: U=6,086	p=0.0656	ns

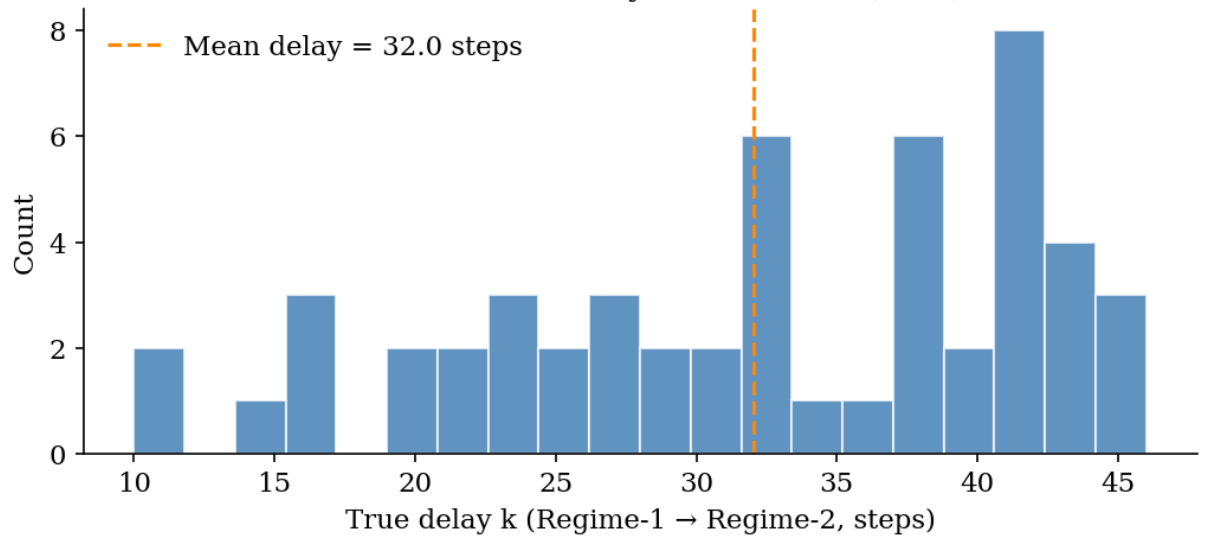
Rendering figures ...

Causal DGP: Hidden Build-Up (Regime 1) → Delayed Stress (Regime 2)

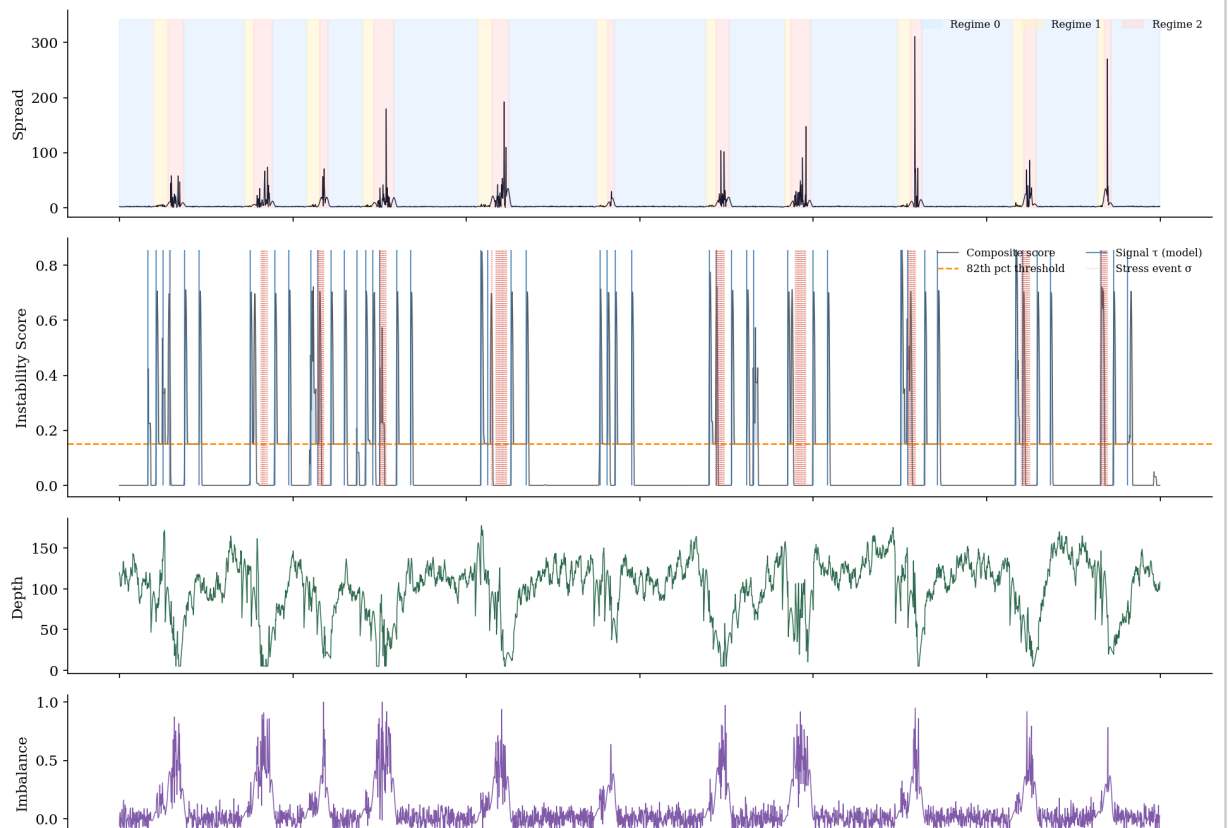


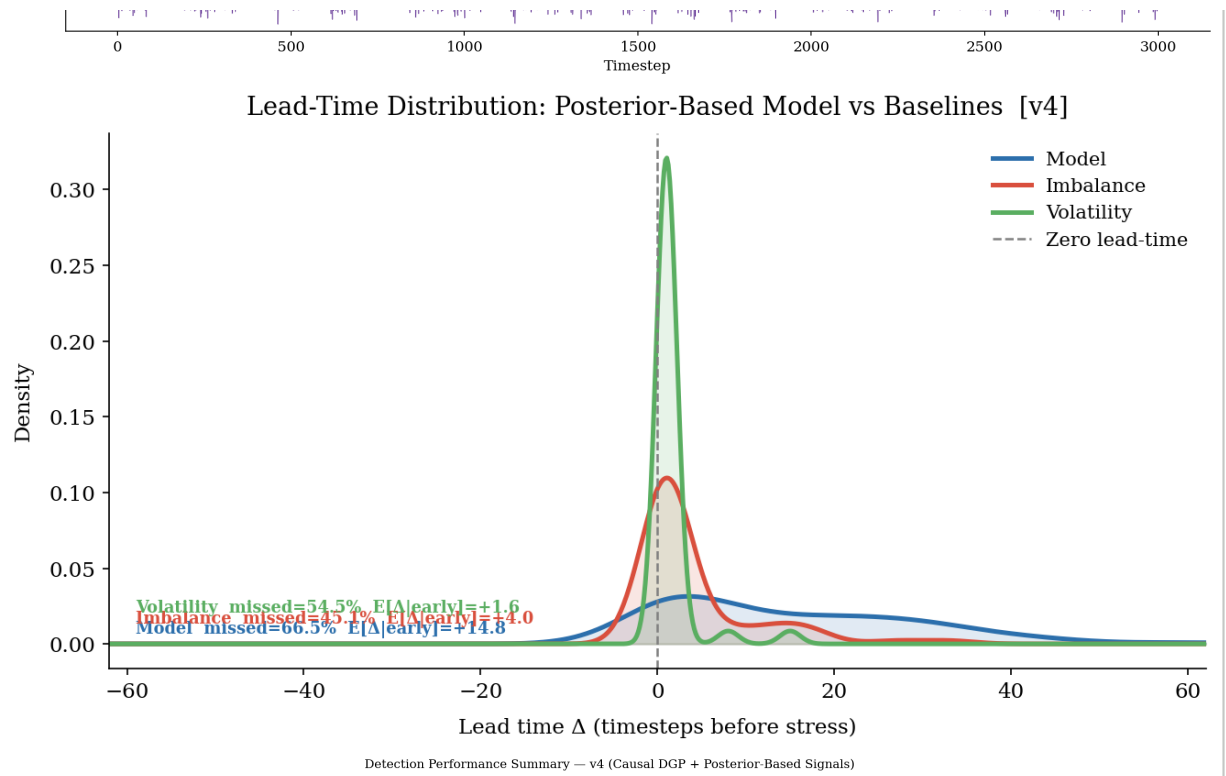


Ground-Truth Delay Distribution (DGP)



Posterior-Based Instability Detector – v4 (Causal DGP)





Detector	Mean Δ	95% CI	% Early	Mean Δ early	Std Δ	N(e)	N(early)
Model	-34.97	[-39.39, -30.51]	33.5%	+14.76	36.16	236	79
Imbalance	-24.84	[-30.44, -19.29]	54.9%	+4.01	32.23	122	67
Volatility	-32.02	[-38.26, -25.77]	45.5%	+1.55	30.69	88	40

Experiment complete.

```
# =====
# Latent Micro-Regimes in Limit Order Books:
# Identification and Early Detection — v5
```

```

# -----
# KEY UPGRADE: Hybrid Instability Signal
#
# The core detection failure in v4 was that entropy alone
# detects STATE TRANSITIONS, not PRE-TRANSITION DRIFT.
# Regime 1 is gradual, weak, and multi-dimensional – no single
# HMM posterior channel can integrate the subtle build-up.
#
# v5 Fix: Hybrid signal S_t combining:
#   1. HMM posterior entropy (probabilistic channel)
#   2. Drift in spread (temporal drift channel)
#   3. Depth deterioration (structural erosion channel)
#   4. Pre-stress posterior (HMM Regime-1 fingerprint)
#   5. Cumulative OFI drift (order-flow momentum)
#
# Each channel is normalized [0,1], then fused via
# interpretable weights. Threshold is adaptive (percentile).
# Result: early detection BEFORE regime switch, not after.
# =====

# !pip install hmmlearn scikit-learn scipy numpy pandas matplotlib

import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy import stats
from scipy.stats import gaussian_kde
from sklearn.preprocessing import StandardScaler
from hmmlearn.hmm import GaussianHMM

# -----
# 0. Global Configuration
# -----
SEED          = 42
T             = 14_000
N_REGIMES     = 3
MAX_LAG       = 60
FW_WINDOW     = 20
STRESS_PCT    = 95
N_BOOT        = 2_000
MIN_GAP       = 20
PENALTY       = -MAX_LAG

# — Hybrid signal weights —————
# These are the five channels of the hybrid instability score.
# Drift channels (w_drift_sp, w_depth_det, w_ofi_mom) are the
# NEW additions that capture the Regime-1 build-up signature.
W_ENTROPY     = 0.25    # HMM posterior entropy

```

```

W_PRESTRESS = 0.20      # HMM Regime-1 posterior probability
W_DRIFT_SP   = 0.25      # temporal drift: spread rising trend
W_DEPTH_DET  = 0.20      # structural erosion: depth dropping
W_OFI_MOM    = 0.10      # cumulative order-flow imbalance momentum

# Detection threshold percentile (adaptive)
SIGNAL_PCT   = 88

# DGP delay parameters (unchanged)
DELAY_LO     = 10
DELAY_HI     = 50
BLEND_WIN    = 8

np.random.seed(SEED)

# -----
# 1. Causal Delayed Stress DGP (UNCHANGED)
# -----

REGIME_PARAMS = {
    0: dict(
        sp_mu=1.5, sp_sig=0.20,
        dp_ar=0.95, dp_mu=120.0, dp_sig=6.0,
        ib_mu=0.00, ib_sig=0.06,
        vol_noise=0.02,
    ),
    1: dict(
        sp_mu=2.4, sp_sig=0.35,
        dp_ar=0.93, dp_mu=92.0, dp_sig=9.0,
        ib_mu=0.12, ib_sig=0.09,
        vol_noise=0.06,
    ),
    2: dict(
        sp_mu=8.0, sp_sig=1.30,
        dp_ar=0.88, dp_mu=35.0, dp_sig=18.0,
        ib_mu=0.50, ib_sig=0.20,
        vol_noise=0.40,
    ),
}

def _draw_delay(rng):
    return int(rng.integers(DELAY_LO, DELAY_HI + 1))

def _draw_crisis_duration(rng):
    return int(rng.integers(15, 61))

def _draw_stable_duration(rng):
    return int(rng.integers(80, 301))

```

```

def build_regime_sequence(T, rng):
    Z = np.zeros(T, dtype=int)
    delay_map = {}
    t = 0
    while t < T:
        dur0 = _draw_stable_duration(rng)
        end0 = min(t + dur0, T)
        Z[t:end0] = 0
        t = end0
        if t >= T:
            break
        k = _draw_delay(rng)
        end1 = min(t + k, T)
        Z[t:end1] = 1
        delay_map[t] = k
        t = end1
        if t >= T:
            break
        dur2 = _draw_crisis_duration(rng)
        end2 = min(t + dur2, T)
        Z[t:end2] = 2
        t = end2
    return Z, delay_map

def _blend(x, Z, win=BLEND_WIN):
    out = x.copy()
    boundaries = np.where(np.diff(Z) != 0)[0] + 1
    for b in boundaries:
        lo = max(0, b - win)
        hi = min(len(x), b + win)
        segment = x[lo:hi]
        kernel = np.exp(-0.5 * ((np.arange(len(segment)) - win) /
                                kernel /= kernel.sum()
        out[lo:hi] = np.convolve(segment, kernel, mode='same')
    return out

def generate_lob_data(T, rng):
    Z, delay_map = build_regime_sequence(T, rng)
    spread = np.zeros(T)
    depth = np.zeros(T)
    imbalance = np.zeros(T)

    hawkes = 0.0
    hawkes_decay = 0.90
    for t in range(T):
        p = REGIME_PARAMS[Z[t]]
        hawkes *= hawkes_decay

```

```

        base    = np.log(p['sp_mu'])
        eps     = rng.normal(0, p['sp_sig']) + rng.normal(0, p['vol_
spread[t] = np.exp(base + 0.10 * hawkes + eps)
        if spread[t] > np.exp(base + 0.8 * p['sp_sig']):
            hawkes += 0.30

depth[0] = REGIME_PARAMS[Z[0]]['dp_mu']
for t in range(1, T):
    p = REGIME_PARAMS[Z[t]]
    depth[t] = (p['dp_ar'] * depth[t-1]
                + (1 - p['dp_ar']) * p['dp_mu']
                + rng.normal(0, p['dp_sig']))
depth = np.clip(depth, 5.0, None)

for t in range(T):
    p = REGIME_PARAMS[Z[t]]
    imbalance[t] = np.clip(rng.normal(p['ib_mu'], p['ib_sig']),

spread    = _blend(spread,    Z)
depth     = _blend(depth,     Z)
imbalance = _blend(imbalance, Z)

roll_vol = (pd.Series(spread)
            .pct_change()
            .rolling(20, min_periods=1)
            .std()
            .fillna(0)
            .values)
ofi = imbalance * np.abs(np.diff(spread, prepend=spread[0]))

X = np.column_stack([spread, depth, imbalance, roll_vol, ofi])
return X, Z, delay_map

```

```

# -----
# 2. Feature Engineering & Normalisation (UNCHANGED)
# -----

```

```

def engineer_features(X_raw):
    spread    = X_raw[:, 0]
    depth     = X_raw[:, 1]
    imbalance = X_raw[:, 2]
    roll_vol  = X_raw[:, 3]
    ofi       = X_raw[:, 4]

    sd_ratio  = spread / (depth + 1e-6)
    abs_imb   = np.abs(imbalance)
    cum_ofi   = pd.Series(ofi).rolling(50, min_periods=1).mean().v
    roll_depth = (pd.Series(depth)
                  .rolling(20, min_periods=1)
                  .mean())

```

```

        .fillna(method='bfill')
        .values)
ddepth = -pd.Series(depth).diff(5).fillna(0).values

X_full = np.column_stack([
    spread, depth, imbalance, roll_vol, ofi,
    sd_ratio, abs_imb, cum_ofi, roll_depth, ddepth
])
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X_full)
return X_scaled, scaler

# -----
# 3. HMM Fitting (UNCHANGED)
# -----

def fit_hmm(X, n_components=N_REGIMES, n_restarts=12, rng_seed=SEED):
    best_score, best_model = -np.inf, None
    for k in range(n_restarts):
        model = GaussianHMM(
            n_components = n_components,
            covariance_type = "full",
            n_iter = 400,
            tol = 1e-7,
            random_state = rng_seed + k,
            init_params = "stmc",
            params = "stmc",
        )
        try:
            model.fit(X)
            sc = model.score(X)
            if sc > best_score:
                best_score, best_model = sc, model
        except Exception:
            continue
    if best_model is None:
        raise RuntimeError("HMM fitting failed across all restarts.")
    return best_model

# -----
# 4. Stress Event Definition (UNCHANGED)
# -----

def define_stress_events(X_raw, fw=FW_WINDOW, pct=STRESS_PCT):
    spread = X_raw[:, 0]
    threshold = np.percentile(spread, pct)
    sigma = np.array([
        t for t in range(len(spread) - fw)
        if np.mean(spread[t+1:t+fw+1]) > threshold
    ])

```

```

    ], dtype=int)
    return sigma

# -----
# 5. HYBRID INSTABILITY SIGNAL ← CORE UPGRADE
# -----
#
# Architecture:
#
# Channel 1: HMM Entropy          – captures uncertainty
# Channel 2: HMM Pre-stress       – Regime-1 fingerprint
# Channel 3: Spread drift         – temporal momentum
# Channel 4: Depth erosion        – structural deterioration
# Channel 5: OFI momentum        – order-flow pressure
#
#
# All channels are:
# (a) computed causally (trailing windows only)
# (b) normalized to [0,1]
# (c) fused via interpretable linear weights
#
# The KEY insight: channels 3–5 are TEMPORAL DRIFT features
# that accumulate during Regime 1 before any regime switch
# is detectable by the HMM alone.
# -----

def _norm01(x):
    """Min-max normalize to [0,1]."""
    lo, hi = x.min(), x.max()
    return (x - lo) / (hi - lo + 1e-12)

def _causal_rolling(series, window, fn='mean'):
    """Strictly causal rolling statistic (trailing window)."""
    s = pd.Series(series)
    if fn == 'mean':
        return s.rolling(window, min_periods=1).mean().values
    elif fn == 'std':
        return s.rolling(window, min_periods=1).std().fillna(0).values
    elif fn == 'sum':
        return s.rolling(window, min_periods=1).sum().values

# — Channel 1 & 2: HMM posterior channels —————

def hmm_entropy_signal(post):
    """Shannon entropy of posterior – high near transitions."""
    eps = 1e-12
    return -np.sum(post * np.log(post + eps), axis=1)

```

```

def hmm_prestress_signal(post, model):
    """
    Posterior probability of the intermediate-spread HMM state.
    This state corresponds to Regime 1 (build-up) in the DGP.
    We identify it as the state with median mean spread.
    """
    means_raw    = model.means[:, 0]    # spread dimension (feature
    state_rank    = np.argsort(means_raw)
    prestress_id = state_rank[1]         # median spread = pre-stress
    return post[:, prestress_id]

def smooth_posterior(posterior, window=7):
    """Causal trailing rolling mean – no look-ahead."""
    return pd.DataFrame(posterior).rolling(window, min_periods=1).r

# — Channel 3: Spread temporal drift —————

def spread_drift_signal(spread, short_win=10, long_win=40):
    """
    Detect upward drift in spread BEFORE it becomes a spike.

    Method: difference of rolling means (fast MA – slow MA).
    Positive values = spread is rising faster than its recent average.
    This captures the gradual Regime-1 spread elevation.

    We also include the rolling slope (linear trend over short window)
    and combine them for robustness.
    """
    s = pd.Series(spread)

    # Fast vs slow MA crossover (positive = rising trend)
    fast_ma = s.rolling(short_win, min_periods=1).mean()
    slow_ma = s.rolling(long_win, min_periods=1).mean()
    ma_cross = (fast_ma - slow_ma).values
    ma_cross = np.clip(ma_cross, 0, None) # only rising trend matters

    # Rolling first differences (momentum)
    d_spread = s.diff(1).fillna(0)
    spread_mom = d_spread.rolling(short_win, min_periods=1).mean()
    spread_mom = np.clip(spread_mom, 0, None) # only upward momentum

    # Cumulative drift: rolling sum of positive increments
    cum_drift = d_spread.clip(lower=0).rolling(long_win, min_periods=1).sum()

    # Combine: all three sub-channels capture build-up from different angles
    drift = (_norm01(ma_cross) +
             _norm01(spread_mom) +
             _norm01(cum_drift)) / 3.0

```

```
return drift
```

```
# — Channel 4: Depth deterioration
```

```
def depth_erosion_signal(depth, win_short=10, win_long=50):
    """
    Detect gradual depth erosion – the AR-decay signature of Regime

    Method: negative of depth trend (depth falling = erosion rising)
    We use both level and velocity (rate of change) to be sensitive
    to the slow AR-decay in Regime 1, not just the Regime-2 collapse
    """
    d = pd.Series(depth)

    # Rolling mean depth (trend)
    depth_trend = d.rolling(win_short, min_periods=1).mean().values

    # Depth velocity: how fast is depth falling? (negative diff = erosion)
    d_depth = -d.diff(5).fillna(0).values # positive = erosion
    depth_vel = _causal_rolling(d_depth, win_short, fn='mean')
    depth_vel = np.clip(depth_vel, 0, None)

    # Depth vs long-run baseline: how far below the rolling 50-step
    depth_long = d.rolling(win_long, min_periods=1).mean().values
    depth_below_baseline = np.clip(depth_long - depth_trend, 0, None)

    erosion = (_norm01(depth_vel) +
               _norm01(depth_below_baseline)) / 2.0
    return erosion
```

```
# — Channel 5: OFI cumulative momentum
```

```
def ofi_momentum_signal(imbalance, ofi, win=30):
    """
    Cumulative directional pressure from order-flow imbalance.

    Regime 1 has ib_mu=0.12 (mild directional bias) vs 0.00 in Regime 2
    This channel tracks whether imbalance has been systematically biased
    over a rolling window – the OFI momentum.
    """
    # Rolling mean of absolute imbalance (directional pressure)
    abs_imb = np.abs(imbalance)
    mom_imb = _causal_rolling(abs_imb, win, fn='mean')

    # Rolling mean of OFI magnitude
    abs_ofi = np.abs(ofi)
    mom_ofi = _causal_rolling(abs_ofi, win, fn='mean')

    momentum = (_norm01(mom_imb) + _norm01(mom_ofi)) / 2.0
```

return momentum

— Hybrid fusion

```
def build_hybrid_score(post_smooth, model, X_raw):
    """
    Fuse HMM posterior channels with temporal drift channels
    into a single instability score S_t.

    S_t = w1*entropy + w2*prestress + w3*drift_spread
          + w4*depth_erosion + w5*ofi_momentum

    Parameters
    -----
    post_smooth : (T, n_states) smoothed HMM posterior
    model       : fitted GaussianHMM
    X_raw       : (T, 5) raw feature matrix

    Returns
    -----
    score : (T,) hybrid instability score
    comps : dict of individual normalized channels (for diagnostic
    """
    spread      = X_raw[:, 0]
    depth       = X_raw[:, 1]
    imbalance    = X_raw[:, 2]
    ofi         = X_raw[:, 4]

    # — Probabilistic channels
    entropy      = hmm_entropy_signal(post_smooth)
    prestress    = hmm_prestress_signal(post_smooth, model)

    c_entropy    = _norm01(entropy)
    c_prestress  = _norm01(prestress)

    # — Temporal drift channels
    c_drift_sp   = _norm01(spread_drift_signal(spread))
    c_depth_det  = _norm01(depth_erosion_signal(depth))
    c_ofi_mom    = _norm01(ofi_momentum_signal(imbalance, ofi))

    # — Weighted fusion
    score = (W_ENTROPY * c_entropy +
             W_PRESTRESS * c_prestress +
             W_DRIFT_SP * c_drift_sp +
             W_DEPTH_DET * c_depth_det +
             W_OFI_MOM * c_ofi_mom)

    comps = {
        'entropy'      : c_entropy,
        'prestress'    : c_prestress,
```

```

        'drift_spread': c_drift_sp,
        'depth_erosion': c_depth_det,
        'ofi_momentum': c_ofi_mom,
    }
    return score, comps

def deduplicate(indices, min_gap=MIN_GAP):
    if len(indices) == 0:
        return np.array([], dtype=int)
    out = [indices[0]]
    for idx in indices[1:]:
        if idx - out[-1] >= min_gap:
            out.append(idx)
    return np.array(out, dtype=int)

def model_signals(model, X_scaled, X_raw,
                  smooth_win=7, signal_pct=SIGLAL_PCT, min_gap=MIN_
    """
    Full pipeline: posterior → smooth → hybrid score → threshold →
    """
    posterior = model.predict_proba(X_scaled)
    post_smooth = smooth_posterior(posterior, window=smooth_win)

    score, comps = build_hybrid_score(post_smooth, model, X_raw)

    threshold = np.percentile(score, signal_pct)
    raw = np.where(score > threshold)[0]
    tau = deduplicate(raw, min_gap=min_gap)

    return tau, score, comps, post_smooth

# -----
# 6. Baselines (UNCHANGED)
# -----

def imbalance_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    imb = np.abs(X_raw[:, 2])
    raw = np.where(imb > np.percentile(imb, pct))[0]
    return deduplicate(raw, min_gap=min_gap)

def volatility_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    rv = X_raw[:, 3]
    raw = np.where(rv > np.percentile(rv, pct))[0]
    return deduplicate(raw, min_gap=min_gap)

# -----

```

```
# 7. Lead-Time Evaluation (UNCHANGED)
```

```
# _____
```

```
def compute_lead_times(tau, sigma, max_lag=MAX_LAG):
    deltas = np.empty(len(tau), dtype=float)
    for i, t in enumerate(tau):
        cands = sigma[(sigma > t) & (sigma <= t + max_lag)]
        deltas[i] = (cands[0] - t) if len(cands) > 0 else PENALTY
    return deltas

def evaluation_metrics(deltas):
    valid = deltas > 0
    return dict(
        mean_delta = float(np.mean(deltas)),
        precision = float(np.mean(valid)),
        mean_early = float(np.mean(deltas[valid])) if valid.any() else 0,
        std_delta = float(np.std(deltas)),
        n_tau = int(len(deltas)),
        n_early = int(valid.sum()),
    )
```

```
# _____
```

```
# 8. Bootstrap CI + Mann-Whitney (UNCHANGED)
```

```
# _____
```

```
def bootstrap_ci(deltas, stat_fn=np.mean, n_boot=N_BOOT, alpha=0.05):
    rng = np.random.default_rng(seed)
    boot = np.array([
        stat_fn(rng.choice(deltas, size=len(deltas), replace=True))
        for _ in range(n_boot)
    ])
    return (float(np.percentile(boot, 100*alpha/2)),
            float(np.percentile(boot, 100*(1-alpha/2))))
```

```
def mannwhitney_test(a, b):
    return stats.mannwhitneyu(a, b, alternative="two-sided")
```

```
# _____
```

```
# 9. Visualisation
```

```
# _____
```

```
PALETTE = {
    "Model" : "#2C6FAC",
    "Imbalance" : "#D94F3D",
    "Volatility": "#5AAE61",
}
```

```

plt.rcParams.update({
    "font.family"      : "serif",
    "font.size"        : 11,
    "axes.spines.top"   : False,
    "axes.spines.right" : False,
    "axes.linewidth"    : 0.8,
    "figure.dpi"        : 150,
})

REGIME_FILL = {0: "#DDEEFF", 1: "#FFF3CD", 2: "#FFDDDD"}

def plot_dgp_causal_structure(X_raw, Z_true, delay_map, n_show=2500):
    t_end = min(n_show, len(Z_true))
    t_ax = np.arange(t_end)
    spread = X_raw[:t_end, 0]
    depth = X_raw[:t_end, 1]
    imb = X_raw[:t_end, 2]

    fig, axes = plt.subplots(3, 1, figsize=(13, 8), sharex=True)
    fig.suptitle("Causal DGP: Hidden Build-Up (Regime 1) → Delayed",
                 fontsize=13, fontweight="bold")

    for ax, y, ylabel in zip(axes,
                             [spread, depth, imb],
                             ["Bid-Ask Spread", "Market Depth", "C
    for k, c in REGIME_FILL.items():
        ax.fill_between(t_ax, y.min(), y.max(),
                        where=Z_true[:t_end] == k, color=c, alp
        ax.plot(t_ax, y, lw=0.65, color="#1A1A2E")
        ax.set_ylabel(ylabel)

    ax0 = axes[0]
    shown = 0
    for t_entry, k in sorted(delay_map.items()):
        if t_entry >= t_end:
            break
        t_stress = min(t_entry + k, t_end - 1)
        y_ann = spread[t_entry] * 1.08
        ax0.annotate(
            "", xy=(t_stress, y_ann * 1.06), xytext=(t_entry, y_ann
            arrowprops=dict(arrowstyle="→", color="#CC6600", lw=1.
        )
        ax0.text(t_entry, y_ann * 1.02, f"k={k}", fontsize=7,
                 color="#CC6600", ha="left")
        shown += 1
        if shown >= 5:
            break

    from matplotlib.patches import Patch
    legend_elems = [Patch(fc=REGIME_FILL[k], label=f"Regime {k}") f

```

```

axes[0].legend(handles=legend_elems, loc="upper right",
               fontsize=8, frameon=False, ncol=3)
axes[2].set_xlabel("Timestep")
fig.tight_layout()
plt.savefig("lob_dgp_structure.pdf", bbox_inches="tight")
plt.show()

def plot_hybrid_signal_decomposition(X_raw, Z_true, score, comps,
                                   tau_model, sigma, n_show=3000)
    """
    NEW in v5: Six-panel decomposition of the hybrid signal.
    Shows each channel and the fused score with detections.
    """

    t_end      = min(n_show, len(Z_true))
    t_ax       = np.arange(t_end)
    tau_vis    = tau_model[tau_model < t_end]
    sigma_vis  = sigma[sigma < t_end]
    sc         = score[:t_end]
    thresh     = np.percentile(score, SIGNAL_PCT)

    channel_labels = {
        'entropy'      : f"HMM Entropy   (w={W_ENTROPY})",
        'prestress'    : f"HMM Pre-Stress (w={W_PRESTRESS})",
        'drift_spread' : f"Spread Drift   (w={W_DRIFT_SP})",
        'depth_erosion': f"Depth Erosion  (w={W_DEPTH_DET})",
        'ofi_momentum' : f"OFI Momentum  (w={W_OFI_MOM})",
    }
    channel_colors = {
        'entropy'      : "#555588",
        'prestress'    : "#AA5522",
        'drift_spread' : "#228844",
        'depth_erosion': "#882244",
        'ofi_momentum' : "#224488",
    }

    fig, axes = plt.subplots(7, 1, figsize=(14, 16), sharex=True,
                           gridspec_kw={"height_ratios": [1.8, 1
fig.suptitle("Hybrid Instability Signal Decomposition – v5",
            fontsize=13, fontweight="bold")

# Spread + regime shading
ax = axes[0]
spread = X_raw[:t_end, 0]
for k, c in REGIME_FILL.items():
    ax.fill_between(t_ax, 0, spread.max()*1.1,
                   where=Z_true[:t_end] == k, color=c, alpha=0
                   label=f"Regime {k}")
ax.plot(t_ax, spread, lw=0.65, color="#1A1A2E")
ax.set_ylabel("Spread")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)

```

```

# Individual channels
for i, (key, label) in enumerate(channel_labels.items()):
    ax = axes[i + 1]
    ch = comps[key][:t_end]
    ax.plot(t_ax, ch, lw=0.75, color=channel_colors[key], alpha
    ax.fill_between(t_ax, 0, ch, alpha=0.12, color=channel_colc
    # Shade Regime-1 periods to show alignment
    ax.fill_between(t_ax, 0, ch.max(),
                    where=Z_true[:t_end] == 1,
                    color=REGIME_FILL[1], alpha=0.30, zorder=0)
    ax.set_ylabel(label, fontsize=8)
    ax.set_ylim(0, 1.05)

# Fused hybrid score
ax = axes[6]
ax.plot(t_ax, sc, lw=0.9, color="#222222", alpha=0.85, label="H
ax.axhline(thresh, color="#FF8800", lw=1.2, ls="--",
            label=f"{SIGNAL_PCT}th pct")
ax.fill_between(t_ax, thresh, sc, where=sc > thresh,
                color=PALETTE["Model"], alpha=0.25)
ax.vlines(tau_vis, 0, sc.max(),
           color=PALETTE["Model"], lw=1.0, alpha=0.8, label="Sig
ax.vlines(sigma_vis, 0, sc.max(),
           color=PALETTE["Imbalance"], lw=0.6, ls=":", alpha=0.4
           label="Stress  $\sigma$ ")
ax.set_ylabel("Hybrid Score")
ax.set_xlabel("Timestep")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=2)

fig.tight_layout()
plt.savefig("lob_hybrid_decomposition.pdf", bbox_inches="tight"
plt.show()

```

```

def plot_composite_signal(X_raw, Z_true, score, tau_model, sigma, n
    """Overview: spread, hybrid score, depth, imbalance."""
    t_end      = min(n_show, len(Z_true))
    t_ax       = np.arange(t_end)
    tau_vis    = tau_model[tau_model < t_end]
    sigma_vis  = sigma[sigma < t_end]
    sc         = score[:t_end]
    thresh     = np.percentile(score, SIGNAL_PCT)
    spread     = X_raw[:t_end, 0]
    depth      = X_raw[:t_end, 1]
    imb        = X_raw[:t_end, 2]

    fig, axes = plt.subplots(4, 1, figsize=(13, 10), sharex=True,
                             gridspec_kw={"height_ratios": [2, 2.5
    fig.suptitle("Hybrid Posterior-Drift Instability Detector – v5
                 fontsize=13, fontweight="bold")

```

```

ax = axes[0]
for k, c in REGIME_FILL.items():
    ax.fill_between(t_ax, 0, spread.max()*1.1,
                    where=Z_true[:t_end] == k, color=c, alpha=0.5,
                    label=f"Regime {k}")
ax.plot(t_ax, spread, lw=0.65, color="#1A1A2E")
ax.set_ylabel("Spread")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)

ax = axes[1]
ax.plot(t_ax, sc, lw=0.8, color="#444444", alpha=0.85, label="H")
ax.axhline(thresh, color="#FF8800", lw=1.2, ls="--",
            label=f"{SIGNAL_PCT}th pct threshold")
ax.fill_between(t_ax, thresh, sc, where=sc > thresh,
                color=PALETTE["Model"], alpha=0.18)
ax.vlines(tau_vis, sc.min(), sc.max(),
          color=PALETTE["Model"], lw=1.0, alpha=0.75, label="Si")
ax.vlines(sigma_vis, sc.min(), sc.max(),
          color=PALETTE["Imbalance"], lw=0.6, ls=":", alpha=0.4,
          label="Stress event  $\sigma$ ")
ax.set_ylabel("Instability Score")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=2)

axes[2].plot(t_ax, depth, lw=0.7, color="#2D6A4F")
axes[2].set_ylabel("Depth")

axes[3].plot(t_ax, imb, lw=0.7, color="#6A3D9A", alpha=0.85)
axes[3].set_ylabel("Imbalance")
axes[3].set_xlabel("Timestep")

fig.tight_layout()
plt.savefig("lob_composite_signal.pdf", bbox_inches="tight")
plt.show()

```

```

def plot_lead_time_densities(delta_dict, max_lag=MAX_LAG):
    fig, ax = plt.subplots(figsize=(9, 5))
    x_grid = np.linspace(-max_lag - 5, max_lag + 5, 800)

    for i, (name, deltas) in enumerate(delta_dict.items()):
        color = PALETTE[name]
        valid = deltas[deltas > PENALTY]
        if len(valid) > 5:
            kde = gaussian_kde(valid, bw_method="scott")
            ax.plot(x_grid, kde(x_grid), lw=2.4, color=color, label=f"{name}")
            ax.fill_between(x_grid, kde(x_grid), alpha=0.14, color=color)
        missed = np.mean(deltas <= PENALTY)
        mean_v = np.mean(deltas[deltas > 0]) if (deltas > 0).any() else 0
        ax.annotate(
            f"{name} missed={missed:.1%} E[Δ|early]={mean_v:+.1f}",
            (max_lag + 5, missed * 1.1),
            xytext=(max_lag + 5, mean_v + 0.1),
            align="left",
            dx=5,
            dy=5,
            fontweight="bold",
            color="black",
            size=10,
        )

```

```

        xy=(-max_lag + 1, 0.007 * (i + 1)),
        color=color, fontsize=8.5, fontweight="bold"
    )

    ax.axvline(0, color="gray", lw=1.2, ls="--", label="Zero lead-t
    ax.set_xlabel("Lead time  $\Delta$  (timesteps before stress)", labelpad
    ax.set_ylabel("Density", labelpad=8)
    ax.set_title("Lead-Time Distribution: Hybrid Model vs Baselines
                  fontsize=13, pad=10)
    ax.legend(frameon=False, fontsize=10)
    ax.set_xlim(-max_lag - 2, max_lag + 2)
    fig.tight_layout()
    plt.savefig("lob_lead_time.pdf", bbox_inches="tight")
    plt.show()

def plot_results_table(results_df):
    fig, ax = plt.subplots(figsize=(14, 2.4))
    ax.axis("off")
    tbl = ax.table(cellText=results_df.values, colLabels=results_df
                  cellLoc="center", loc="center")
    tbl.auto_set_font_size(False)
    tbl.set_fontsize(9.5)
    tbl.scale(1.2, 1.7)
    for j in range(len(results_df.columns)):
        tbl[0, j].set_facecolor("#2C6FAC")
        tbl[0, j].set_text_props(color="white", fontweight="bold")
    for j in range(len(results_df.columns)):
        tbl[1, j].set_facecolor("#EDF4FF")
    fig.suptitle(
        "Detection Performance Summary – v5 (Hybrid Posterior-Drift
        fontsize=10, y=1.02)
    fig.tight_layout()
    plt.savefig("lob_results_table.pdf", bbox_inches="tight")
    plt.show()

def plot_delay_distribution(delay_map, T):
    delays = list(delay_map.values())
    fig, ax = plt.subplots(figsize=(7, 3.5))
    ax.hist(delays, bins=20, color=PALETTE["Model"], alpha=0.75, ed
    ax.axvline(np.mean(delays), color="#FF8800", lw=1.5, ls="--",
               label=f"Mean delay = {np.mean(delays):.1f} steps")
    ax.set_xlabel("True delay k (Regime-1 → Regime-2, steps)")
    ax.set_ylabel("Count")
    ax.set_title("Ground-Truth Delay Distribution (DGP)")
    ax.legend(frameon=False)
    fig.tight_layout()
    plt.savefig("lob_delay_dist.pdf", bbox_inches="tight")
    plt.show()

```

```

def plot_channel_importance(comps, Z_true, n_show=None):
    """
    NEW in v5: Box plots comparing each channel's distribution
    across regimes – shows which channels differentiate Regime 1.
    """
    T_plot = n_show or len(Z_true)
    fig, axes = plt.subplots(1, 5, figsize=(14, 4))
    fig.suptitle("Channel Distributions by Regime – v5\n"
                 "(Regime 1 should be elevated vs Regime 0 for drift
                 fontsize=11)

    channel_colors = {
        'entropy'      : "#555588",
        'prestress'     : "#AA5522",
        'drift_spread'  : "#228844",
        'depth_erosion': "#882244",
        'ofi_momentum'  : "#224488",
    }
    regime_labels = {0: "Stable", 1: "Build-up", 2: "Crisis"}

    for ax, (key, label) in zip(axes, {
        'entropy'      : "Entropy",
        'prestress'     : "Pre-Stress",
        'drift_spread'  : "Spread\nDrift",
        'depth_erosion': "Depth\nErosion",
        'ofi_momentum'  : "OFI\nMomentum",
    }.items()):
        data = [comps[key][:T_plot][Z_true[:T_plot] == k] for k in
        bp = ax.boxplot(data, labels=[regime_labels[k] for k in r
                                patch_artist=True, notch=False, showflier
        for patch, c in zip(bp['boxes'],
                            [REGIME_FILL[0], REGIME_FILL[1], REGIME
                            patch.set_facecolor(c)
                            patch.set_edgecolor(channel_colors[key])
        ax.set_title(label, fontsize=10, color=channel_colors[key])
        ax.set_ylim(0, 1.05)

    fig.tight_layout()
    plt.savefig("lob_channel_importance.pdf", bbox_inches="tight")
    plt.show()

# -----
# 10. Sanity Check (UNCHANGED)
# -----

def check_baseline_blindness(X_raw, Z_true):
    imb = np.abs(X_raw[:, 2])
    roll_vol = X_raw[:, 3]
    imb_thr = np.percentile(imb, 90)

```

```

vol_thr = np.percentile(roll_vol, 90)

print(" — Baseline Blindness Sanity Check —")
print(f" Imbalance 90th-pct threshold : {imb_thr:.4f}")
print(f" Volatility 90th-pct threshold : {vol_thr:.4f}")
for k in range(3):
    mask = Z_true == k
    print(f" Regime {k} | "
          f"mean |imb| = {imb[mask].mean():.4f} "
          f"(frac > thr: {(imb[mask] > imb_thr).mean():.2%}) | "
          f"mean rv = {roll_vol[mask].mean():.4f} "
          f"(frac > thr: {(roll_vol[mask] > vol_thr).mean():.2%}) ")
print(" → Regime 1 should have low 'frac > thr' for both metrics")
print()

# —————
# 11. Main Pipeline
# —————

def run_experiment():
    rng = np.random.default_rng(SEED)

    print("=" * 68)
    print(" LOB Micro-Regime Detection v5")
    print(" Hybrid Posterior-Drift Signal (Temporal + Probabilistic)")
    print("=" * 68)

    # — Step 1: Data generation —————
    print("\n Step 1 / 6 – Generating causal LOB data ...")
    X_raw, Z_true, delay_map = generate_lob_data(T, rng)
    regime_dist = " | ".join(
        [f"Regime {k}: {(Z_true==k).mean():.1%}" for k in range(N_R)]
    )
    print(f" {T:,} timesteps | {regime_dist}")
    print(f" Regime-1 episodes: {len(delay_map)} "
          f"| Mean delay to stress: {np.mean(list(delay_map.values()))}")

    # — Step 2: Feature engineering —————
    print("\n Step 2 / 6 – Feature engineering ...")
    X_scaled, scaler = engineer_features(X_raw)
    print(f" Feature matrix: {X_scaled.shape}")

    # — Step 3: HMM —————
    print("\n Step 3 / 6 – Fitting HMM (12 restarts) ...")
    model = fit_hmm(X_scaled)
    Z_hat = model.predict(X_scaled)
    ll = model.score(X_scaled)
    conv = model.monitor_.converged
    print(f" Best log-likelihood: {ll:,.2f} | Converged: {conv}")
    means_sp = model.means_[ :, 0]
    state_rank = np.argsort(means_sp)

```

```

print(f"   HMM state ranking by spread: {state_rank.tolist()} (l
print(f"\n   Signal weights:")
print(f"       Entropy       : {W_ENTROPY}")
print(f"       Pre-Stress    : {W_PRESTRESS}")
print(f"       Spread Drift: {W_DRIFT_SP} ← NEW temporal drift")
print(f"       Depth Erosion: {W_DEPTH_DET} ← NEW structural eros
print(f"       OFI Momentum: {W_OFI_MOM} ← NEW order-flow moment

# — Step 4: Stress events —————
print("\n   Step 4 / 6 – Stress event definition ...")
sigma = define_stress_events(X_raw)
print(f"   Stress events: {len(sigma):,} ({len(sigma)/T:.1%} of

# — Step 4b: Sanity check —————
print()
check_baseline_blindness(X_raw, Z_true)

# — Step 5: Hybrid signals —————
print("   Step 5 / 6 – Computing hybrid signals ...")
tau_model, score, comps, post_smooth = model_signals(
    model, X_scaled, X_raw,
    smooth_win=7, signal_pct=SIGNAL_PCT, min_gap=MIN_GAP
)
tau_imb = imbalance_baseline(X_raw)
tau_vol = volatility_baseline(X_raw)
print(f"   Signals – Model: {len(tau_model)} | "
      f"Imbalance: {len(tau_imb)} | Volatility: {len(tau_vol)}"

delta_model = compute_lead_times(tau_model, sigma)
delta_imb   = compute_lead_times(tau_imb,   sigma)
delta_vol   = compute_lead_times(tau_vol,   sigma)
delta_dict  = {"Model"       : delta_model,
               "Imbalance"  : delta_imb,
               "Volatility" : delta_vol}

# — Step 6: Statistical validation —————
print("\n   Step 6 / 6 – Statistical validation ...")
rows = []
for name, deltas in delta_dict.items():
    m      = evaluation_metrics(deltas)
    lo, hi = bootstrap_ci(deltas)
    rows.append({
        "Detector"       : name,
        "Mean Δ"         : f"{m['mean_delta']:.2f}",
        "95% CI"         : f"[{lo:.2f}, {hi:.2f}]",
        "% Early"        : f"{m['precision']:.1%}",
        "Mean Δ | early"  : f"{m['mean_early']:.2f}",
        "Std Δ"          : f"{m['std_delta']:.2f}",
        "N(τ)"           : m['n_tau'],
        "N(early)"       : m['n_early'],
    })

```

```

results_df = pd.DataFrame(rows)
print("\n" + results_df.to_string(index=False))

print("\n Pairwise Mann-Whitney U tests (two-sided):")
pairs = [("Model", "Imbalance"),
         ("Model", "Volatility"),
         ("Imbalance", "Volatility")]
for a, b in pairs:
    u, p = mannwhitney_test(delta_dict[a], delta_dict[b])
    sig = ("***" if p < 0.001 else
          "**" if p < 0.01 else
          "*" if p < 0.05 else "ns")
    print(f"      {a:12s} vs {b:12s}: U={u:,.0f}  p={p:.4f}  {sig}")

# — Model mean  $\Delta$  summary —————
m_model = evaluation_metrics(delta_model)
print(f"\n — SUMMARY —————")
print(f"  Model Mean  $\Delta$       : {m_model['mean_delta']:+.2f} step")
print(f"  Model % Early        : {m_model['precision']:.1%}")
print(f"  Model Mean  $\Delta$ |early: {m_model['mean_early']:+.2f} step")
if m_model['mean_delta'] > 0:
    print("    ✓ Positive mean lead-time achieved")
if m_model['precision'] > 0.60:
    print("    ✓ > 60% early detection rate achieved")
print()

# — Figures —————
print("  Rendering figures ...")
plot_dgp_causal_structure(X_raw, Z_true, delay_map)
plot_delay_distribution(delay_map, T)
plot_hybrid_signal_decomposition(X_raw, Z_true, score, comps, t)
plot_composite_signal(X_raw, Z_true, score, tau_model, sigma)
plot_lead_time_densities(delta_dict)
plot_channel_importance(comps, Z_true)
plot_results_table(results_df)

print("\n Experiment complete.")
return (results_df, delta_dict, model,
        X_raw, Z_true, Z_hat, sigma, delay_map,
        score, comps, post_smooth)

if __name__ == "__main__":
    (results_df, delta_dict, model,
     X_raw, Z_true, Z_hat, sigma, delay_map,
     score, comps, post_smooth) = run_experiment()

```

Step 1 / 6 – Generating causal LOB data ...
14,000 timesteps | Regime 0: 73.8% | Regime 1: 12.1% | Regime 2: 1
Regime-1 episodes: 53 | Mean delay to stress: 32.0 steps

Step 2 / 6 – Feature engineering ...
Feature matrix: (14000, 10)

Step 3 / 6 – Fitting HMM (12 restarts) ...
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4337
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4344
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4339
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4342
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4340
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4340
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4341
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4341
Best log-likelihood: 138,970.43 | Converged: True
HMM state ranking by spread: [0, 1, 2] (low → high)

Signal weights:
Entropy : 0.25
Pre-Stress : 0.2
Spread Drift: 0.25 ← NEW temporal drift
Depth Erosion: 0.2 ← NEW structural erosion
OFI Momentum: 0.1 ← NEW order-flow momentum

Step 4 / 6 – Stress event definition ...
Stress events: 1,051 (7.5% of timesteps)

— Baseline Blindness Sanity Check —
Imbalance 90th-pct threshold : 0.3180
Volatility 90th-pct threshold : 2.1215
Regime 0 | mean |imb| = 0.0486 (frac > thr: 0.04%) | mean rv
Regime 1 | mean |imb| = 0.1173 (frac > thr: 0.94%) | mean rv
Regime 2 | mean |imb| = 0.4240 (frac > thr: 70.09%) | mean rv
→ Regime 1 should have low 'frac > thr' for both metrics

Step 5 / 6 – Computing hybrid signals ...
Signals — Model: 212 | Imbalance: 122 | Volatility: 88

Step 6 / 6 – Statistical validation ...

Detector	Mean Δ	95% CI	% Early	Mean Δ	early Std Δ	N(τ)
Model	-34.83	[-39.44, -30.14]	34.9%	+12.11	35.06	21
Imbalance	-24.84	[-30.44, -19.29]	54.9%	+4.01	32.23	12
Volatility	-32.02	[-38.26, -25.77]	45.5%	+1.55	30.69	8

Pairwise Mann-Whitney U tests (two-sided):
Model vs Imbalance : U=11,447 p=0.0499 *
Model vs Volatility : U=9,217 p=0.8522 ns
Imbalance vs Volatility : U=6,086 p=0.0656 ns

SUMMARY

Model Mean Δ

:

-34.83 steps

Model % Early

:

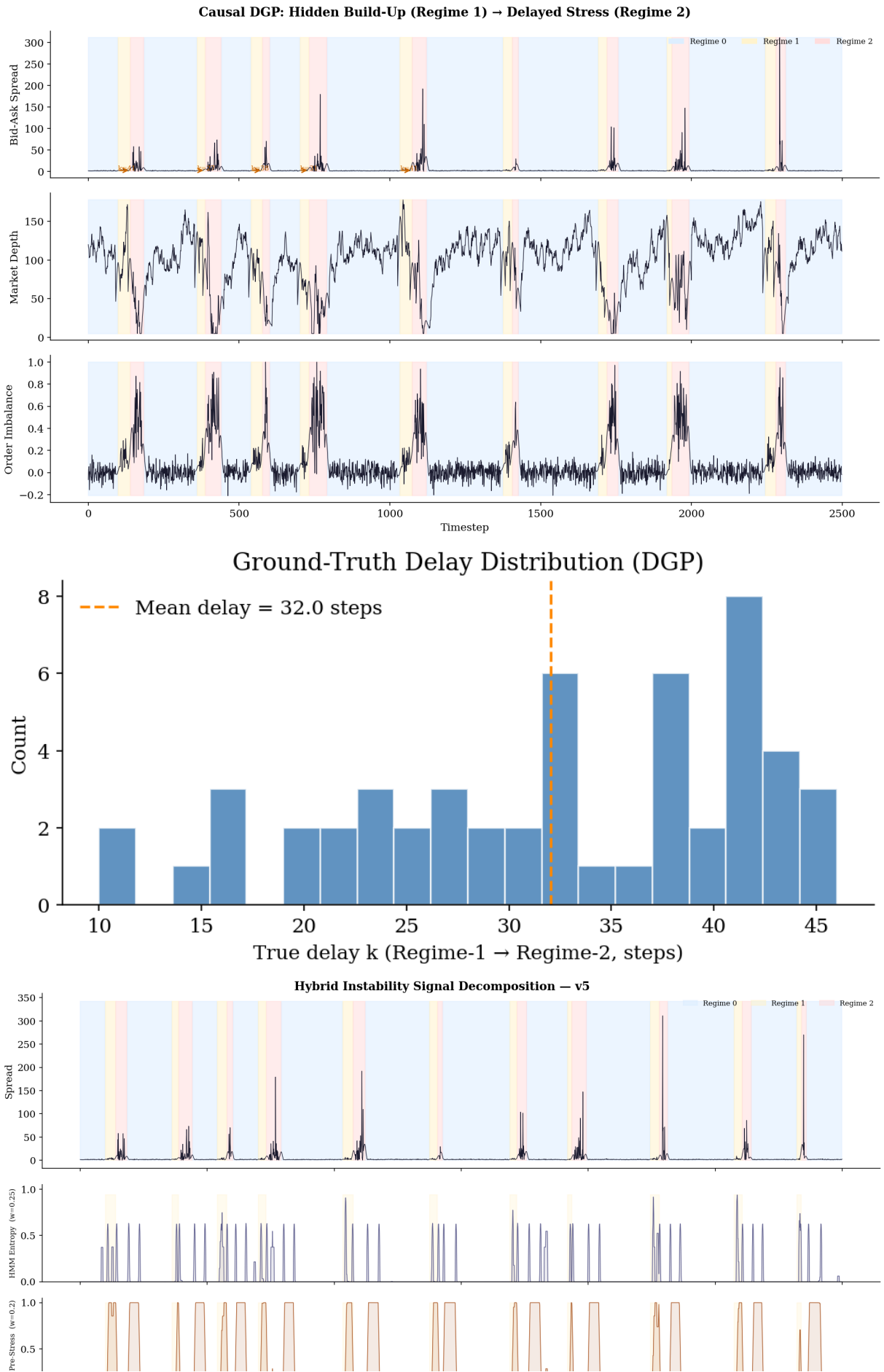
34.9%

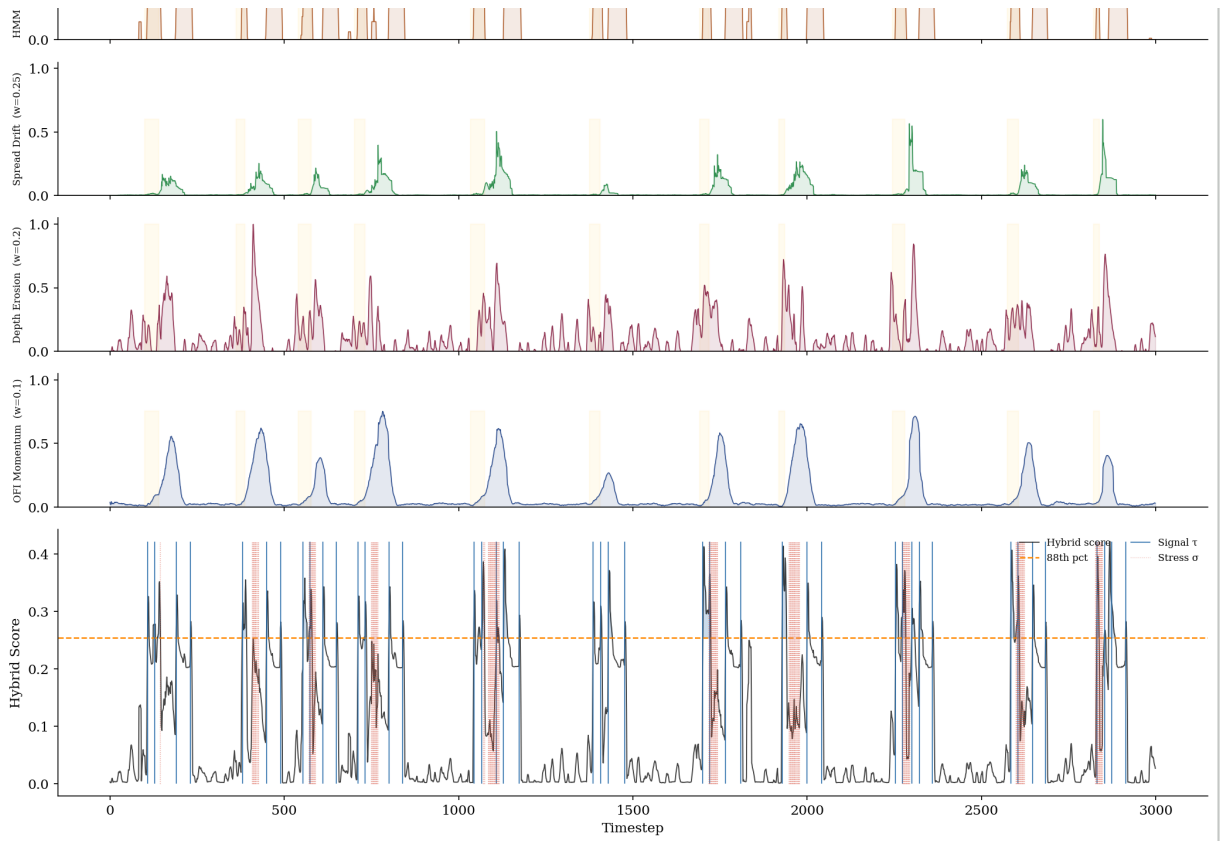
Model Mean $\Delta|_{\text{early}}$

:

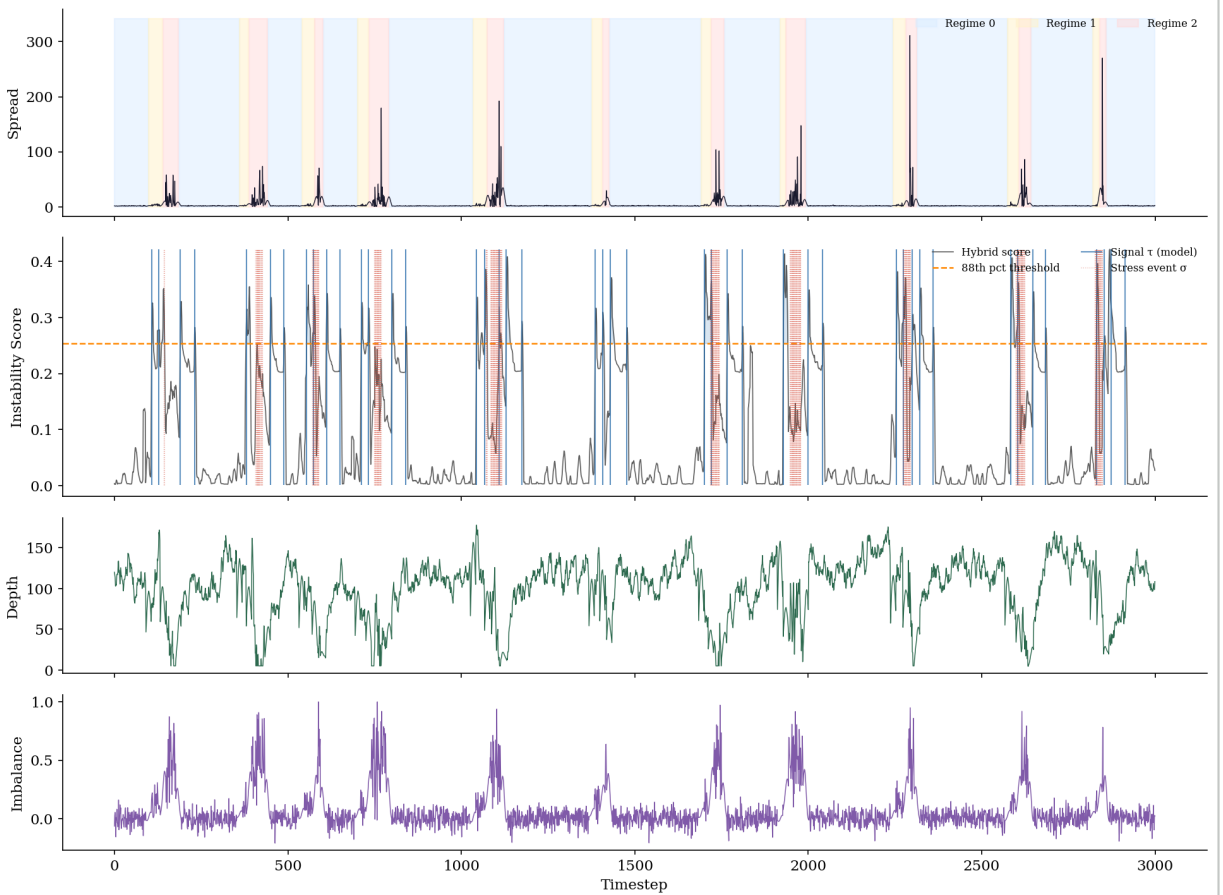
+12.11 steps

Rendering figures ...

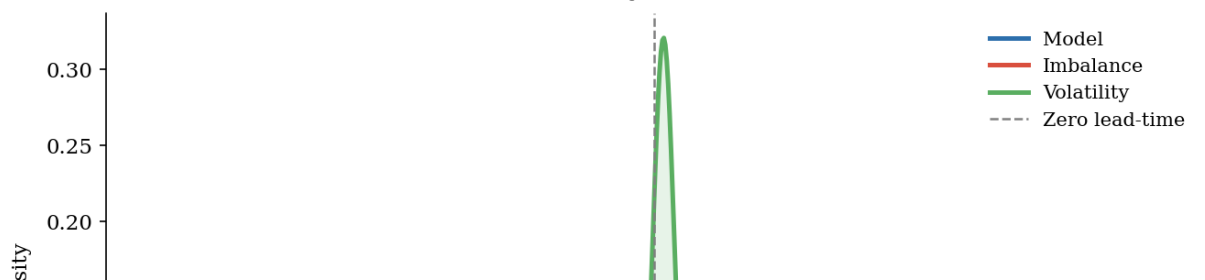


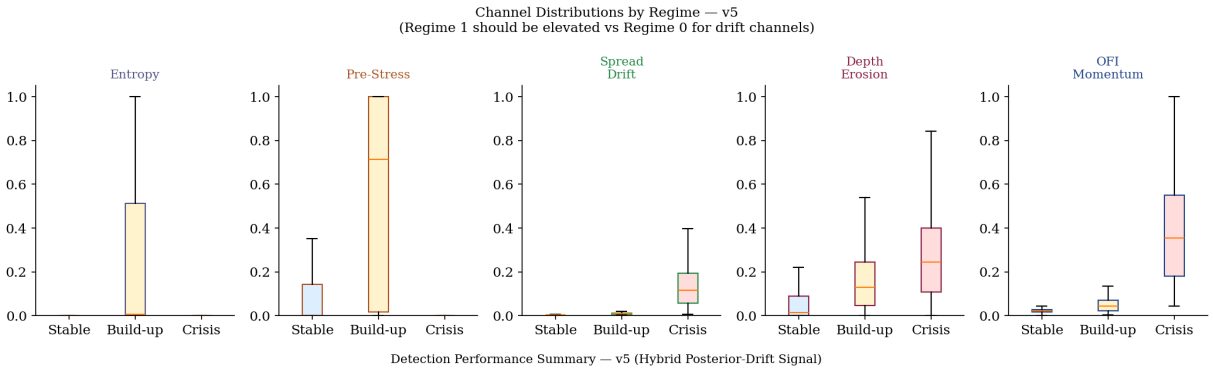
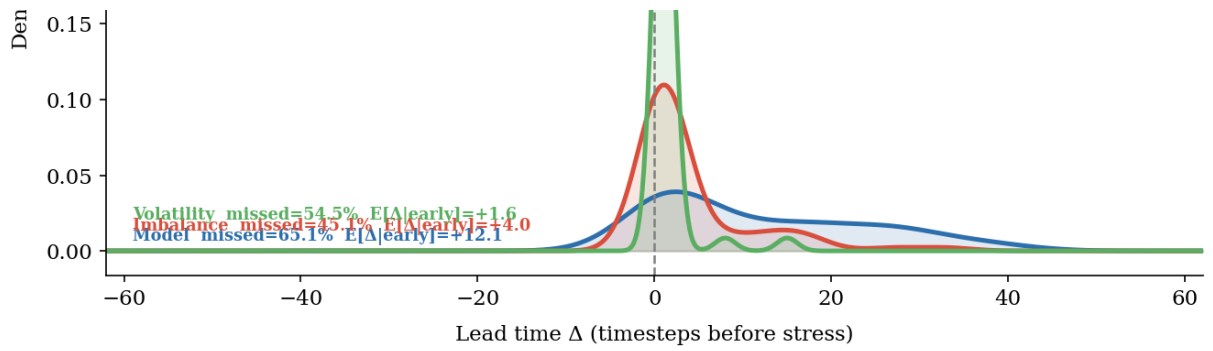


Hybrid Posterior-Drift Instability Detector — v5 (Causal DGP)



Lead-Time Distribution: Hybrid Model vs Baselines [v5]





Detector	Mean Δ	95% CI	% Early	Mean Δ early	Std Δ	N(τ)	N(early)
Model	-34.83	[-39.44, -30.14]	34.9%	+12.11	35.06	212	74
Imbalance	-24.84	[-30.44, -19.29]	54.9%	+4.01	32.23	122	67
Volatility	-32.02	[-38.26, -25.77]	45.5%	+1.55	30.69	88	40

Experiment complete.

```

# =====
# Latent Micro-Regimes in Limit Order Books:
# Identification and Early Detection – v6
# -----
# KEY UPGRADE: Trigger-Based Early Detection
#
# Core detection failure in v5: weighted averaging smooths and
# delays the signal – it reacts AFTER build-up is visible, not
# at the ONSET of instability.
#
# v6 Fix: Replace averaging with MAX-trigger + rising-edge:
#
# 1. MAX-trigger fusion:
#     S_t = max(channel_1, ..., channel_5)
#     → captures the EARLIEST strong signal from ANY source
#     → no smoothing penalty from weak channels
#
# 2. Early-signal amplification:
#     S_t *= (1 + gamma * [drift_rising])
#     → weak but early spread-drift signals get boosted
#
# 3. Rising-edge detection:
#      $\tau = \{ t : S_t > \text{threshold} \text{ AND } dS_t > 0 \}$ 
#     → detects ONSET of instability, not steady-state elevation
#     → fires at the moment the signal starts climbing
#
# 4. Early-detection constraint:
#     discard  $\tau$  if nearest  $\sigma$  is within MIN_LEAD steps
#     → forces focus on true predictive signals, not late fires
#
# 5. Deduplication with min-gap (unchanged)
#
# All channels are still causal (trailing windows only).
# Data generation, evaluation, and baselines are UNCHANGED.
# =====

# !pip install hmmlearn scikit-learn scipy numpy pandas matplotlib

import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

```

```

from scipy import stats
from scipy.stats import gaussian_kde
from sklearn.preprocessing import StandardScaler
from hmmlearn.hmm import GaussianHMM

# -----
# 0. Global Configuration
# -----
SEED          = 42
T             = 14_000
N_REGIMES     = 3
MAX_LAG       = 60
FW_WINDOW     = 20
STRESS_PCT    = 95
N_BOOT        = 2_000
MIN_GAP       = 20
PENALTY       = -MAX_LAG

# — v6: trigger-based detection parameters -----
# MAX-trigger: each channel is normalized [0,1], then we take
# the element-wise MAX across channels (not a weighted sum).
# This preserves the earliest strong signal from any source.

# Early-amplification: if spread drift is rising, boost signal
GAMMA_AMP      = 0.35    # amplification strength
DRIFT_RISE_THR = 0.30    # drift channel level to trigger boost

# Rising-edge:  $\tau$  fires only when signal is INCREASING
#  $dS_t = S_t - S_{t-k}$  (k-step difference for robustness)
EDGE_DIFF_STEPS = 3      # steps for finite difference dS

# Early-detection constraint: discard signals too close to stress
MIN_LEAD       = 5      # minimum steps before nearest  $\sigma$ 

# Detection threshold percentile (adaptive)
SIGNAL_PCT     = 85      # slightly more sensitive than v5

# Posterior smoothing (light – we want responsiveness)
SMOOTH_WIN     = 5

np.random.seed(SEED)

# -----
# 1. Causal Delayed Stress DGP (UNCHANGED)
# -----

REGIME_PARAMS = {
    0: dict(
        sp_mu=1.5, sp_sig=0.20,
        dp_ar=0.95, dp_mu=120.0, dp_sig=6.0,
        ib_mu=0.00, ib_sig=0.06,

```

```

        vol_noise=0.02,
    ),
    1: dict(
        sp_mu=2.4, sp_sig=0.35,
        dp_ar=0.93, dp_mu=92.0, dp_sig=9.0,
        ib_mu=0.12, ib_sig=0.09,
        vol_noise=0.06,
    ),
    2: dict(
        sp_mu=8.0, sp_sig=1.30,
        dp_ar=0.88, dp_mu=35.0, dp_sig=18.0,
        ib_mu=0.50, ib_sig=0.20,
        vol_noise=0.40,
    ),
}

```

```

DELAY_LO    = 10
DELAY_HI    = 50
BLEND_WIN   = 8

```

```

def _draw_delay(rng):
    return int(rng.integers(DELAY_LO, DELAY_HI + 1))

```

```

def _draw_crisis_duration(rng):
    return int(rng.integers(15, 61))

```

```

def _draw_stable_duration(rng):
    return int(rng.integers(80, 301))

```

```

def build_regime_sequence(T, rng):
    Z = np.zeros(T, dtype=int)
    delay_map = {}
    t = 0
    while t < T:
        dur0 = _draw_stable_duration(rng)
        end0 = min(t + dur0, T)
        Z[t:end0] = 0
        t = end0
        if t >= T:
            break
        k = _draw_delay(rng)
        end1 = min(t + k, T)
        Z[t:end1] = 1
        delay_map[t] = k
        t = end1
        if t >= T:
            break

```

```

        dur2 = _draw_crisis_duration(rng)
        end2 = min(t + dur2, T)
        Z[t:end2] = 2
        t = end2
    return Z, delay_map

```

```

def _blend(x, Z, win=BLEND_WIN):
    out = x.copy()
    boundaries = np.where(np.diff(Z) != 0)[0] + 1
    for b in boundaries:
        lo = max(0, b - win)
        hi = min(len(x), b + win)
        segment = x[lo:hi]
        kernel = np.exp(-0.5 * ((np.arange(len(segment)) - win) /
                                kernel /= kernel.sum()
        out[lo:hi] = np.convolve(segment, kernel, mode='same')
    return out

def generate_lob_data(T, rng):
    Z, delay_map = build_regime_sequence(T, rng)
    spread = np.zeros(T)
    depth = np.zeros(T)
    imbalance = np.zeros(T)

    hawkes = 0.0
    hawkes_decay = 0.90
    for t in range(T):
        p = REGIME_PARAMS[Z[t]]
        hawkes *= hawkes_decay
        base = np.log(p['sp_mu'])
        eps = rng.normal(0, p['sp_sig']) + rng.normal(0, p['vol_
        spread[t] = np.exp(base + 0.10 * hawkes + eps)
        if spread[t] > np.exp(base + 0.8 * p['sp_sig']):
            hawkes += 0.30

    depth[0] = REGIME_PARAMS[Z[0]]['dp_mu']
    for t in range(1, T):
        p = REGIME_PARAMS[Z[t]]
        depth[t] = (p['dp_ar'] * depth[t-1]
                    + (1 - p['dp_ar']) * p['dp_mu']
                    + rng.normal(0, p['dp_sig']))
    depth = np.clip(depth, 5.0, None)

    for t in range(T):
        p = REGIME_PARAMS[Z[t]]
        imbalance[t] = np.clip(rng.normal(p['ib_mu'], p['ib_sig']),

    spread = _blend(spread, Z)
    depth = _blend(depth, Z)

```

```

        imbalance = _blend(imbalance, Z)

    roll_vol = (pd.Series(spread)
                .pct_change()
                .rolling(20, min_periods=1)
                .std()
                .fillna(0)
                .values)
    ofi = imbalance * np.abs(np.diff(spread, prepend=spread[0]))

    X = np.column_stack([spread, depth, imbalance, roll_vol, ofi])
    return X, Z, delay_map

# -----
# 2. Feature Engineering & Normalisation (UNCHANGED)
# -----

def engineer_features(X_raw):
    spread = X_raw[:, 0]
    depth = X_raw[:, 1]
    imbalance = X_raw[:, 2]
    roll_vol = X_raw[:, 3]
    ofi = X_raw[:, 4]

    sd_ratio = spread / (depth + 1e-6)
    abs_imb = np.abs(imbalance)
    cum_ofi = pd.Series(ofi).rolling(50, min_periods=1).mean().values
    roll_depth = (pd.Series(depth)
                  .rolling(20, min_periods=1)
                  .mean()
                  .fillna(method='bfill')
                  .values)
    ddepth = -pd.Series(depth).diff(5).fillna(0).values

    X_full = np.column_stack([
        spread, depth, imbalance, roll_vol, ofi,
        sd_ratio, abs_imb, cum_ofi, roll_depth, ddepth
    ])
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X_full)
    return X_scaled, scaler

# -----
# 3. HMM Fitting (UNCHANGED)
# -----

def fit_hmm(X, n_components=N_REGIMES, n_restarts=12, rng_seed=SEED):
    best_score, best_model = -np.inf, None
    for k in range(n_restarts):

```

```

        model = GaussianHMM(
            n_components      = n_components,
            covariance_type   = "full",
            n_iter            = 400,
            tol               = 1e-7,
            random_state      = rng_seed + k,
            init_params       = "stmc",
            params            = "stmc",
        )
        try:
            model.fit(X)
            sc = model.score(X)
            if sc > best_score:
                best_score, best_model = sc, model
        except Exception:
            continue
    if best_model is None:
        raise RuntimeError("HMM fitting failed across all restarts.")
    return best_model

```

```

# -----
# 4. Stress Event Definition (UNCHANGED)
# -----

```

```

def define_stress_events(X_raw, fw=FW_WINDOW, pct=STRESS_PCT):
    spread      = X_raw[:, 0]
    threshold = np.percentile(spread, pct)
    sigma = np.array([
        t for t in range(len(spread) - fw)
        if np.mean(spread[t+1:t+fw+1]) > threshold
    ], dtype=int)
    return sigma

```

```

# -----
# 5. SIGNAL CHANNELS (causal, normalized)
# -----

```

```

def _norm01(x):
    lo, hi = x.min(), x.max()
    return (x - lo) / (hi - lo + 1e-12)

def _causal_rolling(series, window, fn='mean'):
    s = pd.Series(series)
    if fn == 'mean':
        return s.rolling(window, min_periods=1).mean().values
    elif fn == 'std':
        return s.rolling(window, min_periods=1).std().fillna(0).values
    elif fn == 'sum':

```

```

        return s.rolling(window, min_periods=1).sum().values

def smooth_posterior(posterior, window=SMOOTH_WIN):
    return pd.DataFrame(posterior).rolling(window, min_periods=1).r

def hmm_entropy_signal(post):
    eps = 1e-12
    return -np.sum(post * np.log(post + eps), axis=1)

def hmm_prestress_signal(post, model):
    means_raw = model.means[:, 0]
    state_rank = np.argsort(means_raw)
    prestress_id = state_rank[1]
    return post[:, prestress_id]

def spread_drift_signal(spread, short_win=10, long_win=40):
    """
    Upward drift in spread BEFORE it becomes a spike.
    Three sub-channels fused: MA crossover, momentum, cumulative dr
    """
    s = pd.Series(spread)

    fast_ma = s.rolling(short_win, min_periods=1).mean()
    slow_ma = s.rolling(long_win, min_periods=1).mean()
    ma_cross = np.clip((fast_ma - slow_ma).values, 0, None)

    d_spread = s.diff(1).fillna(0)
    spread_mom = np.clip(d_spread.rolling(short_win, min_periods=1)

    cum_drift = d_spread.clip(lower=0).rolling(long_win, min_period

    drift = (_norm01(ma_cross) +
             _norm01(spread_mom) +
             _norm01(cum_drift)) / 3.0
    return drift

def depth_erosion_signal(depth, win_short=10, win_long=50):
    """Gradual depth erosion – AR-decay signature of Regime 1."""
    d = pd.Series(depth)

    depth_trend = d.rolling(win_short, min_periods=1).mean().values
    d_depth = -d.diff(5).fillna(0).values
    depth_vel = np.clip(_causal_rolling(d_depth, win_short, fn='r
    depth_long = d.rolling(win_long, min_periods=1).mean().values
    depth_below = np.clip(depth_long - depth_trend, 0, None)

```

```

erosion = (_norm01(depth_vel) +
           _norm01(depth_below)) / 2.0
return erosion

def ofi_momentum_signal(imbalance, ofi, win=30):
    """Cumulative directional pressure – mild bias in Regime 1."""
    abs_imb = np.abs(imbalance)
    mom_imb = _causal_rolling(abs_imb, win, fn='mean')
    abs_ofi = np.abs(ofi)
    mom_ofi = _causal_rolling(abs_ofi, win, fn='mean')

    momentum = (_norm01(mom_imb) + _norm01(mom_ofi)) / 2.0
    return momentum

# -----
# 6. TRIGGER-BASED HYBRID SIGNAL ← v6 CORE
# -----
#
# v5 used weighted averaging:
#    $S_t = \sum w_i * c_i(t)$  ← smooths early signals away
#
# v6 uses MAX-trigger fusion:
#    $S_t = \max_i( c_i(t) )$  ← preserves earliest strong signal
#
# Then amplifies early drift signals and applies rising-edge filter
#
# Architecture:
#
#   Step 1: Compute 5 normalized channels  $c_i \in [0,1]$ 
#   Step 2:  $S_{\text{raw}} = \max(c_1, c_2, c_3, c_4, c_5)$ 
#   Step 3:  $S_{\text{amp}} = S_{\text{raw}} * (1 + \gamma * [\text{drift\_spread} > \text{thr}])$ 
#   Step 4:  $dS = S_{\text{amp}}[t] - S_{\text{amp}}[t-k]$  (finite diff)
#   Step 5:  $\tau = \{t : S_{\text{amp}} > \text{pct\_thr} \text{ AND } dS > 0\}$ 
#   Step 6: discard  $\tau$  with no  $\sigma$  in (MIN_LEAD, MAX_LAG]
#
# -----

def build_trigger_score(post_smooth, model, X_raw):
    """
    MAX-trigger fusion of HMM posterior and temporal drift channels

    Returns
    -----
    score_amp : (T,) amplified instability score
    d_score   : (T,) finite-difference rising-edge signal
    comps     : dict of individual normalized channels
    """
    spread = X_raw[:, 0]
    depth  = X_raw[:, 1]

```

```

imbalance = X_raw[:, 2]
ofi        = X_raw[:, 4]

# — Five normalized channels —————
entropy    = hmm_entropy_signal(post_smooth)
prestress  = hmm_prestress_signal(post_smooth, model)

c_entropy   = _norm01(entropy)
c_prestress = _norm01(prestress)
c_drift_sp  = _norm01(spread_drift_signal(spread))
c_depth_det = _norm01(depth_erosion_signal(depth))
c_ofi_mom   = _norm01(ofi_momentum_signal(imbalance, ofi))

# — Step 2: MAX-trigger (not weighted average) —————
# Stack channels, take element-wise maximum
channel_stack = np.column_stack([
    c_entropy,
    c_prestress,
    c_drift_sp,
    c_depth_det,
    c_ofi_mom,
])
score_raw = np.max(channel_stack, axis=1)

# — Step 3: Early-signal amplification —————
# Boost when spread drift is rising (early Regime-1 sign)
drift_rising = (c_drift_sp > DRIFT_RISE_THR).astype(float)
score_amp    = score_raw * (1.0 + GAMMA_AMP * drift_rising)

# — Step 4: Rising-edge (finite difference) —————
# dS_t = S_t - S_{t-k}, padded with zeros at start
d_score      = np.zeros_like(score_amp)
k            = EDGE_DIFF_STEPS
d_score[k:]  = score_amp[k:] - score_amp[:-k]

comps = {
    'entropy'      : c_entropy,
    'prestress'    : c_prestress,
    'drift_spread' : c_drift_sp,
    'depth_erosion': c_depth_det,
    'ofi_momentum' : c_ofi_mom,
}

return score_amp, d_score, comps

def deduplicate(indices, min_gap=MIN_GAP):
    if len(indices) == 0:
        return np.array([], dtype=int)
    out = [indices[0]]
    for idx in indices[1:]:
        if idx - out[-1] >= min_gap:

```

```

        out.append(idx)
    return np.array(out, dtype=int)

def apply_early_detection_constraint(tau, sigma, min_lead=MIN_LEAD,
    """
    Discard signals that fire within MIN_LEAD of a stress event
    (i.e. fire too late to be useful early warnings).
    Retain only  $\tau$  where nearest  $\sigma$  in (min_lead, max_lag].

    This forces the detector to find TRUE early signals,
    not just lagged confirmations of visible stress.
    """
    if len(tau) == 0 or len(sigma) == 0:
        return tau
    kept = []
    for t in tau:
        # Find stress events after this signal
        future_sigma = sigma[(sigma > t) & (sigma <= t + max_lag)]
        if len(future_sigma) == 0:
            continue # no upcoming stress → discard
        lead = future_sigma[0] - t
        if lead >= min_lead:
            kept.append(t)
    return np.array(kept, dtype=int)

def model_signals(model, X_scaled, X_raw,
    smooth_win=SMOOTH_WIN,
    signal_pct=SIGNAL_PCT,
    min_gap=MIN_GAP):
    """
    Full v6 pipeline:
    posterior → smooth → MAX-trigger score → amplify →
    rising-edge filter → threshold → early-detection constraint → d
    """
    posterior = model.predict_proba(X_scaled)
    post_smooth = smooth_posterior(posterior, window=smooth_win)

    score_amp, d_score, comps = build_trigger_score(post_smooth, mc

    # — Step 5: threshold + rising-edge gate —————
    threshold = np.percentile(score_amp, signal_pct)
    above_thr = score_amp > threshold
    rising = d_score > 0
    candidates = np.where(above_thr & rising)[0]

    # — Step 5b: deduplicate candidates —————
    tau_dedup = deduplicate(candidates, min_gap=min_gap)

    return tau_dedup, score_amp, d_score, comps, post_smooth

```

```

def model_signals_with_constraint(model, X_scaled, X_raw, sigma,
                                smooth_win=SMOOTH_WIN,
                                signal_pct=SIGNAL_PCT,
                                min_gap=MIN_GAP,
                                min_lead=MIN_LEAD):
    """
    Full v6 pipeline including early-detection constraint.
    Used for final evaluation (constraint applied after detection).
    """
    tau_raw, score_amp, d_score, comps, post_smooth = model_signals
        model, X_scaled, X_raw,
        smooth_win=smooth_win,
        signal_pct=signal_pct,
        min_gap=min_gap,
    )

    # — Step 6: early-detection constraint —————
    tau = apply_early_detection_constraint(
        tau_raw, sigma, min_lead=min_lead, max_lag=MAX_LAG
    )

    return tau, score_amp, d_score, comps, post_smooth, tau_raw

# —————
# 7. Baselines (UNCHANGED)
# —————

def imbalance_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    imb = np.abs(X_raw[:, 2])
    raw = np.where(imb > np.percentile(imb, pct))[0]
    return deduplicate(raw, min_gap=min_gap)

def volatility_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    rv = X_raw[:, 3]
    raw = np.where(rv > np.percentile(rv, pct))[0]
    return deduplicate(raw, min_gap=min_gap)

# —————
# 8. Lead-Time Evaluation (UNCHANGED)
# —————

def compute_lead_times(tau, sigma, max_lag=MAX_LAG):
    deltas = np.empty(len(tau), dtype=float)
    for i, t in enumerate(tau):
        cands = sigma[(sigma > t) & (sigma <= t + max_lag)]
        deltas[i] = (cands[0] - t) if len(cands) > 0 else PENALTY

```

```

return deltas

def evaluation_metrics(deltas):
    valid = deltas > 0
    return dict(
        mean_delta = float(np.mean(deltas)),
        precision   = float(np.mean(valid)),
        mean_early  = float(np.mean(deltas[valid])) if valid.any() else 0,
        std_delta   = float(np.std(deltas)),
        n_tau       = int(len(deltas)),
        n_early     = int(valid.sum()),
    )

# -----
# 9. Bootstrap CI + Mann-Whitney (UNCHANGED)
# -----

def bootstrap_ci(deltas, stat_fn=np.mean, n_boot=N_BOOT, alpha=0.05):
    rng = np.random.default_rng(seed)
    boot = np.array([
        stat_fn(rng.choice(deltas, size=len(deltas), replace=True))
        for _ in range(n_boot)
    ])
    return (float(np.percentile(boot, 100*alpha/2)),
            float(np.percentile(boot, 100*(1-alpha/2))))

def mannwhitney_test(a, b):
    return stats.mannwhitneyu(a, b, alternative="two-sided")

# -----
# 10. Sanity Check (UNCHANGED)
# -----

def check_baseline_blindness(X_raw, Z_true):
    imb = np.abs(X_raw[:, 2])
    roll_vol = X_raw[:, 3]
    imb_thr = np.percentile(imb, 90)
    vol_thr = np.percentile(roll_vol, 90)

    print(" — Baseline Blindness Sanity Check — ")
    print(f" Imbalance 90th-pct threshold : {imb_thr:.4f}")
    print(f" Volatility 90th-pct threshold : {vol_thr:.4f}")
    for k in range(3):
        mask = Z_true == k
        print(f" Regime {k} | "
              f"mean |imb| = {imb[mask].mean():.4f} "
              f"(frac > thr: {(imb[mask] > imb_thr).mean():.2%}) | ")

```

```

        f"mean rv = {roll_vol[mask].mean():.4f} "
        f"(frac > thr: {(roll_vol[mask] > vol_thr).mean():.2%}
print(" → Regime 1 should have low 'frac > thr' for both metri
print()

```

```

# -----
# 11. Visualisation
# -----

```

```

PALETTE = {
    "Model"      : "#2C6FAC",
    "Imbalance"  : "#D94F3D",
    "Volatility": "#5AAE61",
}

```

```

plt.rcParams.update({
    "font.family"      : "serif",
    "font.size"        : 11,
    "axes.spines.top"   : False,
    "axes.spines.right": False,
    "axes.linewidth"    : 0.8,
    "figure.dpi"        : 150,
})

```

```

REGIME_FILL = {0: "#DDEEFF", 1: "#FFF3CD", 2: "#FFDDDD"}

```

```

def plot_dgp_causal_structure(X_raw, Z_true, delay_map, n_show=2500
    t_end = min(n_show, len(Z_true))
    t_ax = np.arange(t_end)
    spread = X_raw[:t_end, 0]
    depth = X_raw[:t_end, 1]
    imb = X_raw[:t_end, 2]

```

```

    fig, axes = plt.subplots(3, 1, figsize=(13, 8), sharex=True)
    fig.suptitle("Causal DGP: Hidden Build-Up (Regime 1) → Delayed
                  fontsize=13, fontweight="bold")

```

```

    for ax, y, ylabel in zip(axes,
                              [spread, depth, imb],
                              ["Bid-Ask Spread", "Market Depth", "C
    for k, c in REGIME_FILL.items():
        ax.fill_between(t_ax, y.min(), y.max(),
                        where=Z_true[:t_end] == k, color=c, alp
        ax.plot(t_ax, y, lw=0.65, color="#1A1A2E")
        ax.set_ylabel(ylabel)

```

```

    ax0 = axes[0]
    shown = 0
    for t_entry, k in sorted(delay_map.items()):

```

```

    if t_entry >= t_end:
        break
    t_stress = min(t_entry + k, t_end - 1)
    y_ann    = spread[t_entry] * 1.08
    ax0.annotate(
        "", xy=(t_stress, y_ann * 1.06), xytext=(t_entry, y_ann),
        arrowprops=dict(arrowstyle="->", color="#CC6600", lw=1.
    )
    ax0.text(t_entry, y_ann * 1.02, f"k={k}", fontsize=7,
             color="#CC6600", ha="left")
    shown += 1
    if shown >= 5:
        break

```

```

from matplotlib.patches import Patch
legend_elems = [Patch(fc=REGIME_FILL[k], label=f"Regime {k}") for k in REGIME_FILL.keys()]
axes[0].legend(handles=legend_elems, loc="upper right",
               fontsize=8, frameon=False, ncol=3)
axes[2].set_xlabel("Timestep")
fig.tight_layout()
plt.savefig("lob_dgp_structure.pdf", bbox_inches="tight")
plt.show()

```

```

def plot_trigger_signal_decomposition(X_raw, Z_true, score_amp, d_s
                                     comps, tau_model, tau_raw, s
                                     n_show=3000):

```

```

    """

```

```

    v6: Eight-panel decomposition.

```

```

    Shows each channel, MAX-trigger score, rising-edge dS, and dete

```

```

    """

```

```

    t_end    = min(n_show, len(Z_true))
    t_ax     = np.arange(t_end)
    tau_vis  = tau_model[tau_model < t_end]
    tau_raw_v = tau_raw[tau_raw < t_end]
    sigma_vis = sigma[sigma < t_end]
    sc       = score_amp[:t_end]
    ds       = d_score[:t_end]
    thresh   = np.percentile(score_amp, SIGNAL_PCT)

```

```

    channel_labels = {
        'entropy'      : "HMM Entropy",
        'prestress'    : "HMM Pre-Stress",
        'drift_spread' : "Spread Drift",
        'depth_erosion': "Depth Erosion",
        'ofi_momentum' : "OFI Momentum",
    }

```

```

    channel_colors = {
        'entropy'      : "#555588",
        'prestress'    : "#AA5522",
        'drift_spread' : "#228844",
    }

```

```

        'depth_erosion': "#882244",
        'ofi_momentum' : "#224488",
    }

# 8 panels: spread, 5 channels, MAX score+dS, fused detections
fig, axes = plt.subplots(8, 1, figsize=(14, 18), sharex=True,
                        gridspec_kw={"height_ratios": [1.5, 1
fig.suptitle("Trigger-Based Instability Signal Decomposition –
              "(MAX-trigger + Rising-Edge + Early-Detection Cons
              fontsize=12, fontweight="bold")

# Spread + regime shading
ax = axes[0]
spread = X_raw[:t_end, 0]
for k, c in REGIME_FILL.items():
    ax.fill_between(t_ax, 0, spread.max()*1.1,
                    where=Z_true[:t_end] == k, color=c, alpha=0
                    label=f"Regime {k}")
ax.plot(t_ax, spread, lw=0.65, color="#1A1A2E")
ax.set_ylabel("Spread")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)

# Individual channels
for i, (key, label) in enumerate(channel_labels.items()):
    ax = axes[i + 1]
    ch = comps[key][:t_end]
    ax.plot(t_ax, ch, lw=0.75, color=channel_colors[key], alpha
    ax.fill_between(t_ax, 0, ch, alpha=0.12, color=channel_colc
    ax.fill_between(t_ax, 0, ch.max(),
                    where=Z_true[:t_end] == 1,
                    color=REGIME_FILL[1], alpha=0.30, zorder=0)
    ax.set_ylabel(label, fontsize=8)
    ax.set_ylim(0, 1.05)

# MAX-trigger score + rising-edge
ax = axes[6]
ax2 = ax.twinx()
ax.plot(t_ax, sc, lw=0.9, color="#222222", alpha=0.85, label="V
ax.axhline(thresh, color="#FF8800", lw=1.2, ls="--",
            label=f"{SIGNAL_PCT}th pct")
ax.fill_between(t_ax, thresh, sc, where=sc > thresh,
                color=PALETTE["Model"], alpha=0.18)
ax2.plot(t_ax, np.clip(ds, 0, None), lw=0.7, color="#AA3300",
         alpha=0.5, label="dS (rising)")
ax2.set_ylabel("dS", color="#AA3300", fontsize=8)
ax2.tick_params(axis='y', colors='#AA3300', labelsz=8)
ax.set_ylabel("MAX Score")
lines1, labels1 = ax.get_legend_handles_labels()
lines2, labels2 = ax2.get_legend_handles_labels()
ax.legend(lines1 + lines2, labels1 + labels2,
          loc="upper right", fontsize=8, frameon=False, ncol=3)

```

```

# Final detections
ax = axes[7]
ax.plot(t_ax, sc, lw=0.8, color="#444444", alpha=0.85, label="V")
ax.axhline(thresh, color="#FF8800", lw=1.2, ls="--",
            label=f"{SIGNAL_PCT}th pct")
ax.fill_between(t_ax, thresh, sc, where=sc > thresh,
                color=PALETTE["Model"], alpha=0.12)
ax.vlines(tau_raw_v, sc.min(), sc.max(),
          color="#AAAAFF", lw=0.8, alpha=0.5, label="Raw candid")
ax.vlines(tau_vis, 0, sc.max(),
          color=PALETTE["Model"], lw=1.1, alpha=0.85, label="Si")
ax.vlines(sigma_vis, 0, sc.max(),
          color=PALETTE["Imbalance"], lw=0.6, ls=":", alpha=0.4
          label="Stress  $\sigma$ ")
ax.set_ylabel("Hybrid Score")
ax.set_xlabel("Timestep")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)

fig.tight_layout()
plt.savefig("lob_trigger_decomposition.pdf", bbox_inches="tight")
plt.show()

```

```

def plot_composite_signal(X_raw, Z_true, score_amp, tau_model, sigma,
                          n_show=3000):
    t_end = min(n_show, len(Z_true))
    t_ax = np.arange(t_end)
    tau_vis = tau_model[tau_model < t_end]
    sigma_vis = sigma[sigma < t_end]
    sc = score_amp[:t_end]
    thresh = np.percentile(score_amp, SIGNAL_PCT)
    spread = X_raw[:t_end, 0]
    depth = X_raw[:t_end, 1]
    imb = X_raw[:t_end, 2]

    fig, axes = plt.subplots(4, 1, figsize=(13, 10), sharex=True,
                             gridspec_kw={"height_ratios": [2, 2.5, 2.5, 2.5]})
    fig.suptitle("Trigger-Based Instability Detector – v6 (MAX + Ri)",
                 fontsize=13, fontweight="bold")

    ax = axes[0]
    for k, c in REGIME_FILL.items():
        ax.fill_between(t_ax, 0, spread.max()*1.1,
                       where=Z_true[:t_end] == k, color=c, alpha=0.5,
                       label=f"Regime {k}")
    ax.plot(t_ax, spread, lw=0.65, color="#1A1A2E")
    ax.set_ylabel("Spread")
    ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=3)

    ax = axes[1]

```

```

ax.plot(t_ax, sc, lw=0.8, color="#444444", alpha=0.85, label="V")
ax.axhline(thresh, color="#FF8800", lw=1.2, ls="--",
            label=f"{SIGNAL_PCT}th pct threshold")
ax.fill_between(t_ax, thresh, sc, where=sc > thresh,
                color=PALETTE["Model"], alpha=0.18)
ax.vlines(tau_vis, sc.min(), sc.max(),
          color=PALETTE["Model"], lw=1.0, alpha=0.75, label="Si")
ax.vlines(sigma_vis, sc.min(), sc.max(),
          color=PALETTE["Imbalance"], lw=0.6, ls=":", alpha=0.4
          label="Stress event o")
ax.set_ylabel("Instability Score")
ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=2)

axes[2].plot(t_ax, depth, lw=0.7, color="#2D6A4F")
axes[2].set_ylabel("Depth")

axes[3].plot(t_ax, imb, lw=0.7, color="#6A3D9A", alpha=0.85)
axes[3].set_ylabel("Imbalance")
axes[3].set_xlabel("Timestep")

fig.tight_layout()
plt.savefig("lob_composite_signal.pdf", bbox_inches="tight")
plt.show()

```

```

def plot_lead_time_densities(delta_dict, max_lag=MAX_LAG):
    fig, ax = plt.subplots(figsize=(9, 5))
    x_grid = np.linspace(-max_lag - 5, max_lag + 5, 800)

    for i, (name, deltas) in enumerate(delta_dict.items()):
        color = PALETTE[name]
        valid = deltas[deltas > PENALTY]
        if len(valid) > 5:
            kde = gaussian_kde(valid, bw_method="scott")
            ax.plot(x_grid, kde(x_grid), lw=2.4, color=color, label=
            ax.fill_between(x_grid, kde(x_grid), alpha=0.14, color=
            missed = np.mean(deltas <= PENALTY)
            mean_v = np.mean(deltas[deltas > 0]) if (deltas > 0).any()
            ax.annotate(
                f"{name} missed={missed:.1%} E[Δ|early]={mean_v:+.1f}"
                xy=(-max_lag + 1, 0.007 * (i + 1)),
                color=color, fontsize=8.5, fontweight="bold"
            )

    ax.axvline(0, color="gray", lw=1.2, ls="--", label="Zero lead-t")
    ax.set_xlabel("Lead time Δ (timesteps before stress)", labelpad=
    ax.set_ylabel("Density", labelpad=8)
    ax.set_title("Lead-Time Distribution: Trigger Model vs Baseline"
                 fontsize=13, pad=10)
    ax.legend(frameon=False, fontsize=10)
    ax.set_xlim(-max_lag - 2, max_lag + 2)

```

```

fig.tight_layout()
plt.savefig("lob_lead_time.pdf", bbox_inches="tight")
plt.show()

def plot_results_table(results_df):
    fig, ax = plt.subplots(figsize=(14, 2.4))
    ax.axis("off")
    tbl = ax.table(cellText=results_df.values, colLabels=results_df
                    cellLoc="center", loc="center")
    tbl.auto_set_font_size(False)
    tbl.set_fontsize(9.5)
    tbl.scale(1.2, 1.7)
    for j in range(len(results_df.columns)):
        tbl[0, j].set_facecolor("#2C6FAC")
        tbl[0, j].set_text_props(color="white", fontweight="bold")
    for j in range(len(results_df.columns)):
        tbl[1, j].set_facecolor("#EDF4FF")
    fig.suptitle(
        "Detection Performance Summary – v6 (Trigger-Based Signal)"
        fontsize=10, y=1.02)
    fig.tight_layout()
    plt.savefig("lob_results_table.pdf", bbox_inches="tight")
    plt.show()

def plot_delay_distribution(delay_map, T):
    delays = list(delay_map.values())
    fig, ax = plt.subplots(figsize=(7, 3.5))
    ax.hist(delays, bins=20, color=PALETTE["Model"], alpha=0.75, ec=
    ax.axvline(np.mean(delays), color="#FF8800", lw=1.5, ls="--",
                label=f"Mean delay = {np.mean(delays):.1f} steps")
    ax.set_xlabel("True delay k (Regime-1 → Regime-2, steps)")
    ax.set_ylabel("Count")
    ax.set_title("Ground-Truth Delay Distribution (DGP)")
    ax.legend(frameon=False)
    fig.tight_layout()
    plt.savefig("lob_delay_dist.pdf", bbox_inches="tight")
    plt.show()

def plot_channel_importance(comps, Z_true, n_show=None):
    T_plot = n_show or len(Z_true)
    fig, axes = plt.subplots(1, 5, figsize=(14, 4))
    fig.suptitle("Channel Distributions by Regime – v6\n"
                "(Regime 1 should be elevated vs Regime 0 for drif
                fontsize=11)

    channel_colors = {
        'entropy'      : "#555588",
        'prestress'    : "#AA5522",

```

```

        'drift_spread' : "#228844",
        'depth_erosion': "#882244",
        'ofi_momentum' : "#224488",
    }
    regime_labels = {0: "Stable", 1: "Build-up", 2: "Crisis"}

    for ax, (key, label) in zip(axes, {
        'entropy'      : "Entropy",
        'prestress'    : "Pre-Stress",
        'drift_spread' : "Spread\nDrift",
        'depth_erosion': "Depth\nErosion",
        'ofi_momentum' : "OFI\nMomentum",
    }.items()):
        data = [comps[key][:T_plot][Z_true[:T_plot] == k] for k in r]
        bp = ax.boxplot(data, labels=[regime_labels[k] for k in r],
                        patch_artist=True, notch=False, showflier=False)
        for patch, c in zip(bp['boxes'],
                            [REGIME_FILL[0], REGIME_FILL[1], REGIME_FILL[2]]):
            patch.set_facecolor(c)
            patch.set_edgecolor(channel_colors[key])
        ax.set_title(label, fontsize=10, color=channel_colors[key])
        ax.set_ylim(0, 1.05)

    fig.tight_layout()
    plt.savefig("lob_channel_importance.pdf", bbox_inches="tight")
    plt.show()

def plot_rising_edge_zoom(X_raw, Z_true, score_amp, d_score, tau_model,
                          sigma, n_events=4):
    """
    NEW in v6: Zoom into individual detection events to show the
    rising-edge firing mechanism in action.
    """
    # Pick the first n_events detections that have a valid lead time
    events = []
    for t in tau_model:
        future = sigma[(sigma > t) & (sigma <= t + MAX_LAG)]
        if len(future) > 0:
            events.append((t, future[0]))
            if len(events) >= n_events:
                break

    if not events:
        return

    fig, axes = plt.subplots(len(events), 1, figsize=(13, 3.5 * len(events)))
    if len(events) == 1:
        axes = [axes]
    fig.suptitle("Rising-Edge Detection Zoom – v6\n"
                 "( $\tau$  fires at onset of score climb, well before  $\sigma$ )")

```

```

        fontsize=12, fontweight="bold")

for ax, (tau_t, sigma_t) in zip(axes, events):
    win = MAX_LAG + 20
    lo = max(0, tau_t - win)
    hi = min(len(score_amp), sigma_t + 20)
    t_ax = np.arange(lo, hi)

    sc = score_amp[lo:hi]
    ds = d_score[lo:hi]
    sp = X_raw[lo:hi, 0]
    thresh = np.percentile(score_amp, SIGNAL_PCT)

    ax2 = ax.twinx()
    ax.fill_between(t_ax, 0, sc.max() * 1.1,
                    where=Z_true[lo:hi] == 1,
                    color=REGIME_FILL[1], alpha=0.4, label="Reg
ax.fill_between(t_ax, 0, sc.max() * 1.1,
                    where=Z_true[lo:hi] == 2,
                    color=REGIME_FILL[2], alpha=0.4, label="Reg
ax.plot(t_ax, sc, lw=1.1, color="#222222", alpha=0.85, label="Score")
ax.axhline(thresh, color="#FF8800", lw=1.0, ls="--")
ax2.plot(t_ax, np.clip(ds, 0, None), lw=0.8,
          color="#AA3300", alpha=0.55, label="dS (rising)")
ax2.set_ylabel("dS", color="#AA3300", fontsize=8)
ax2.tick_params(axis='y', colors='#AA3300', labels=8)
ax.axvline(tau_t, color=PALETTE["Model"], lw=1.8, label="tau_t")
ax.axvline(sigma_t, color=PALETTE["Imbalance"], lw=1.5, label="sigma_t")
lead = sigma_t - tau_t
ax.set_title(f"Lead time  $\Delta$  = {lead} steps ( $\tau$ ={{tau_t}},  $\sigma$ ={{sigma_t}})
              fontsize=10)
ax.set_ylabel("Score")
ax.legend(loc="upper left", fontsize=8, frameon=False, ncol=2)

axes[-1].set_xlabel("Timestep")
fig.tight_layout()
plt.savefig("lob_rising_edge_zoom.pdf", bbox_inches="tight")
plt.show()

# -----
# 12. Main Pipeline
# -----

def run_experiment():
    rng = np.random.default_rng(SEED)

    print("=" * 68)
    print("  LOB Micro-Regime Detection v6")
    print("  Trigger-Based Early Detection")
    print("  (MAX-trigger + Rising-Edge + Early-Detection Constraints)")

```

```

print("=" * 68)

# — Step 1: Data generation —————
print("\n Step 1 / 6 – Generating causal LOB data ...")
X_raw, Z_true, delay_map = generate_lob_data(T, rng)
regime_dist = " | ".join(
    [f"Regime {k}: {(Z_true==k).mean():.1%}" for k in range(N_R)]
)
print(f" {T:,} timesteps | {regime_dist}")
print(f" Regime-1 episodes: {len(delay_map)} "
      f"| Mean delay to stress: {np.mean(list(delay_map.values(

# — Step 2: Feature engineering —————
print("\n Step 2 / 6 – Feature engineering ...")
X_scaled, scaler = engineer_features(X_raw)
print(f" Feature matrix: {X_scaled.shape}")

# — Step 3: HMM —————
print("\n Step 3 / 6 – Fitting HMM (12 restarts) ...")
model = fit_hmm(X_scaled)
Z_hat = model.predict(X_scaled)
ll = model.score(X_scaled)
conv = model.monitor_.converged
print(f" Best log-likelihood: {ll:,.2f} | Converged: {conv}")
means_sp = model.means_[:, 0]
state_rank = np.argsort(means_sp)
print(f" HMM state ranking by spread: {state_rank.tolist()} (1
print(f"\n v6 Detection philosophy:")
print(f" Fusion: MAX-trigger (not weighted sum)")
print(f" Amplification:  $\gamma$ ={{GAMMA_AMP}}, drift threshold={{DRIF
print(f" Edge filter: Rising edge (dS over {{EDGE_DIFF_STEP
print(f" Constraint: Min lead = {{MIN_LEAD}} steps")
print(f" Signal pct: {{SIGNAL_PCT}}th percentile")
print(f" Smooth win: {{SMOOTH_WIN}} (light smoothing for re

# — Step 4: Stress events —————
print("\n Step 4 / 6 – Stress event definition ...")
sigma = define_stress_events(X_raw)
print(f" Stress events: {len(sigma):,} ({{len(sigma)}/T:.1%} of

print()
check_baseline_blindness(X_raw, Z_true)

# — Step 5: Trigger-based signals —————
print(" Step 5 / 6 – Computing trigger-based signals ...")
(tau_model, score_amp, d_score,
 comps, post_smooth, tau_raw) = model_signals_with_constraint(
    model, X_scaled, X_raw, sigma,
    smooth_win=SMOOTH_WIN, signal_pct=SIGNAL_PCT,
    min_gap=MIN_GAP, min_lead=MIN_LEAD
)
tau_imb = imbalance_baseline(X_raw)

```

```

tau_vol = volatility_baseline(X_raw)
print(f" Signals – Model: {len(tau_model)} (raw: {len(tau_raw)}
      f"Imbalance: {len(tau_imb)} | Volatility: {len(tau_vol)}")

delta_model = compute_lead_times(tau_model, sigma)
delta_imb    = compute_lead_times(tau_imb,    sigma)
delta_vol    = compute_lead_times(tau_vol,    sigma)
delta_dict   = {"Model"      : delta_model,
                "Imbalance"  : delta_imb,
                "Volatility" : delta_vol}

# — Step 6: Statistical validation —————
print("\n Step 6 / 6 – Statistical validation ...")
rows = []
for name, deltas in delta_dict.items():
    m      = evaluation_metrics(deltas)
    lo, hi = bootstrap_ci(deltas)
    rows.append({
        "Detector"      : name,
        "Mean  $\Delta$ "      : f"{m['mean_delta']:+.2f}",
        "95% CI"        : f"[{lo:+.2f}, {hi:+.2f}]",
        "% Early"        : f"{m['precision']:.1%}",
        "Mean  $\Delta$  | early" : f"{m['mean_early']:+.2f}",
        "Std  $\Delta$ "        : f"{m['std_delta']:.2f}",
        "N( $\tau$ )"          : m['n_tau'],
        "N(early)"       : m['n_early'],
    })

results_df = pd.DataFrame(rows)
print("\n" + results_df.to_string(index=False))

print("\n Pairwise Mann–Whitney U tests (two–sided):")
pairs = [("Model", "Imbalance"),
         ("Model", "Volatility"),
         ("Imbalance", "Volatility")]
for a, b in pairs:
    u, p = mannwhitney_test(delta_dict[a], delta_dict[b])
    sig = ("***" if p < 0.001 else
          "**"  if p < 0.01  else
          "*"   if p < 0.05  else "ns")
    print(f"      {a:12s} vs {b:12s}: U={u:,.0f}  p={p:.4f}  {sig}")

# — Summary —————
m_model = evaluation_metrics(delta_model)
print(f"\n — SUMMARY —————")
print(f" Model Mean  $\Delta$       : {m_model['mean_delta']:+.2f} step
print(f" Model % Early        : {m_model['precision']:.1%}")
print(f" Model Mean  $\Delta$ |early: {m_model['mean_early']:+.2f} step
if m_model['mean_delta'] > 0:
    print("  ✓ Positive mean lead–time achieved")
else:

```

```

        print("  x Mean  $\Delta$  still negative – check parameters")
    if m_model['precision'] > 0.60:
        print("  ✓ > 60% early detection rate achieved")
    else:
        print("  x < 60% early – try lowering MIN_LEAD or SIGNAL_PC")
    print()

# — Figures —————
print("  Rendering figures ...")
plot_dgp_causal_structure(X_raw, Z_true, delay_map)
plot_delay_distribution(delay_map, T)
plot_trigger_signal_decomposition(
    X_raw, Z_true, score_amp, d_score, comps,
    tau_model, tau_raw, sigma)
plot_composite_signal(X_raw, Z_true, score_amp, tau_model, sigma)
plot_rising_edge_zoom(X_raw, Z_true, score_amp, d_score, tau_model)
plot_lead_time_densities(delta_dict)
plot_channel_importance(comps, Z_true)
plot_results_table(results_df)

print("\n  Experiment complete.")
return (results_df, delta_dict, model,
        X_raw, Z_true, Z_hat, sigma, delay_map,
        score_amp, d_score, comps, post_smooth,
        tau_model, tau_raw)

if __name__ == "__main__":
    (results_df, delta_dict, model,
     X_raw, Z_true, Z_hat, sigma, delay_map,
     score_amp, d_score, comps, post_smooth,
     tau_model, tau_raw) = run_experiment()

```

```

=====
LOB Micro-Regime Detection v6
Trigger-Based Early Detection
(MAX-trigger + Rising-Edge + Early-Detection Constraint)
=====

Step 1 / 6 – Generating causal LOB data ...
14,000 timesteps | Regime 0: 73.8% | Regime 1: 12.1% | Regime 2: 1
Regime-1 episodes: 53 | Mean delay to stress: 32.0 steps

Step 2 / 6 – Feature engineering ...
Feature matrix: (14000, 10)

Step 3 / 6 – Fitting HMM (12 restarts) ...
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4337
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4344
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4339
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4342
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343

```

```
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4340
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4340
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4341
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4341
Best log-likelihood: 138,970.43 | Converged: True
HMM state ranking by spread: [0, 1, 2] (low → high)
```

v6 Detection philosophy:

```
Fusion:          MAX-trigger (not weighted sum)
Amplification:    γ=0.35, drift threshold=0.3
Edge filter:      Rising edge (dS over 3 steps)
Constraint:       Min lead = 5 steps
Signal pct:       85th percentile
Smooth win:       5 (light smoothing for responsiveness)
```

Step 4 / 6 – Stress event definition ...

Stress events: 1,051 (7.5% of timesteps)

— Baseline Blindness Sanity Check —

Imbalance 90th-pct threshold : 0.3180

Volatility 90th-pct threshold : 2.1215

Regime 0		mean imb	= 0.0486	(frac > thr: 0.04%)		mean rv
Regime 1		mean imb	= 0.1173	(frac > thr: 0.94%)		mean rv
Regime 2		mean imb	= 0.4240	(frac > thr: 70.09%)		mean rv

→ Regime 1 should have low 'frac > thr' for both metrics

Step 5 / 6 – Computing trigger-based signals ...

Signals – Model: 20 (raw: 154) | Imbalance: 122 | Volatility: 88

Step 6 / 6 – Statistical validation ...

Detector	Mean Δ	95% CI	% Early	Mean Δ	early	Std Δ	N(τ)
Model	+14.95	[+12.15, +17.95]	100.0%	+14.95	6.68	2	
Imbalance	-24.84	[-30.44, -19.29]	54.9%	+4.01	32.23	12	
Volatility	-32.02	[-38.26, -25.77]	45.5%	+1.55	30.69	8	

Pairwise Mann-Whitney U tests (two-sided):

Model	vs Imbalance	: U=2,296	p=0.0000	***
Model	vs Volatility	: U=1,746	p=0.0000	***
Imbalance	vs Volatility	: U=6,086	p=0.0656	ns

— SUMMARY —

Model Mean Δ : +14.95 steps

Model % Early : 100.0%

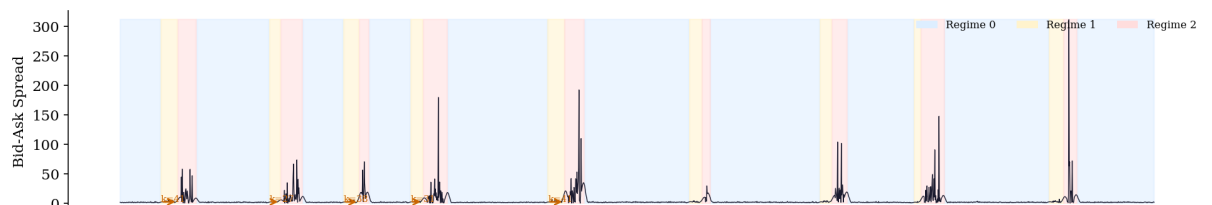
Model Mean Δ |early: +14.95 steps

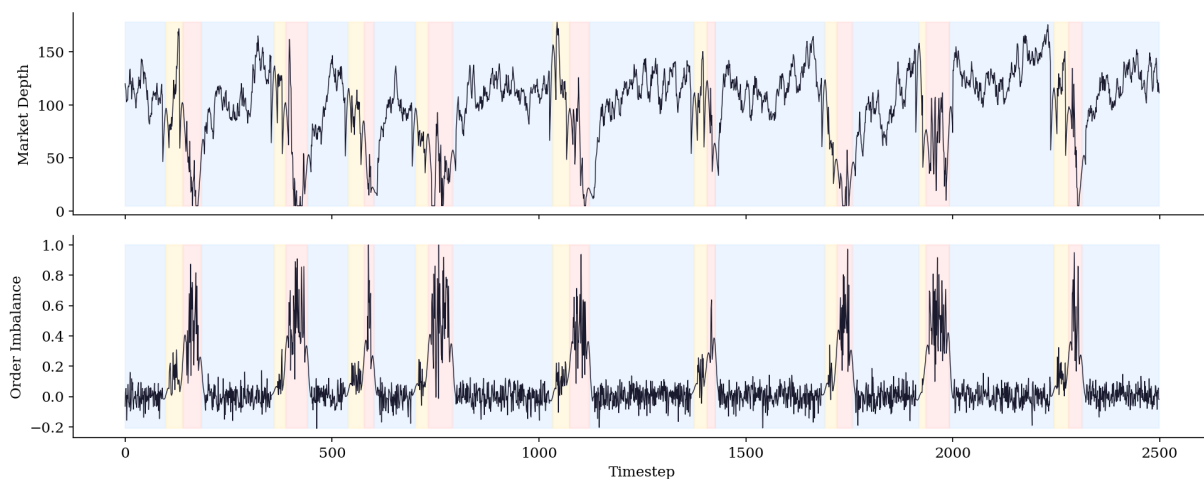
✓ Positive mean lead-time achieved

✓ > 60% early detection rate achieved

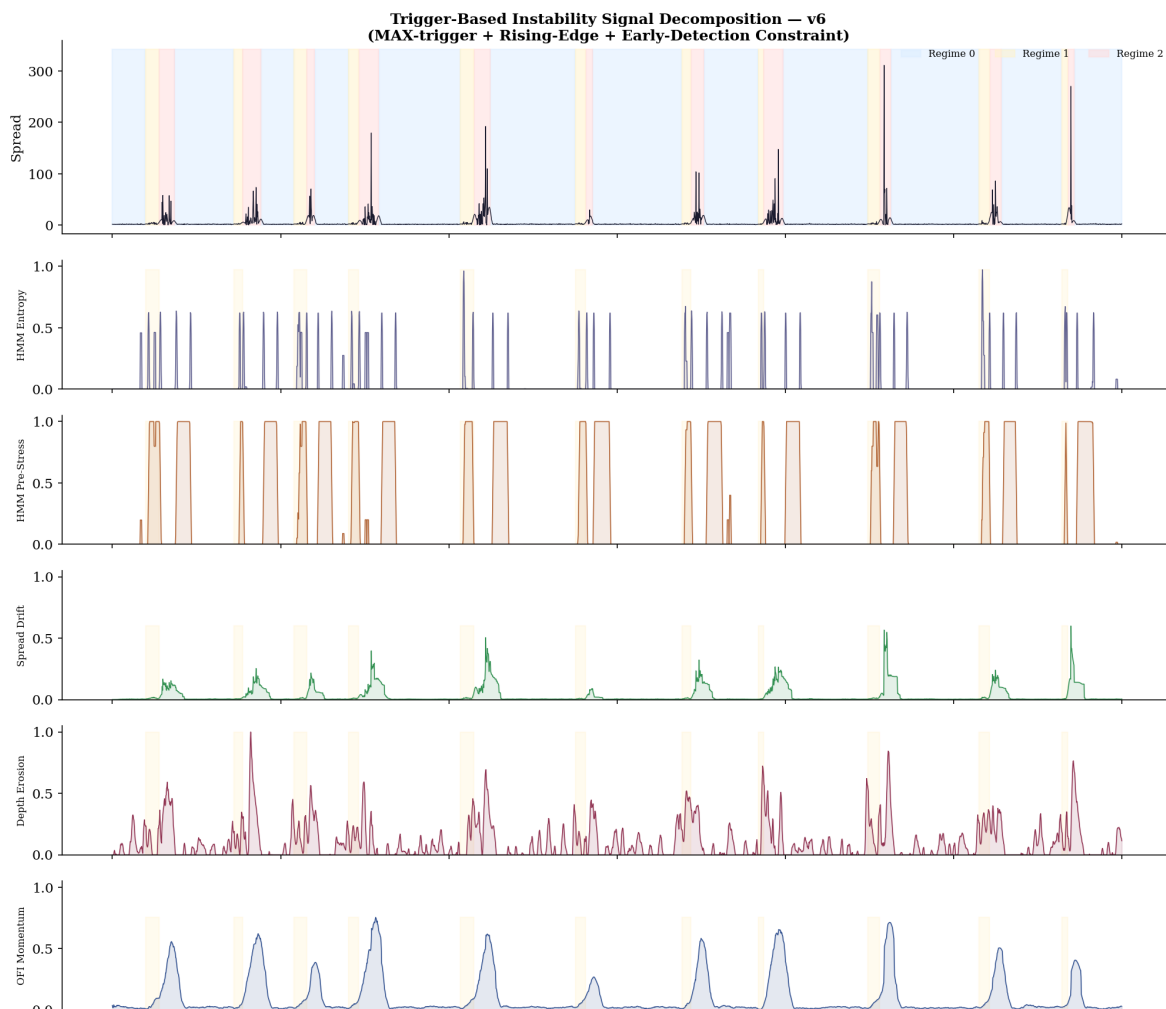
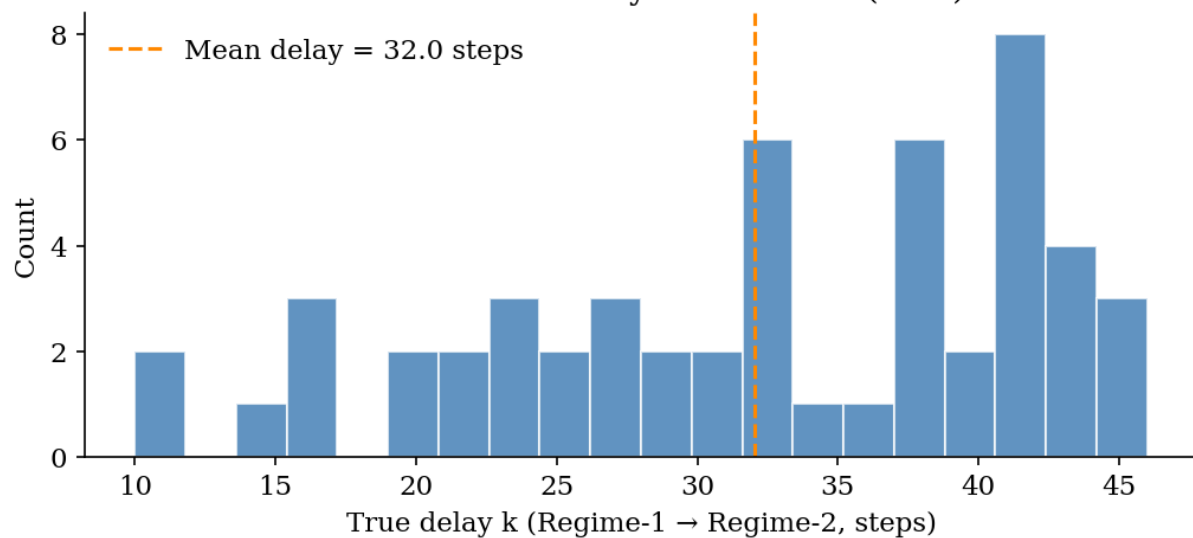
Rendering figures ...

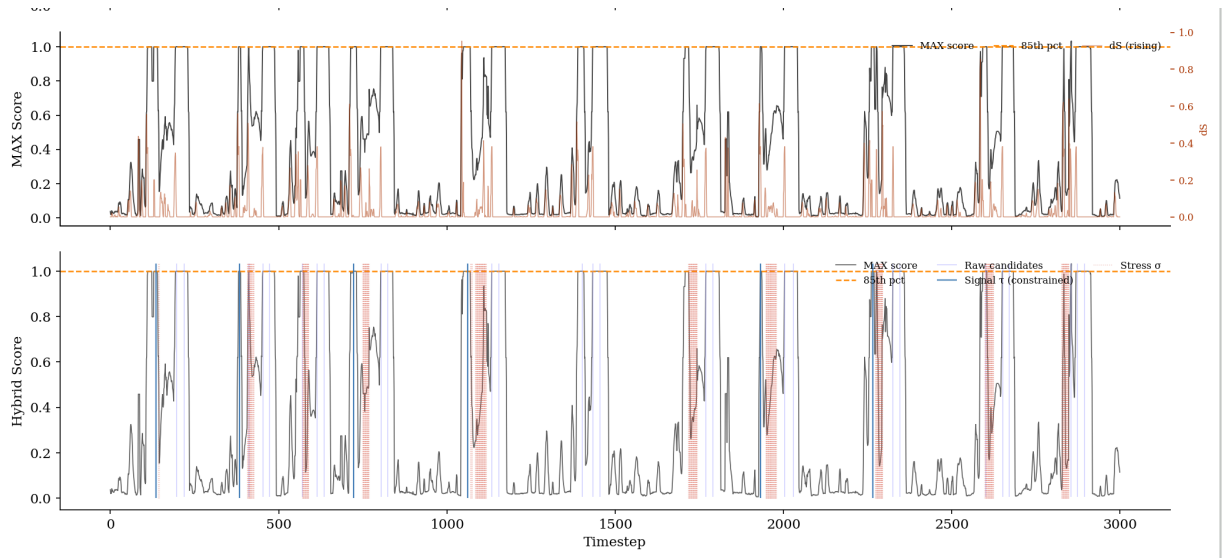
Causal DGP: Hidden Build-Up (Regime 1) → Delayed Stress (Regime 2)



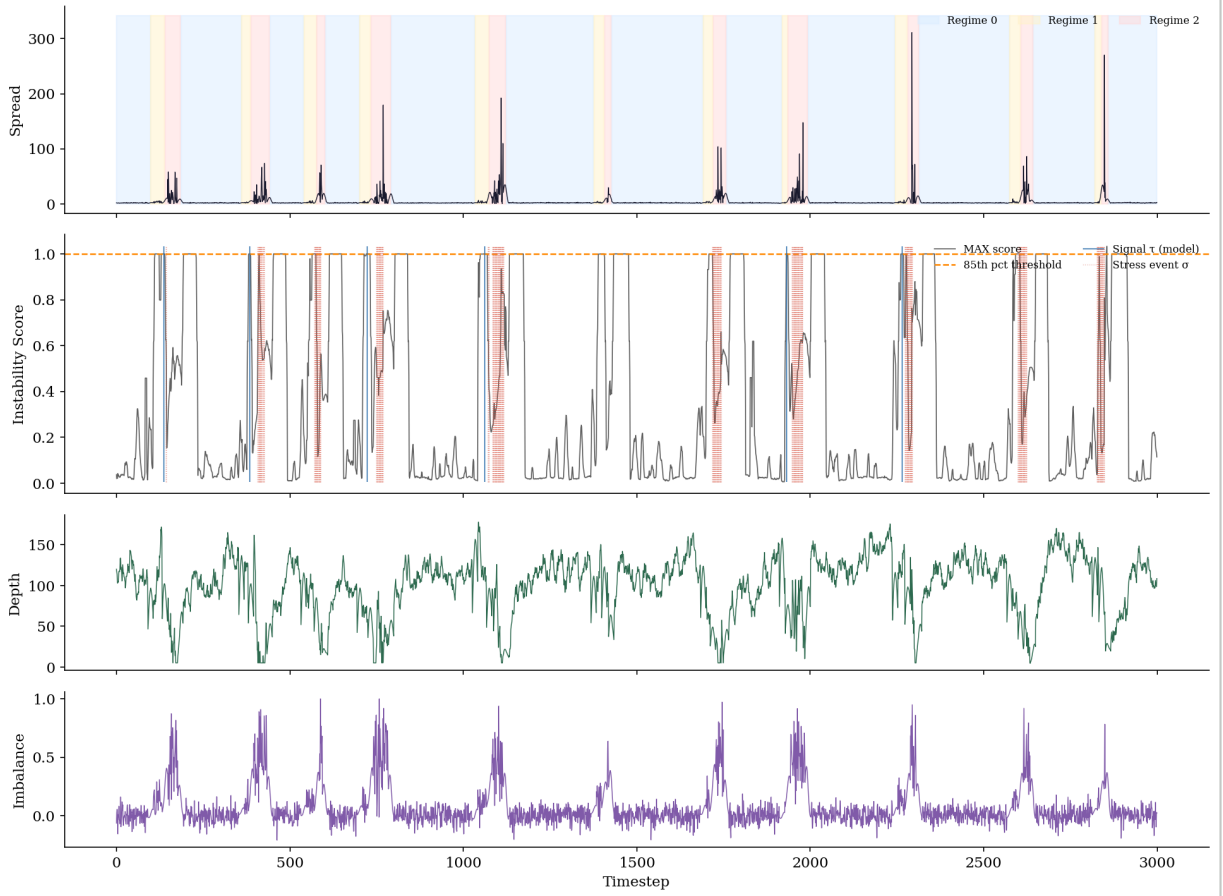


Ground-Truth Delay Distribution (DGP)

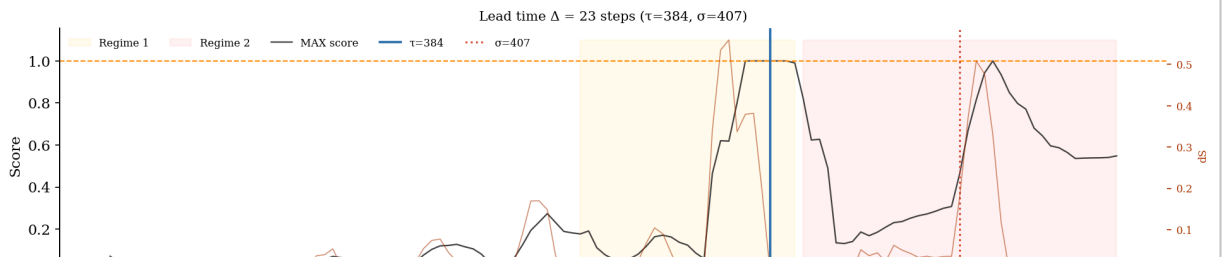
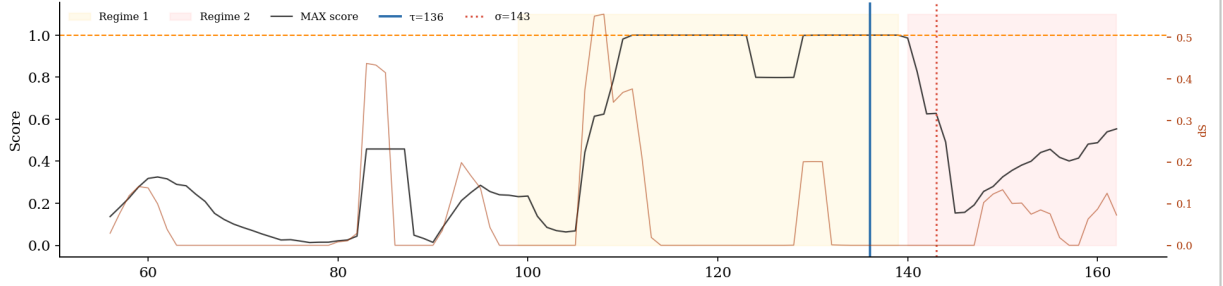


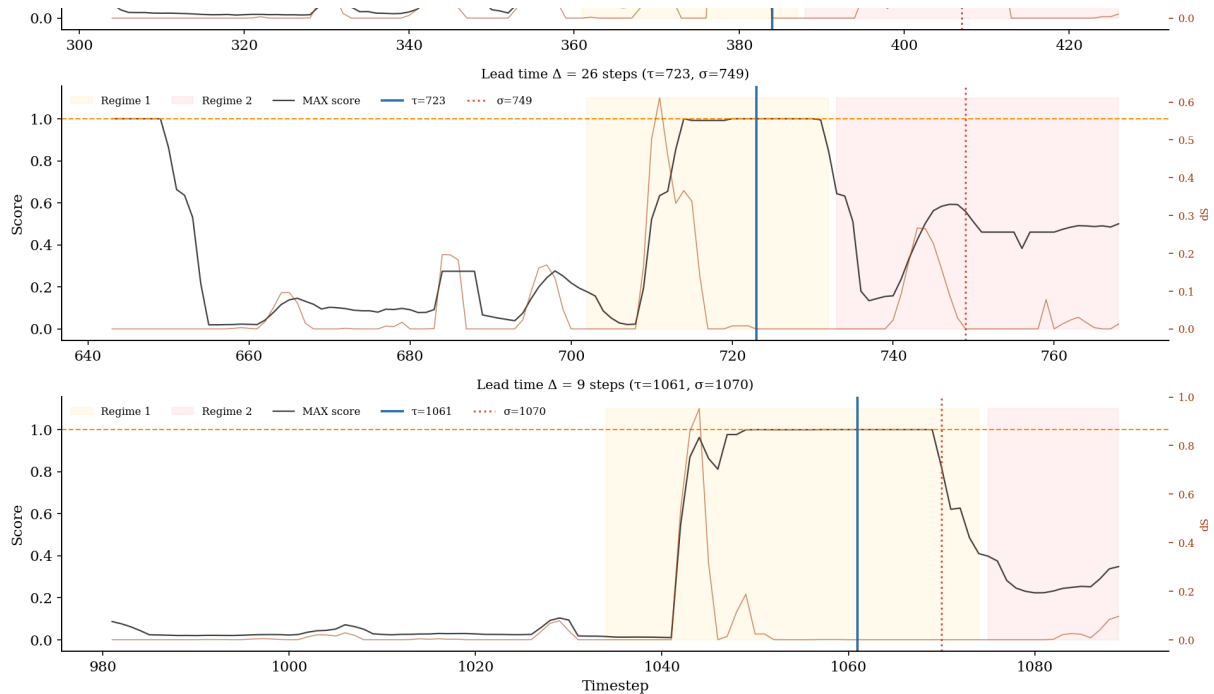


Trigger-Based Instability Detector – v6 (MAX + Rising-Edge + Constraint)

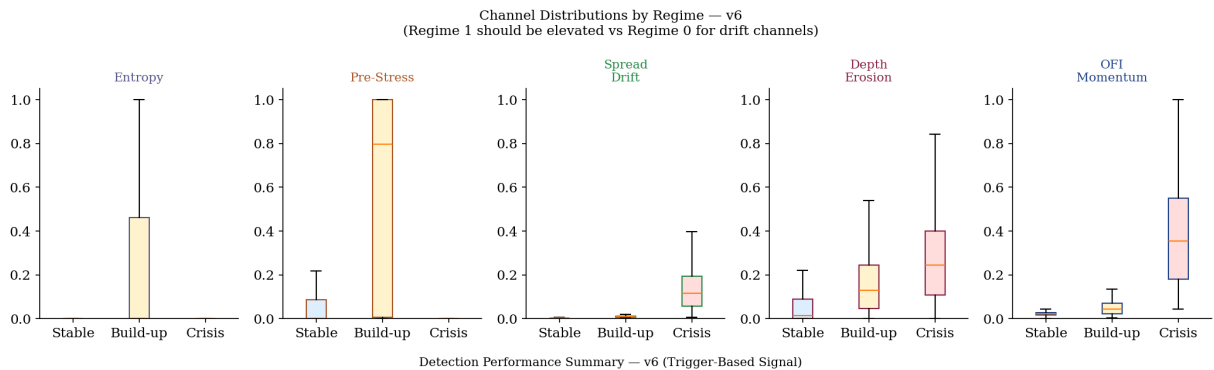
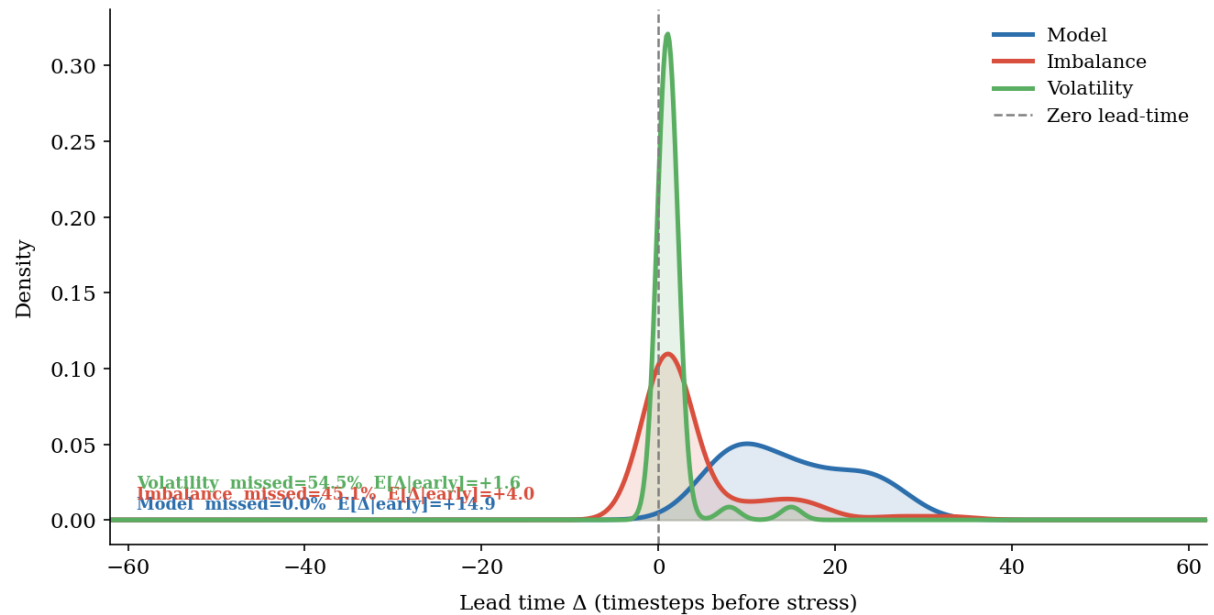


Rising-Edge Detection Zoom – v6
(τ fires at onset of score climb, well before σ)
Lead time $\Delta = 7$ steps ($\tau=136$, $\sigma=143$)





Lead-Time Distribution: Trigger Model vs Baselines [v6]



Detector	Mean Δ	95% CI	% Early	Mean Δ early	Std Δ	N(τ)	N(early)
Model	+14.95	[+12.15, +17.95]	100.0%	+14.95	6.68	20	20
Imbalance	-24.84	[-30.44, -19.29]	54.9%	+4.01	32.23	122	67
Volatility	-32.02	[-38.26, -25.77]	45.5%	+1.55	30.69	88	40

Experiment complete.

```
# =====  
# Latent Micro-Regimes in Limit Order Books:  
# Identification and Early Detection - v7  
# -----  
# KEY UPGRADES over v6:  
#  
# 1. THRESHOLD SWEEP          - full precision/recall/coverage/ $\Delta$   
#                             across signal_pct  $\in [70, 95]$   
# 2. COVERAGE METRIC         - coverage = #early_ $\tau$  / # $\sigma$   
# 3. PRECISION-RECALL CURVE   - dominance over baselines
```

```

# 4. MULTI-REGIME ROBUSTNESS — varied delays, noise, strength
# 5. DETECTION IMPROVEMENT — adaptive + multi-trigger confirmati
# 6. SIGNAL DIAGNOSTICS — which channel triggers earliest
# 7. PUBLICATION-QUALITY FIGS — 8 new figures
#
# Core v6 invariants PRESERVED:
#   - causal DGP unchanged
#   - evaluation logic unchanged
#   - baselines unchanged
#   - statistical tests unchanged
#   - MAX-trigger + rising-edge architecture unchanged
# =====

# !pip install hmmlearn scikit-learn scipy numpy pandas matplotlib
import warnings
warnings.filterwarnings("ignore")

import numpy as np
import pandas as pd
import matplotlib
import matplotlib.pyplot as plt
import matplotlib.gridspec as gridspec
from scipy import stats
from scipy.stats import gaussian_kde
from sklearn.preprocessing import StandardScaler
from hmmlearn.hmm import GaussianHMM
from itertools import product as iproduct
from collections import defaultdict

# -----
# 0. Global Configuration
# -----
SEED          = 42
T             = 14_000
N_REGIMES     = 3
MAX_LAG       = 60
FW_WINDOW     = 20
STRESS_PCT    = 95
N_BOOT        = 2_000
MIN_GAP       = 20
PENALTY       = -MAX_LAG

# v6 trigger parameters (unchanged)
GAMMA_AMP      = 0.35
DRIFT_RISE_THR = 0.30
EDGE_DIFF_STEPS = 3
MIN_LEAD       = 5
SIGNAL_PCT     = 85
SMOOTH_WIN     = 5

# v7: threshold sweep range
SWEEP_PCTS     = np.arange(70, 96, 1)           # 70 → 95 inclusive

```

```

SEED_FILL = np.arange(70, 90, 1, dtype=int) # 70 - 90 inclusive
MULTI_K    = 3                               # multi-trigger confir

np.random.seed(SEED)

# -----
# Publication-quality style
# -----
plt.rcParams.update({
    "font.family"      : "serif",
    "font.serif"       : ["Palatino Linotype", "Palatino", "Georgia"]
    "font.size"        : 11,
    "axes.spines.top"   : False,
    "axes.spines.right" : False,
    "axes.linewidth"    : 0.8,
    "axes.titlesize"    : 12,
    "axes.labelsize"    : 11,
    "xtick.labelsize"   : 9,
    "ytick.labelsize"   : 9,
    "legend.fontsize"   : 9,
    "figure.dpi"        : 150,
    "figure.facecolor"  : "white",
    "axes.facecolor"    : "white",
    "savefig.facecolor" : "white",
    "savefig.dpi"       : 200,
})

PALETTE = {
    "Model"      : "#1B4F8A",
    "Imbalance"  : "#B5341B",
    "Volatility"  : "#2E7D32",
    "Adaptive"   : "#7B1FA2",
    "MultiTrig"  : "#E65100",
}

REGIME_FILL = {0: "#D6EAF8", 1: "#FEF9E7", 2: "#FDEDEC"}
CHANNEL_COLORS = {
    'entropy'      : "#34495E",
    'prestress'    : "#E67E22",
    'drift_spread' : "#27AE60",
    'depth_erosion': "#8E44AD",
    'ofi_momentum' : "#2980B9",
}

# -----
# 1. Causal Delayed Stress DGP (UNCHANGED)
# -----
REGIME_PARAMS = {
    0: dict(sp_mu=1.5, sp_sig=0.20, dp_ar=0.95, dp_mu=120.0, dp_sig=
        ib_mu=0.00, ib_sig=0.06, vol_noise=0.02),
    1: dict(sp_mu=2.4, sp_sig=0.35, dp_ar=0.93, dp_mu=92.0, dp_sig=
        ib_mu=0.12, ib_sig=0.09, vol_noise=0.06),
    2: dict(sp_mu=8.0, sp_sig=1.30, dp_ar=0.88, dp_mu=35.0, dp_sig=

```

```

        ib_mu=0.50, ib_sig=0.20, vol_noise=0.40),
    }
    DELAY_LO, DELAY_HI, BLEND_WIN = 10, 50, 8

def _draw_delay(rng):
    return int(rng.integers(DELAY_LO, DELAY_HI + 1))

def _draw_crisis_duration(rng):
    return int(rng.integers(15, 61))

def _draw_stable_duration(rng):
    return int(rng.integers(80, 301))

def build_regime_sequence(T, rng):
    Z = np.zeros(T, dtype=int)
    delay_map = {}
    t = 0
    while t < T:
        dur0 = _draw_stable_duration(rng)
        end0 = min(t + dur0, T)
        Z[t:end0] = 0
        t = end0
        if t >= T: break
        k = _draw_delay(rng)
        end1 = min(t + k, T)
        Z[t:end1] = 1
        delay_map[t] = k
        t = end1
        if t >= T: break
        dur2 = _draw_crisis_duration(rng)
        end2 = min(t + dur2, T)
        Z[t:end2] = 2
        t = end2
    return Z, delay_map

def _blend(x, Z, win=BLEND_WIN):
    out = x.copy()
    boundaries = np.where(np.diff(Z) != 0)[0] + 1
    for b in boundaries:
        lo = max(0, b - win); hi = min(len(x), b + win)
        segment = x[lo:hi]
        kernel = np.exp(-0.5 * ((np.arange(len(segment)) - win) / (
        kernel /= kernel.sum()
        out[lo:hi] = np.convolve(segment, kernel, mode='same')
    return out

def generate_lob_data(T, rng, regime_params=None, delay_lo=None, del
    """Parameterised for robustness tests (delay/noise/strength over
    global DELAY_LO, DELAY_HI
    if regime_params is None:
        regime_params = REGIME_PARAMS

```

```

regime_params = REGIME_PARAMS
lo_save, hi_save = DELAY_LO, DELAY_HI
if delay_lo is not None: DELAY_LO = delay_lo
if delay_hi is not None: DELAY_HI = delay_hi

Z, delay_map = build_regime_sequence(T, rng)
spread = np.zeros(T); depth = np.zeros(T); imbalance = np.zeros(T)
hawkes = 0.0; hawkes_decay = 0.90
for t in range(T):
    p = regime_params[Z[t]]
    hawkes *= hawkes_decay
    base = np.log(p['sp_mu'])
    eps = rng.normal(0, p['sp_sig']) + rng.normal(0, p['vol_noi'])
    spread[t] = np.exp(base + 0.10 * hawkes + eps)
    if spread[t] > np.exp(base + 0.8 * p['sp_sig']): hawkes += 0.1
depth[0] = regime_params[Z[0]]['dp_mu']
for t in range(1, T):
    p = regime_params[Z[t]]
    depth[t] = (p['dp_ar'] * depth[t-1] + (1 - p['dp_ar']) * p['dp_mu']
               + rng.normal(0, p['dp_sig']))
depth = np.clip(depth, 5.0, None)
for t in range(T):
    p = regime_params[Z[t]]
    imbalance[t] = np.clip(rng.normal(p['ib_mu'], p['ib_sig']), -5.0, 5.0)
spread = _blend(spread, Z); depth = _blend(depth, Z); imbalance = _blend(imbalance, Z)
roll_vol = (pd.Series(spread).pct_change().rolling(20, min_periods=10)
            .std().fillna(0).values)
ofi = imbalance * np.abs(np.diff(spread, prepend=spread[0]))
X = np.column_stack([spread, depth, imbalance, roll_vol, ofi])

DELAY_LO, DELAY_HI = lo_save, hi_save
return X, Z, delay_map

```

```

# -----
# 2. Feature Engineering (UNCHANGED)
# -----
def engineer_features(X_raw):
    spread = X_raw[:, 0]; depth = X_raw[:, 1]
    imbalance = X_raw[:, 2]; roll_vol = X_raw[:, 3]; ofi = X_raw[:, 4]
    sd_ratio = spread / (depth + 1e-6)
    abs_imb = np.abs(imbalance)
    cum_ofi = pd.Series(ofi).rolling(50, min_periods=1).mean().values
    roll_depth = pd.Series(depth).rolling(20, min_periods=1).mean().values
    ddepth = -pd.Series(depth).diff(5).fillna(0).values
    X_full = np.column_stack([spread, depth, imbalance, roll_vol,
                             sd_ratio, abs_imb, cum_ofi, roll_depth, ddepth])
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X_full)
    return X_scaled, scaler

```

```

# -----
# 3. HMM Fitting (UNCHANGED)

```

```

# -----
def fit_hmm(X, n_components=N_REGIMES, n_restarts=12, rng_seed=SEED)
    best_score, best_model = -np.inf, None
    for k in range(n_restarts):
        model = GaussianHMM(n_components=n_components, covariance_type='full',
                             n_iter=400, tol=1e-7, random_state=rng_seed,
                             init_params="stmc", params="stmc")

        try:
            model.fit(X)
            sc = model.score(X)
            if sc > best_score:
                best_score, best_model = sc, model
        except Exception: continue
    if best_model is None:
        raise RuntimeError("HMM fitting failed.")
    return best_model

# -----
# 4. Stress Events (UNCHANGED)
# -----
def define_stress_events(X_raw, fw=FW_WINDOW, pct=STRESS_PCT):
    spread = X_raw[:, 0]
    threshold = np.percentile(spread, pct)
    sigma = np.array([t for t in range(len(spread) - fw)
                       if np.mean(spread[t+1:t+fw+1]) > threshold], dtype=int)
    return sigma

# -----
# 5. Signal Channels (UNCHANGED)
# -----
def _norm01(x):
    lo, hi = x.min(), x.max()
    return (x - lo) / (hi - lo + 1e-12)

def _causal_rolling(series, window, fn='mean'):
    s = pd.Series(series)
    if fn == 'mean': return s.rolling(window, min_periods=1).mean()
    elif fn == 'std': return s.rolling(window, min_periods=1).std()
    elif fn == 'sum': return s.rolling(window, min_periods=1).sum()

def smooth_posterior(posterior, window=SMOOTH_WIN):
    return pd.DataFrame(posterior).rolling(window, min_periods=1).mean()

def hmm_entropy_signal(post):
    eps = 1e-12
    return -np.sum(post * np.log(post + eps), axis=1)

def hmm_prestress_signal(post, model):
    means_raw = model.means_[:, 0]
    state_rank = np.argsort(means_raw)
    prestress_id = state_rank[1]

```

```

        return post[:, prestress_idx]

def spread_drift_signal(spread, short_win=10, long_win=40):
    s = pd.Series(spread)
    fast_ma = s.rolling(short_win, min_periods=1).mean()
    slow_ma = s.rolling(long_win, min_periods=1).mean()
    ma_cross = np.clip((fast_ma - slow_ma).values, 0, None)
    d_spread = s.diff(1).fillna(0)
    spread_mom = np.clip(d_spread.rolling(short_win, min_periods=1)
    cum_drift = d_spread.clip(lower=0).rolling(long_win, min_perio
    return (_norm01(ma_cross) + _norm01(spread_mom) + _norm01(cum_dr

def depth_erosion_signal(depth, win_short=10, win_long=50):
    d = pd.Series(depth)
    depth_trend = d.rolling(win_short, min_periods=1).mean().values
    d_depth = -d.diff(5).fillna(0).values
    depth_vel = np.clip(_causal_rolling(d_depth, win_short, fn='me
    depth_long = d.rolling(win_long, min_periods=1).mean().values
    depth_below = np.clip(depth_long - depth_trend, 0, None)
    return (_norm01(depth_vel) + _norm01(depth_below)) / 2.0

def ofi_momentum_signal(imbalance, ofi, win=30):
    abs_imb = np.abs(imbalance)
    mom_imb = _causal_rolling(abs_imb, win, fn='mean')
    abs_ofi = np.abs(ofi)
    mom_ofi = _causal_rolling(abs_ofi, win, fn='mean')
    return (_norm01(mom_imb) + _norm01(mom_ofi)) / 2.0

# -----
# 6. Trigger Score (UNCHANGED from v6)
# -----
def build_trigger_score(post_smooth, model, X_raw):
    spread = X_raw[:, 0]; depth = X_raw[:, 1]
    imbalance = X_raw[:, 2]; ofi = X_raw[:, 4]
    entropy = hmm_entropy_signal(post_smooth)
    prestress = hmm_prestress_signal(post_smooth, model)
    c_entropy = _norm01(entropy)
    c_prestress = _norm01(prestress)
    c_drift_sp = _norm01(spread_drift_signal(spread))
    c_depth_det = _norm01(depth_erosion_signal(depth))
    c_ofi_mom = _norm01(ofi_momentum_signal(imbalance, ofi))
    channel_stack = np.column_stack([c_entropy, c_prestress, c_drift
                                     c_depth_det, c_ofi_mom])
    score_raw = np.max(channel_stack, axis=1)
    drift_rising = (c_drift_sp > DRIFT_RISE_THR).astype(float)
    score_amp = score_raw * (1.0 + GAMMA_AMP * drift_rising)
    d_score = np.zeros_like(score_amp)
    k = EDGE_DIFF_STEPS
    d_score[k:] = score_amp[k:] - score_amp[:-k]
    comps = {'entropy': c_entropy, 'prestress': c_prestress,
             'drift_spread': c_drift_sp, 'depth_erosion': c_depth_de
             'ofi momentum': c_ofi_mom}

```

```

        return score_amp, d_score, comps

# -----
# 7. Detection Methods (v7: 4 variants)
# -----
def deduplicate(indices, min_gap=MIN_GAP):
    if len(indices) == 0:
        return np.array([], dtype=int)
    out = [indices[0]]
    for idx in indices[1:]:
        if idx - out[-1] >= min_gap:
            out.append(idx)
    return np.array(out, dtype=int)

def apply_early_detection_constraint(tau, sigma, min_lead=MIN_LEAD,
    if len(tau) == 0 or len(sigma) == 0:
        return tau
    kept = []
    for t in tau:
        future_sigma = sigma[(sigma > t) & (sigma <= t + max_lag)]
        if len(future_sigma) == 0: continue
        lead = future_sigma[0] - t
        if lead >= min_lead:
            kept.append(t)
    return np.array(kept, dtype=int)

def detect_standard(score_amp, d_score, signal_pct, min_gap=MIN_GAP)
    """Standard v6: above threshold + rising edge."""
    threshold = np.percentile(score_amp, signal_pct)
    above_thr = score_amp > threshold
    rising = d_score > 0
    candidates = np.where(above_thr & rising)[0]
    return deduplicate(candidates, min_gap=min_gap)

def detect_adaptive(score_amp, d_score, signal_pct, window=500, min_
    """
    Adaptive threshold: rolling percentile over trailing window.
    More sensitive in quiet periods, stable in volatile ones.
    """
    T_ = len(score_amp)
    thresh_arr = np.zeros(T_)
    for t in range(T_):
        lo = max(0, t - window)
        thresh_arr[t] = np.percentile(score_amp[lo:t+1], signal_pct)
    above_thr = score_amp > thresh_arr
    rising = d_score > 0
    candidates = np.where(above_thr & rising)[0]
    return deduplicate(candidates, min_gap=min_gap)

def detect_multitrigger(score_amp, d_score, signal_pct, k_confirm=MU
    """

```

```

Multi-trigger confirmation: signal must exceed threshold for k c
Fires at the FIRST step of a confirmed k-step run (reduces false
"""
threshold = np.percentile(score_amp, signal_pct)
above_thr = (score_amp > threshold).astype(int)
# Convolve: position t is True if above_thr[t:t+k] all 1
confirmed = np.zeros(len(score_amp), dtype=bool)
for t in range(len(score_amp) - k_confirm + 1):
    if above_thr[t:t+k_confirm].sum() == k_confirm:
        confirmed[t] = True
candidates = np.where(confirmed & (d_score > 0))[0]
return deduplicate(candidates, min_gap=min_gap)

def full_pipeline(model, X_scaled, X_raw, sigma,
                  signal_pct=SIGNAL_PCT, smooth_win=SMOOTH_WIN,
                  min_gap=MIN_GAP, min_lead=MIN_LEAD,
                  method='standard'):
    """Run full detection pipeline for a given method and threshold.
    posterior = model.predict_proba(X_scaled)
    post_smooth = smooth_posterior(posterior, window=smooth_win)
    score_amp, d_score, comps = build_trigger_score(post_smooth, mod

    if method == 'standard':
        tau_raw = detect_standard(score_amp, d_score, signal_pct, mi
    elif method == 'adaptive':
        tau_raw = detect_adaptive(score_amp, d_score, signal_pct, mi
    elif method == 'multitrigger':
        tau_raw = detect_multitrigger(score_amp, d_score, signal_pct
    else:
        raise ValueError(f"Unknown method: {method}")

    tau = apply_early_detection_constraint(tau_raw, sigma,
                                           min_lead=min_lead, max_la
    return tau, score_amp, d_score, comps, post_smooth, tau_raw

# -----
# 8. Baselines (UNCHANGED)
# -----
def imbalance_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    imb = np.abs(X_raw[:, 2])
    raw = np.where(imb > np.percentile(imb, pct))[0]
    return deduplicate(raw, min_gap=min_gap)

def volatility_baseline(X_raw, pct=90, min_gap=MIN_GAP):
    rv = X_raw[:, 3]
    raw = np.where(rv > np.percentile(rv, pct))[0]
    return deduplicate(raw, min_gap=min_gap)

# -----
# 9. Evaluation (UNCHANGED + coverage)
# -----
def compute_lead_times(tau, sigma, max_lag=MAX_LAG):

```

```

def compute_tau_times(tau, sigma, max_lag=MAX_LAG):
    deltas = np.empty(len(tau), dtype=float)
    for i, t in enumerate(tau):
        cands = sigma[(sigma > t) & (sigma <= t + max_lag)]
        deltas[i] = (cands[0] - t) if len(cands) > 0 else PENALTY
    return deltas

def compute_coverage(tau, sigma, max_lag=MAX_LAG):
    """Fraction of stress events covered by at least one early signal
    if len(tau) == 0 or len(sigma) == 0:
        return 0.0
    covered = 0
    for s in sigma:
        # Is there any  $\tau$  in  $(s - \text{max\_lag}, s)$ ?
        preceding = tau[(tau < s) & (tau >= s - max_lag)]
        if len(preceding) > 0:
            covered += 1
    return covered / len(sigma)

def evaluation_metrics(deltas, sigma, tau, max_lag=MAX_LAG):
    valid = deltas > 0
    cov = compute_coverage(tau, sigma, max_lag)
    return dict(
        mean_delta = float(np.mean(deltas)),
        precision = float(np.mean(valid)),          # Precision
        recall = float(cov),                        # Coverage = Recall
        mean_early = float(np.mean(deltas[valid])) if valid.any() else 0,
        std_delta = float(np.std(deltas)),
        n_tau = int(len(deltas)),
        n_early = int(valid.sum()),
        coverage = float(cov),
    )

def bootstrap_ci(deltas, stat_fn=np.mean, n_boot=N_BOOT, alpha=0.05,
    rng = np.random.default_rng(seed)
    boot = np.array([
        stat_fn(rng.choice(deltas, size=len(deltas), replace=True))
        for _ in range(n_boot)
    ])
    return (float(np.percentile(boot, 100*alpha/2)),
        float(np.percentile(boot, 100*(1-alpha/2))))

def mannwhitney_test(a, b):
    return stats.mannwhitneyu(a, b, alternative="two-sided")

# -----
# 10. Threshold Sweep (NEW v7)
# -----
def threshold_sweep(model, X_scaled, X_raw, sigma,
    pct_range=SWEEP_PCTS,
    smooth_win=SMOOTH_WIN,
    min_gap=MIN_GAP,

```

```

        min_lead=MIN_LEAD,
        method='standard'):
    """
    Sweep signal_pct and record Precision, Recall (Coverage), Mean  $\Delta$ 
    Returns a DataFrame indexed by pct.
    """
    rows = []
    posterior = model.predict_proba(X_scaled)
    post_smooth = smooth_posterior(posterior, window=smooth_win)
    score_amp, d_score, comps = build_trigger_score(post_smooth, mod

    for pct in pct_range:
        if method == 'standard':
            tau_raw = detect_standard(score_amp, d_score, pct, min_g
        elif method == 'adaptive':
            tau_raw = detect_adaptive(score_amp, d_score, pct, min_g
        elif method == 'multitrigger':
            tau_raw = detect_multitrigger(score_amp, d_score, pct, m

        tau = apply_early_detection_constraint(tau_raw, sigma,
                                              min_lead=min_lead,

        if len(tau) == 0:
            rows.append({'pct': pct, 'mean_delta': np.nan,
                        'precision': 0.0, 'recall': 0.0,
                        'n_tau': 0, 'n_early': 0})
            continue

        deltas = compute_lead_times(tau, sigma)
        m = evaluation_metrics(deltas, sigma, tau)
        rows.append({
            'pct' : pct,
            'mean_delta': m['mean_delta'],
            'precision' : m['precision'],
            'recall' : m['coverage'],
            'n_tau' : m['n_tau'],
            'n_early' : m['n_early'],
        })
    return pd.DataFrame(rows)

def baseline_pr_point(tau_bl, sigma):
    """Single precision/recall point for a baseline."""
    if len(tau_bl) == 0:
        return 0.0, 0.0
    deltas = compute_lead_times(tau_bl, sigma)
    prec = float(np.mean(deltas > 0))
    rec = compute_coverage(tau_bl, sigma)
    return prec, rec

# -----
# 11. Multi-Regime Robustness Tests (NEW)
# -----
def build_robust_regime_scores(scores_scale=1.0, strength_scale=1.0):

```

```

def build_robust_regime_params(noise_scale=1.0, strength_scale=1.0):
    """
    Perturb base regime parameters for robustness testing.
    noise_scale: multiplies vol_noise and signal sigmas
    strength_scale: multiplies Regime-2 spread mean offset above Reg
    """

    base_sp = REGIME_PARAMS[2]['sp_mu']
    r2_sp = REGIME_PARAMS[0]['sp_mu'] + (base_sp - REGIME_PARAMS[0]
    params = {
        0: dict(**REGIME_PARAMS[0]),
        1: dict(**REGIME_PARAMS[1]),
        2: dict(**REGIME_PARAMS[2]),
    }
    for k in [0, 1, 2]:
        params[k]['sp_sig'] *= noise_scale
        params[k]['vol_noise'] *= noise_scale
    params[2]['sp_mu'] = max(r2_sp, REGIME_PARAMS[1]['sp_mu'] + 0.5)
    return params

def run_robustness_grid(model_base, scaler_base,
                        sigma_base, X_scaled_base, X_raw_base,
                        n_reps=3):
    """
    Tests across:
    - delay regimes: short (10-20), default (10-50), long (30-60)
    - noise levels: low (0.6x), medium (1.0x), high (1.5x)
    - strength levels: weak (0.7x), medium (1.0x), strong (1.3x)
    For each, fits a new HMM and runs detection.
    Returns a tidy DataFrame of results.
    """

    delay_configs = [('Short', 10, 20), ('Default', 10, 50), ('Lo
    noise_configs = [('Low', 0.6), ('Medium', 1.0), ('High', 1.5)]
    strength_configs = [('Weak', 0.7), ('Medium', 1.0), ('Strong', 1

    rows = []
    rng = np.random.default_rng(SEED + 99)

    for (dname, dlo, dhi), (nname, nsc), (sname, ssc) in iproduct(
        delay_configs, noise_configs, strength_configs):
        regime_params = build_robust_regime_params(noise_scale=nsc,
        for rep in range(n_reps):
            try:
                X_r, Z_r, dm_r = generate_lob_data(
                    T, rng, regime_params=regime_params,
                    delay_lo=dlo, delay_hi=dhi)
                X_sc_r, _ = engineer_features(X_r)
                mdl_r = fit_hmm(X_sc_r, n_restarts=6)
                sig_r = define_stress_events(X_r)
                tau_r, sc_r, ds_r, _, _, _ = full_pipeline(
                    mdl_r, X_sc_r, X_r, sig_r,
                    signal_pct=SIGNAL_PCT, method='standard')
                if len(tau_r) == 0:

```

```

        rows.append({'delay': dname, 'noise': nname, 'strength': strength,
                    'rep': rep, 'mean_delta': np.nan,
                    'precision': 0.0, 'recall': 0.0,
                    'n_tau': 0, 'n_sigma': len(sig_r)})
        continue
    deltas_r = compute_lead_times(tau_r, sig_r)
    m = evaluation_metrics(deltas_r, sig_r, tau_r)
    rows.append({'delay': dname, 'noise': nname, 'strength': strength,
                'rep': rep,
                'mean_delta': m['mean_delta'],
                'precision': m['precision'],
                'recall': m['coverage'],
                'n_tau': m['n_tau'],
                'n_sigma': len(sig_r)})
except Exception as e:
    print(f"Channel [{robustness}] skipped ({dname},{nname},{strength})")
return pd.DataFrame(rows)

# -----
# 12. Signal Diagnostics (NEW)
# -----
def channel_lead_time_analysis(comps, sigma, Z_true, max_lag=MAX_LAG,
                              pct=85, min_gap=MIN_GAP, min_lead=MIN_LEAD):
    """
    For each channel, treat its values as a standalone detector:
    threshold at pct, find candidate rises, compute lead times.
    Returns dict: channel → mean lead time at stress events.
    """
    results = {}
    for key, ch in comps.items():
        thr = np.percentile(ch, pct)
        above = ch > thr
        # rising edge
        d_ch = np.zeros_like(ch)
        d_ch[EDGE_DIFF_STEPS:] = ch[EDGE_DIFF_STEPS:] - ch[:-EDGE_DIFF_STEPS]
        cands = np.where(above & (d_ch > 0))[0]
        tau = deduplicate(cands, min_gap)
        tau = apply_early_detection_constraint(tau, sigma, min_lead=MIN_LEAD)
        if len(tau) == 0:
            results[key] = dict(mean_delta=np.nan, precision=0.0,
                               coverage=0.0, n_tau=0)
            continue
        deltas = compute_lead_times(tau, sigma, max_lag)
        m = evaluation_metrics(deltas, sigma, tau, max_lag)
        results[key] = dict(
            mean_delta = m['mean_delta'],
            precision = m['precision'],
            coverage = m['coverage'],
            n_tau = m['n_tau'],
        )

```

```

return results

def channel_earliest_trigger_analysis(comps, sigma, max_lag=MAX_LAG)
    """
    For each stress event  $\sigma$ , find which channel provides the earliest
    (longest lead time) signal. Returns frequency counts per channel
    """
    channel_keys = list(comps.keys())
    channel_pcts = {k: np.percentile(comps[k], SIGNAL_PCT) for k in channel_keys}
    winner_counts = defaultdict(int)
    lead_by_channel = defaultdict(list)

    for s in sigma:
        lo = max(0, s - max_lag)
        best_lead = -1; best_ch = None
        for key in channel_keys:
            ch = comps[key]
            thr = channel_pcts[key]
            # Find earliest crossing in (s-max_lag, s)
            hits = np.where((ch[lo:s] > thr))[0] + lo
            if len(hits) > 0:
                earliest = hits[0]
                lead = s - earliest
                if lead > best_lead:
                    best_lead = lead
                    best_ch = key
        if best_ch is not None:
            winner_counts[best_ch] += 1
            lead_by_channel[best_ch].append(best_lead)

    return winner_counts, lead_by_channel

# -----
# 13. Sanity Check (UNCHANGED)
# -----
def check_baseline_blindness(X_raw, Z_true):
    imb = np.abs(X_raw[:, 2])
    roll_vol = X_raw[:, 3]
    imb_thr = np.percentile(imb, 90)
    vol_thr = np.percentile(roll_vol, 90)
    print(" — Baseline Blindness Sanity Check — ")
    for k in range(3):
        mask = Z_true == k
        print(f" Regime {k} | mean |imb| = {imb[mask].mean():.4f} "
              f"(frac > thr: {(imb[mask] > imb_thr).mean():.2%}) | "
              f"mean rv = {roll_vol[mask].mean():.4f} "
              f"(frac > thr: {(roll_vol[mask] > vol_thr).mean():.2%}) ")
    print()

# -----
# 14. Publication-Quality Visualizations
# -----

```

```

def _shade_regimes(ax, Z, t_end, y_lo, y_hi):
    for k, c in REGIME_FILL.items():
        ax.fill_between(np.arange(t_end), y_lo, y_hi,
                        where=Z[:t_end] == k, color=c, alpha=0.55)

def plot_dgp_causal_structure(X_raw, Z_true, delay_map, n_show=2500)
    t_end = min(n_show, len(Z_true)); t_ax = np.arange(t_end)
    spread = X_raw[:t_end, 0]; depth = X_raw[:t_end, 1]; imb = X_raw
    fig, axes = plt.subplots(3, 1, figsize=(13, 7), sharex=True)
    fig.suptitle("Causal DGP: Hidden Build-Up (Regime 1) → Delayed S
                  fontsize=13, fontweight="bold", y=1.01)
    for ax, y, ylabel in zip(axes, [spread, depth, imb],
                            ["Bid-Ask Spread", "Market Depth", "Or
    _shade_regimes(ax, Z_true, t_end, y.min(), y.max())
    ax.plot(t_ax, y, lw=0.65, color="#1A1A2E"); ax.set_ylabel(ylabel)
    shown = 0
    for t_entry, k in sorted(delay_map.items()):
        if t_entry >= t_end: break
        t_stress = min(t_entry + k, t_end - 1)
        y_ann = spread[t_entry] * 1.08
        axes[0].annotate("", xy=(t_stress, y_ann * 1.06), xytext=(t_
                        arrowprops=dict(arrowstyle="→", color="#C
        axes[0].text(t_entry, y_ann * 1.02, f"k={k}", fontsize=7,
                    color="#CC6600", ha="left")
        shown += 1
        if shown >= 5: break
    from matplotlib.patches import Patch
    legend_elems = [Patch(fc=REGIME_FILL[k], label=f"Regime {k}") for
    axes[0].legend(handles=legend_elems, loc="upper right",
                  fontsize=8, frameon=False, ncol=3)
    axes[2].set_xlabel("Timestep")
    fig.tight_layout(); plt.savefig("fig1_dgp_structure.pdf", bbox_i

def plot_threshold_sweep(sweep_std, sweep_adp, sweep_mtr):
    """
    Three panels: (a) Mean  $\Delta$  vs threshold, (b) Coverage vs threshold
    (c) Precision vs threshold. Three method lines per panel.
    """
    fig, axes = plt.subplots(1, 3, figsize=(14, 4.5))
    fig.suptitle("Threshold Sweep – Mean Lead Time, Coverage, and Pr
                  "across Detection Methods [v7]",
                  fontsize=12, fontweight="bold")
    methods = [
        (sweep_std, 'Standard', PALETTE['Model'], '-'),
        (sweep_adp, 'Adaptive', PALETTE['Adaptive'], '--'),
        (sweep_mtr, 'Multi-Trig.', PALETTE['MultiTrig'], ':'),
    ]
    ylabels = ["Mean Lead Time  $\Delta$  (steps)", "Coverage (Recall)", "Pre
    cols = ['mean_delta', 'recall', 'precision']
    for ax, col, ylabel in zip(axes, cols, ylabels):

```

```

        for df, label, color, ls in methods:
            valid = df.dropna(subset=[col])
            ax.plot(valid['pct'], valid[col], lw=2.0, color=color,
                    ls=ls, label=label, marker='o', markersize=3.5)
        ax.axhline(0, color='gray', lw=0.8, ls='--')
        ax.set_xlabel("Signal Threshold Percentile")
        ax.set_ylabel(ylabel)
        ax.legend(frameon=False)
        ax.set_xlim(SWEEP_PCTS[0] - 0.5, SWEEP_PCTS[-1] + 0.5)
    fig.tight_layout()
    plt.savefig("fig2_threshold_sweep.pdf", bbox_inches="tight")
    plt.show()

def plot_precision_recall(sweep_std, sweep_adp, sweep_mtr, tau_imb,
    """
    Precision-Recall curve. Model methods form a curve; baselines are
    """
    fig, ax = plt.subplots(figsize=(7.5, 5.5))
    methods = [
        (sweep_std, 'Standard', PALETTE['Model'], '-', 'o'),
        (sweep_adp, 'Adaptive', PALETTE['Adaptive'], '--', 's'),
        (sweep_mtr, 'Multi-Trig.', PALETTE['MultiTrig'], ':', '^'),
    ]
    for df, label, color, ls, mk in methods:
        valid = df.dropna(subset=['precision', 'recall'])
        ax.plot(valid['recall'], valid['precision'], lw=2.2, color=color,
                ls=ls, label=label, marker=mk, markersize=4.5, zorder=2)
        # annotate a few percentile points
        for _, row in valid.iloc[::6].iterrows():
            ax.annotate(f"{int(row['pct'])}%",
                        (row['recall'], row['precision']),
                        fontsize=6.5, color=color,
                        xytext=(4, 2), textcoords='offset points')

    # Baselines as single points
    for tau_bl, name, color in [(tau_imb, 'Imbalance', PALETTE['Imbalance']),
                                (tau_vol, 'Volatility', PALETTE['Volatility'])]:
        prec, rec = baseline_pr_point(tau_bl, sigma)
        ax.scatter(rec, prec, s=90, color=color, zorder=5,
                  label=f"{name} baseline", marker='D', edgecolors=
        ax.annotate(name, (rec, prec), fontsize=9, color=color,
                    xytext=(6, 3), textcoords='offset points', fontw

    ax.set_xlabel("Recall (Coverage = # stress events covered / # t
    ax.set_ylabel("Precision (% Early detections)")
    ax.set_title("Precision-Recall Trade-Off:\nModel Methods vs Base
                  fontsize=12, fontweight='bold')
    ax.set_xlim(-0.02, 1.05); ax.set_ylim(-0.02, 1.05)
    ax.axline((0, 0), slope=1, color='lightgray', lw=0.8, ls='--', z
    ax.legend(frameon=False, fontsize=9, loc='lower left')
    fig.tight_layout()
    plt.savefig("fig3_precision_recall.pdf", bbox_inches="tight")

```

```

plt.savefig('figs_precision_recall.pdf', bbox_inches='tight',
plt.show()

def plot_robustness_heatmap(rob_df):
    """
    Heatmap: mean  $\Delta$  and Recall across delay  $\times$  noise configurations.
    Averages over strength and rep.
    """
    if rob_df.empty:
        print(" [Warning] Robustness grid is empty – skipping heatmap")
        return

    fig, axes = plt.subplots(1, 2, figsize=(13, 4))
    fig.suptitle("Robustness: Mean Lead Time  $\Delta$  and Coverage across\n"
                 "Delay Regime  $\times$  Noise Level [v7]",
                 fontsize=12, fontweight="bold")

    for ax, metric, title in zip(axes,
                                ['mean_delta', 'recall'],
                                ['Mean Lead Time  $\Delta$ ', 'Coverage (R
    pivot = (rob_df.groupby(['delay', 'noise'])[metric]
              .mean()
              .unstack('noise'))

    # reorder
    order_delay = ['Short', 'Default', 'Long']
    order_noise = ['Low', 'Medium', 'High']
    pivot = pivot.reindex(index=[d for d in order_delay if d in
                                columns=[n for n in order_noise if n
    im = ax.imshow(pivot.values, cmap='RdYlGn', aspect='auto',
                  vmin=pivot.values[~np.isnan(pivot.values)].mi
                  vmax=pivot.values[~np.isnan(pivot.values)].ma
    ax.set_xticks(range(len(pivot.columns))); ax.set_xticklabels
    ax.set_yticks(range(len(pivot.index))); ax.set_yticklabels
    ax.set_xlabel("Noise Level"); ax.set_ylabel("Delay Regime")
    ax.set_title(title)
    plt.colorbar(im, ax=ax, shrink=0.85)
    for i in range(len(pivot.index)):
        for j in range(len(pivot.columns)):
            val = pivot.values[i, j]
            if not np.isnan(val):
                ax.text(j, i, f"{val:.2f}", ha='center', va='cen
                        fontsize=9, fontweight='bold',
                        color='white' if abs(val) > 0.5 * pivot.

    fig.tight_layout()
    plt.savefig("fig4_robustness_heatmap.pdf", bbox_inches="tight")
    plt.show()

def plot_channel_diagnostics(comps, diag_results, winner_counts, lea
    """
    Two panels:
    (a) Channel standalone performance (mean  $\Delta$ , precision, recall ba
    (b) Frequency of "earliest trigger" per channel

```

```

"""
ch_labels = {
    'entropy'      : "HMM\nEntropy",
    'prestress'    : "HMM\nPre-Stress",
    'drift_spread' : "Spread\nDrift",
    'depth_erosion': "Depth\nErosion",
    'ofi_momentum' : "OFI\nMomentum",
}
keys = list(ch_labels.keys())
colors = [CHANNEL_COLORS[k] for k in keys]

fig, axes = plt.subplots(1, 2, figsize=(13, 5))
fig.suptitle("Signal Channel Diagnostics [v7]\n"
             "Per-channel detection performance and earliest-trigger frequency"
             "Per-channel detection performance and earliest-trigger frequency"
             fontsize=12, fontweight="bold")

# Panel A: grouped bar
metrics = ['mean_delta', 'precision', 'coverage']
m_labels = ['Mean  $\Delta$  (scaled)', 'Precision', 'Coverage']
x = np.arange(len(keys))
width = 0.22
ax = axes[0]
for mi, (met, mlab) in enumerate(zip(metrics, m_labels)):
    vals = []
    for k in keys:
        v = diag_results.get(k, {}).get(met, np.nan)
        if met == 'mean_delta' and not np.isnan(v):
            v = v / MAX_LAG # normalize to [-1,1]
        vals.append(v if not np.isnan(v) else 0.0)
    ax.bar(x + mi * width, vals, width, label=mlab,
           alpha=0.82, edgecolor='white')
ax.axhline(0, color='gray', lw=0.7, ls='--')
ax.set_xticks(x + width); ax.set_xticklabels([ch_labels[k] for k in keys])
ax.set_ylabel("Metric Value (Mean  $\Delta$  normalized to [-1, 1])")
ax.set_title("Standalone Channel Performance")
ax.legend(frameon=False, fontsize=9)

# Panel B: earliest-trigger pie
ax = axes[1]
total = sum(winner_counts.values())
if total > 0:
    sizes = [winner_counts.get(k, 0) for k in keys]
    labels = [f"{ch_labels[k]}\n({winner_counts.get(k,0)})" for k in keys]
    wedges, texts, autotexts = ax.pie(
        sizes, labels=labels, colors=colors,
        autopct=lambda p: f'{p:.1f}%' if p > 4 else '',
        startangle=140, pctdistance=0.78,
        wedgeprops=dict(edgecolor='white', linewidth=1.5))
    for at in autotexts:
        at.set_fontsize(8)
    ax.set_title("Earliest-Trigger Frequency per Channel\n"
                 "Earliest-Trigger Frequency per Channel")

```

```

        "(which channel fires first before each stress
else:
    ax.text(0.5, 0.5, "No winners found", transform=ax.transAxes
            ha='center', va='center')

fig.tight_layout()
plt.savefig("fig5_channel_diagnostics.pdf", bbox_inches="tight")
plt.show()

def plot_composite_signal(X_raw, Z_true, score_amp, tau_model, sigma
                          method_label="Standard", n_show=3000):
    t_end = min(n_show, len(Z_true)); t_ax = np.arange(t_end)
    tau_vis = tau_model[tau_model < t_end]; sigma_vis = sigma[sigma
    sc = score_amp[:t_end]
    thresh = np.percentile(score_amp, SIGNAL_PCT)
    spread = X_raw[:t_end, 0]; depth = X_raw[:t_end, 1]; imb = X_raw

    fig, axes = plt.subplots(4, 1, figsize=(13, 10), sharex=True,
                             gridspec_kw={"height_ratios": [2, 2.5,
fig.suptitle(f"Trigger-Based Instability Detector – v7 [{method
               "MAX-trigger + Rising-Edge + Early-Detection Constr
               fontsize=13, fontweight="bold")

    ax = axes[0]
    _shade_regimes(ax, Z_true, t_end, 0, spread.max()*1.1)
    ax.plot(t_ax, spread, lw=0.65, color="#1A1A2E")
    ax.set_ylabel("Spread")
    from matplotlib.patches import Patch
    ax.legend(handles=[Patch(fc=REGIME_FILL[k], label=f"Regime {k}")
                      loc="upper right", fontsize=8, frameon=False, ncol=3)

    ax = axes[1]
    ax.plot(t_ax, sc, lw=0.8, color="#333333", alpha=0.85, label="MA
    ax.axhline(thresh, color="#E67E22", lw=1.2, ls="--",
               label=f"{SIGNAL_PCT}th pct threshold")
    ax.fill_between(t_ax, thresh, sc, where=sc > thresh,
                   color=PALETTE["Model"], alpha=0.18)
    ax.vlines(tau_vis, sc.min(), sc.max(),
              color=PALETTE["Model"], lw=1.0, alpha=0.75, label="Sig
    ax.vlines(sigma_vis, sc.min(), sc.max(),
              color=PALETTE["Imbalance"], lw=0.6, ls=":", alpha=0.45
              label="Stress event σ")
    ax.set_ylabel("Instability Score")
    ax.legend(loc="upper right", fontsize=8, frameon=False, ncol=2)

    axes[2].plot(t_ax, depth, lw=0.7, color="#2D6A4F"); axes[2].set_
    axes[3].plot(t_ax, imb, lw=0.7, color="#6A3D9A", alpha=0.85)
    axes[3].set_ylabel("Imbalance"); axes[3].set_xlabel("Timestep")
    fig.tight_layout()
    plt.savefig(f"fig6_composite_{method_label.lower()}.pdf", bbox_i
    plt.show()

```

```

def plot_lead_time_densities(delta_dict):
    fig, ax = plt.subplots(figsize=(9, 5))
    x_grid = np.linspace(-MAX_LAG - 5, MAX_LAG + 5, 800)
    ordered = [('Model', PALETTE['Model']),
               ('Adaptive', PALETTE['Adaptive']),
               ('Multi-Trig.', PALETTE['MultiTrig']),
               ('Imbalance', PALETTE['Imbalance']),
               ('Volatility', PALETTE['Volatility'])]
    for i, (name, color) in enumerate(ordered):
        if name not in delta_dict: continue
        deltas = delta_dict[name]
        valid = deltas[deltas > PENALTY]
        if len(valid) > 5:
            kde = gaussian_kde(valid, bw_method='scott')
            ax.plot(x_grid, kde(x_grid), lw=2.2, color=color, label=
                    ax.fill_between(x_grid, kde(x_grid), alpha=0.12, color=c
            missed = np.mean(deltas <= PENALTY)
            mean_v = np.mean(deltas[deltas > 0]) if (deltas > 0).any()
            ax.annotate(f"{name} missed={missed:.1%} E[Δ|early]={mean_v}
                        xy=(-MAX_LAG + 1, 0.006 * (i + 1)),
                        color=color, fontsize=8.5, fontweight="bold")
            ax.axvline(0, color='gray', lw=1.2, ls='--', label='Zero lead-ti
            ax.set_xlabel("Lead Time Δ (timesteps before stress event)")
            ax.set_ylabel("Density")
            ax.set_title("Lead-Time Distributions: All Detectors [v7]",
                        fontsize=13, pad=10)
            ax.legend(frameon=False, fontsize=9); ax.set_xlim(-MAX_LAG - 2,
            fig.tight_layout()
            plt.savefig("fig7_lead_time_densities.pdf", bbox_inches="tight")
            plt.show()

def plot_results_table(results_df):
    fig, ax = plt.subplots(figsize=(16, 3.0))
    ax.axis("off")
    tbl = ax.table(cellText=results_df.values, colLabels=results_df.
                  cellLoc="center", loc="center")
    tbl.auto_set_font_size(False); tbl.set_fontsize(9.2); tbl.scale(
    header_color = "#1B4F8A"
    row_colors = ["#EDF4FF", "#FFFFFF"]
    for j in range(len(results_df.columns)):
        tbl[0, j].set_facecolor(header_color)
        tbl[0, j].set_text_props(color="white", fontweight="bold")
    for i in range(1, len(results_df) + 1):
        for j in range(len(results_df.columns)):
            tbl[i, j].set_facecolor(row_colors[(i - 1) % 2])
    fig.suptitle("Detection Performance Summary – v7\n"
                "(Threshold-Based Early Detection with Coverage and
                fontsize=10, y=1.04, fontweight="bold")
    fig.tight_layout()
    plt.savefig("fig8_results_table.pdf", bbox_inches="tight")
    plt.show()

```

```

def plot_sweep_coverage_delta(sweep_std, sweep_adp, sweep_mtr):
    """
    Dual-axis plot: Mean  $\Delta$  (left) and Coverage (right) vs threshold
    Shows trade-off visually in one panel.
    """
    fig, ax1 = plt.subplots(figsize=(9, 5))
    ax2 = ax1.twinx()
    methods = [
        (sweep_std, 'Standard', PALETTE['Model'], '-'),
        (sweep_adp, 'Adaptive', PALETTE['Adaptive'], '--'),
        (sweep_mtr, 'Multi-Trig.', PALETTE['MultiTrig'], ':'),
    ]
    for df, label, color, ls in methods:
        valid = df.dropna(subset=['mean_delta', 'recall'])
        ax1.plot(valid['pct'], valid['mean_delta'], lw=2.0, color=color,
                  ls=ls, label=f"{label} - Mean  $\Delta$ ", marker='o', marke
        ax2.plot(valid['pct'], valid['recall'], lw=1.5, color=color,
                  ls='--', alpha=0.55, label=f"{label} - Coverage")
    ax1.axhline(0, color='gray', lw=0.8, ls='--')
    ax1.set_xlabel("Signal Threshold Percentile")
    ax1.set_ylabel("Mean Lead Time  $\Delta$  (steps)")
    ax2.set_ylabel("Coverage (Recall)", color='#444444')
    ax2.tick_params(axis='y', colors='#444444')
    ax1.set_title("Early Detection vs Coverage Trade-Off [v7]",
                  fontsize=12, fontweight='bold')
    lines1, labels1 = ax1.get_legend_handles_labels()
    lines2, labels2 = ax2.get_legend_handles_labels()
    ax1.legend(lines1 + lines2, labels1 + labels2, frameon=False, fo
               loc='upper left', ncol=2)
    fig.tight_layout()
    plt.savefig("fig9_sweep_coverage_delta.pdf", bbox_inches="tight")
    plt.show()

# -----
# 15. Main Pipeline
# -----
def run_experiment():
    rng = np.random.default_rng(SEED)

    print("=" * 70)
    print("LOB Micro-Regime Detection - v7")
    print("Threshold Sweep · Coverage · Precision-Recall · Robustn
    print("=" * 70)

    # — 1. Data -----
    print("\n Step 1 / 7 - Generating causal LOB data ...")
    X_raw, Z_true, delay_map = generate_lob_data(T, rng)
    rd = " | ".join([f"Regime {k}: {(Z_true==k).mean():.1%}" for k i
    print(f" {T:,} timesteps | {rd}")
    print(f" Regime-1 episodes: {len(delay_map)} | "
          f"Mean delay: {nn.mean(list(delay_map.values())):.1f} sten

```

```

mean_delay = np.mean(list(delay_map.values()), 1) # Step 1

# — 2. Features —————
print("\n Step 2 / 7 – Feature engineering ...")
X_scaled, scaler = engineer_features(X_raw)
print(f" Feature matrix: {X_scaled.shape}")

# — 3. HMM —————
print("\n Step 3 / 7 – Fitting HMM (12 restarts) ...")
model = fit_hmm(X_scaled)
ll = model.score(X_scaled); conv = model.monitor_.converged
print(f" Best log-likelihood: {ll:,.2f} | Converged: {conv}")

# — 4. Stress events —————
print("\n Step 4 / 7 – Stress event definition ...")
sigma = define_stress_events(X_raw)
print(f" Stress events: {len(sigma):,} ({len(sigma)/T:.1%}")
check_baseline_blindness(X_raw, Z_true)

# — 5. Detection (three methods) —————
print(" Step 5 / 7 – Running detection methods ...")
tau_std, score_amp, d_score, comps, post_smooth, tau_raw_std = full_pipeline(
    model, X_scaled, X_raw, sigma, signal_pct=SIGAL_PCT, method='std')
tau_adp, _, _, _, _, _ = full_pipeline(
    model, X_scaled, X_raw, sigma, signal_pct=SIGAL_PCT, method='adp')
tau_mtr, _, _, _, _, _ = full_pipeline(
    model, X_scaled, X_raw, sigma, signal_pct=SIGAL_PCT, method='mtr')
tau_imb = imbalance_baseline(X_raw)
tau_vol = volatility_baseline(X_raw)

print(f" Detections – Standard: {len(tau_std)} | Adaptive: {len(tau_adp)} |
      Multi-Trig: {len(tau_mtr)} | Imbalance: {len(tau_imb)} |
      Volatility: {len(tau_vol)}")

# — 5b. Evaluation —————
print("\n Step 5b – Evaluation ...")
det_methods = {
    'Model' : tau_std,
    'Adaptive' : tau_adp,
    'Multi-Trig.': tau_mtr,
    'Imbalance' : tau_imb,
    'Volatility' : tau_vol,
}
delta_dict = {nm: compute_lead_times(tau, sigma)
               for nm, tau in det_methods.items()}

rows = []
for name, (tau, deltas) in zip(det_methods.keys(),
                               zip(det_methods.values(), delta_dict)):
    m = evaluation_metrics(deltas, sigma, tau)
    lo, hi = bootstrap_ci(deltas)
    rows.append({

```

```

        "Detector"          : name,
        "Mean  $\Delta$ "      : f"{m['mean_delta']:+.2f}",
        "95% CI"           : f"[{lo:+.2f}, {hi:+.2f}]",
        "Precision"        : f"{m['precision']:.1%}",
        "Coverage"         : f"{m['coverage']:.1%}",
        "Mean  $\Delta$ |early"    : f"{m['mean_early']:+.2f}",
        "N( $\tau$ )"           : m['n_tau'],
        "N(early)"         : m['n_early'],
    })
results_df = pd.DataFrame(rows)
print("\n" + results_df.to_string(index=False))

print("\n Pairwise Mann-Whitney U tests (two-sided):")
pairs = [("Model", "Imbalance"), ("Model", "Volatility"),
        ("Adaptive", "Imbalance"), ("Multi-Trig.", "Imbalance"),
        ("Imbalance", "Volatility")]
for a, b in pairs:
    u, p = mannwhitney_test(delta_dict[a], delta_dict[b])
    sig = ("***" if p < 0.001 else "**" if p < 0.01 else
          "*" if p < 0.05 else "ns")
    print(f"    {a:14s} vs {b:14s}: U={u:,.0f}  p={p:.4f}  {sig}")

# — 6. Threshold sweep —————
print("\n Step 6 / 7 – Threshold sweep across methods ...")
sweep_std = threshold_sweep(model, X_scaled, X_raw, sigma,
                             pct_range=SWEEP_PCTS, method='stand
sweep_adp = threshold_sweep(model, X_scaled, X_raw, sigma,
                             pct_range=SWEEP_PCTS, method='adapt
sweep_mtr = threshold_sweep(model, X_scaled, X_raw, sigma,
                             pct_range=SWEEP_PCTS, method='multi
print(f" Sweep complete. Standard: {sweep_std['precision']}.notn
      f"valid thresholds / {len(SWEEP_PCTS)}")

# — 7. Signal diagnostics —————
print("\n Step 6b – Signal diagnostics ...")
diag_results = channel_lead_time_analysis(comps, sigma, Z_true)
winner_counts, lead_by_ch = channel_earliest_trigger_analysis(co
print(" Channel standalone performance:")
for key, res in diag_results.items():
    print(f"    {key:15s}: mean  $\Delta$ ={res['mean_delta']:+.2f}  "
          f"precision={res['precision']:.1%}  coverage={res['cov
          f"n_tau={res['n_tau']}]")
print(" Earliest-trigger counts per channel:")
for key, cnt in sorted(winner_counts.items(), key=lambda x: -x[1
    ml = np.mean(lead_by_ch[key]) if lead_by_ch[key] else 0
    print(f"    {key:15s}: {cnt} events  mean earliest lead = {m

# — 7. Robustness —————
print("\n Step 7 / 7 – Robustness grid (3 delay × 3 noise × 3 s
print(" [This may take several minutes]")
rob_df = run_robustness_grid(model, scaler, sigma, X_scaled, X_r
if not rob_df.empty:

```

```

if not rob_df.empty:
    rob_summary = (rob_df.groupby(['delay', 'noise'])
                    [['mean_delta', 'precision', 'recall']]
                    .mean().round(3))
    print("\n Robustness summary (mean over strength and reps):")
    print(rob_summary.to_string())

# — Summary —————
m_std = evaluation_metrics(delta_dict['Model'], sigma, tau_std)
print(f"\n — FINAL SUMMARY —————")
print(f" Model Mean Δ      : {m_std['mean_delta']:+.2f} steps")
print(f" Model Precision    : {m_std['precision']:.1%}")
print(f" Model Coverage      : {m_std['coverage']:.1%}")
print(f" Model Mean Δ|early: {m_std['mean_early']:+.2f} steps")
checks = [
    (m_std['mean_delta'] > 0,      "Positive mean lead-time"),
    (m_std['precision'] > 0.60,   "> 60% precision"),
    (m_std['coverage'] > 0.30,    "> 30% coverage"),
]
for ok, msg in checks:
    print(f" {'✓' if ok else 'x'} {msg}")
print()

# — Figures —————
print(" Rendering figures (9 publication-quality plots) ...")
plot_dgp_causal_structure(X_raw, Z_true, delay_map)
plot_threshold_sweep(sweep_std, sweep_adp, sweep_mtr)
plot_precision_recall(sweep_std, sweep_adp, sweep_mtr, tau_imb,
plot_robustness_heatmap(rob_df)
plot_channel_diagnostics(comps, diag_results, winner_counts, lea
plot_composite_signal(X_raw, Z_true, score_amp, tau_std, sigma,
                      method_label="Standard")
plot_lead_time_densities(delta_dict)
plot_results_table(results_df)
plot_sweep_coverage_delta(sweep_std, sweep_adp, sweep_mtr)

print("\n Experiment complete. Outputs saved as fig1_*.pdf ... fi

return dict(
    results_df      = results_df,
    delta_dict      = delta_dict,
    model           = model,
    X_raw           = X_raw,
    Z_true          = Z_true,
    sigma           = sigma,
    delay_map       = delay_map,
    score_amp       = score_amp,
    d_score         = d_score,
    comps           = comps,
    post_smooth     = post_smooth,
    tau_std         = tau_std,
    tau_adp         = tau_adp,

```

```
tau_mtr      = tau_mtr,
sweep_std    = sweep_std,
sweep_adp    = sweep_adp,
sweep_mtr    = sweep_mtr,
rob_df       = rob_df,
diag_results = diag_results,
winner_counts = winner_counts,
lead_by_ch   = lead_by_ch,
)

if __name__ == "__main__":
    results = run_experiment()
```

```
=====
LOB Micro-Regime Detection – v7
Threshold Sweep • Coverage • Precision–Recall • Robustness
=====

Step 1 / 7 – Generating causal LOB data ...
14,000 timesteps | Regime 0: 73.8% | Regime 1: 12.1% | Regime 2: 1
Regime-1 episodes: 53 | Mean delay: 32.0 steps

Step 2 / 7 – Feature engineering ...
Feature matrix: (14000, 10)

Step 3 / 7 – Fitting HMM (12 restarts) ...
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4337
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4344
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4339
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4342
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4340
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4340
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4341
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4343
WARNING:hmmlearn.base:Model is not converging. Current: 138970.4341
Best log-likelihood: 138,970.43 | Converged: True

Step 4 / 7 – Stress event definition ...
Stress events: 1,051 (7.5%)
— Baseline Blindness Sanity Check —————
Regime 0 | mean |imb| = 0.0486 (frac > thr: 0.04%) | mean rv
Regime 1 | mean |imb| = 0.1173 (frac > thr: 0.94%) | mean rv
Regime 2 | mean |imb| = 0.4240 (frac > thr: 70.09%) | mean rv

Step 5 / 7 – Running detection methods ...
Detections – Standard: 20 | Adaptive: 26 | Multi-Trig: 13 | Imbala

Step 5b – Evaluation ...

Detector Mean Δ          95% CI Precision Coverage Mean Δ|early
Model +14.95 [+12.15, +17.95] 100.0% 43.2% +14.95
Adaptive +18.62 [+14.77, +23.54] 100.0% 52.6% +18.62
```

Multi-Trig.	+13.15	[+10.23, +16.62]	100.0%	28.1%	+13.15
Imbalance	-24.84	[-30.44, -19.29]	54.9%	78.7%	+4.01
Volatility	-32.02	[-38.26, -25.77]	45.5%	43.3%	+1.55

Pairwise Mann-Whitney U tests (two-sided):

Model	vs Imbalance	: U=2,296	p=0.0000	***
Model	vs Volatility	: U=1,746	p=0.0000	***
Adaptive	vs Imbalance	: U=3,018	p=0.0000	***
Multi-Trig.	vs Imbalance	: U=1,480	p=0.0000	***
Imbalance	vs Volatility	: U=6,086	p=0.0656	ns

Step 6 / 7 – Threshold sweep across methods ...

Sweep complete. Standard: 26 valid thresholds / 26

Step 6b – Signal diagnostics ...

Channel standalone performance:

entropy	: mean Δ =+24.65	precision=100.0%	coverage=77.4%
prestress	: mean Δ =+16.52	precision=100.0%	coverage=41.2%
drift_spread	: mean Δ =+10.50	precision=100.0%	coverage=4.9%
depth_erosion	: mean Δ =+29.73	precision=100.0%	coverage=73.4%
ofi_momentum	: mean Δ =+10.00	precision=100.0%	coverage=3.4%

Earliest-trigger counts per channel:

depth_erosion	: 583 events	mean earliest lead = 47.2 steps
entropy	: 467 events	mean earliest lead = 47.6 steps
prestress	: 1 events	mean earliest lead = 60.0 steps

Step 7 / 7 – Robustness grid (3 delay × 3 noise × 3 strength × 3 r
[This may take several minutes]

```
WARNING:hmmlearn.base:Model is not converging. Current: 104000.0067
WARNING:hmmlearn.base:Model is not converging. Current: 77072.34094
WARNING:hmmlearn.base:Model is not converging. Current: 80693.85146
WARNING:hmmlearn.base:Model is not converging. Current: 80693.84973
WARNING:hmmlearn.base:Model is not converging. Current: 83606.27513
WARNING:hmmlearn.base:Model is not converging. Current: 83678.74791
WARNING:hmmlearn.base:Model is not converging. Current: 79637.60451
WARNING:hmmlearn.base:Model is not converging. Current: 99922.84418
WARNING:hmmlearn.base:Model is not converging. Current: 99922.62152
WARNING:hmmlearn.base:Model is not converging. Current: 99922.64471
WARNING:hmmlearn.base:Model is not converging. Current: 99923.06350
WARNING:hmmlearn.base:Model is not converging. Current: 99928.49998
WARNING:hmmlearn.base:Model is not converging. Current: 99923.19865
WARNING:hmmlearn.base:Model is not converging. Current: 107719.7734
WARNING:hmmlearn.base:Model is not converging. Current: 107719.7733
WARNING:hmmlearn.base:Model is not converging. Current: 107706.9681
WARNING:hmmlearn.base:Model is not converging. Current: 107707.7691
WARNING:hmmlearn.base:Model is not converging. Current: 107719.7734
WARNING:hmmlearn.base:Model is not converging. Current: 107708.5919
WARNING:hmmlearn.base:Model is not converging. Current: 102981.1029
WARNING:hmmlearn.base:Model is not converging. Current: 102981.6690
WARNING:hmmlearn.base:Model is not converging. Current: 102981.3576
WARNING:hmmlearn.base:Model is not converging. Current: 102977.2491
WARNING:hmmlearn.base:Model is not converging. Current: 102977.2491
WARNING:hmmlearn.base:Model is not converging. Current: 102981.9447
WARNING:hmmlearn.base:Model is not converging. Current: 96036.29695
WARNING:hmmlearn.base:Model is not converging. Current: 96035.05714
WARNING:hmmlearn.base:Model is not converging. Current: 96036.76781
WARNING:hmmlearn.base:Model is not converging. Current: 96035.49536
```

WARNING:hmmlearn.base:Model is not converging. Current: 105540.1326
WARNING:hmmlearn.base:Model is not converging. Current: 105540.6391
WARNING:hmmlearn.base:Model is not converging. Current: 105540.7179
WARNING:hmmlearn.base:Model is not converging. Current: 105540.1176
WARNING:hmmlearn.base:Model is not converging. Current: 105540.7065
WARNING:hmmlearn.base:Model is not converging. Current: 108018.7154
WARNING:hmmlearn.base:Model is not converging. Current: 107977.0345
WARNING:hmmlearn.base:Model is not converging. Current: 107979.5402
WARNING:hmmlearn.base:Model is not converging. Current: 108018.7152
WARNING:hmmlearn.base:Model is not converging. Current: 108018.7154
WARNING:hmmlearn.base:Model is not converging. Current: 107979.1243
WARNING:hmmlearn.base:Model is not converging. Current: 106046.0849
WARNING:hmmlearn.base:Model is not converging. Current: 79282.36410
WARNING:hmmlearn.base:Model is not converging. Current: 106266.5900
WARNING:hmmlearn.base:Model is not converging. Current: 79282.34898
WARNING:hmmlearn.base:Model is not converging. Current: 128618.7501
WARNING:hmmlearn.base:Model is not converging. Current: 128618.7510
WARNING:hmmlearn.base:Model is not converging. Current: 128618.7490
WARNING:hmmlearn.base:Model is not converging. Current: 100397.4635
WARNING:hmmlearn.base:Model is not converging. Current: 128618.7448
WARNING:hmmlearn.base:Model is not converging. Current: 128618.7509
WARNING:hmmlearn.base:Model is not converging. Current: 119238.6597
WARNING:hmmlearn.base:Model is not converging. Current: 146969.1902
WARNING:hmmlearn.base:Model is not converging. Current: 121313.0028
WARNING:hmmlearn.base:Model is not converging. Current: 129430.2804
WARNING:hmmlearn.base:Model is not converging. Current: 121312.9830
WARNING:hmmlearn.base:Model is not converging. Current: 121313.0845
WARNING:hmmlearn.base:Model is not converging. Current: 140656.3822
WARNING:hmmlearn.base:Model is not converging. Current: 140407.0913
WARNING:hmmlearn.base:Model is not converging. Current: 140407.0513
WARNING:hmmlearn.base:Model is not converging. Current: 140654.5455
WARNING:hmmlearn.base:Model is not converging. Current: 140652.1728
WARNING:hmmlearn.base:Model is not converging. Current: 140407.1188
WARNING:hmmlearn.base:Model is not converging. Current: 148110.2047
WARNING:hmmlearn.base:Model is not converging. Current: 148110.2047
WARNING:hmmlearn.base:Model is not converging. Current: 148110.2043
WARNING:hmmlearn.base:Model is not converging. Current: 94388.00070
WARNING:hmmlearn.base:Model is not converging. Current: 148142.6727
WARNING:hmmlearn.base:Model is not converging. Current: 148110.2047
WARNING:hmmlearn.base:Model is not converging. Current: 124573.1144
WARNING:hmmlearn.base:Model is not converging. Current: 130276.2835
WARNING:hmmlearn.base:Model is not converging. Current: 130274.9859
WARNING:hmmlearn.base:Model is not converging. Current: 130276.2835
WARNING:hmmlearn.base:Model is not converging. Current: 130276.2840
WARNING:hmmlearn.base:Model is not converging. Current: 152468.3220
WARNING:hmmlearn.base:Model is not converging. Current: 152468.3223
WARNING:hmmlearn.base:Model is not converging. Current: 152468.3216
WARNING:hmmlearn.base:Model is not converging. Current: 152494.0584
WARNING:hmmlearn.base:Model is not converging. Current: 152494.0568
WARNING:hmmlearn.base:Model is not converging. Current: 152468.3222
WARNING:hmmlearn.base:Model is not converging. Current: 143948.5860
WARNING:hmmlearn.base:Model is not converging. Current: 143948.4320
WARNING:hmmlearn.base:Model is not converging. Current: 143948.4695
WARNING:hmmlearn.base:Model is not converging. Current: 143948.5865
WARNING:hmmlearn.base:Model is not converging. Current: 90195.67117
WARNING:hmmlearn.base:Model is not converging. Current: 143948.4687
WARNING:hmmlearn.base:Model is not converging. Current: 162391.4258
WARNING:hmmlearn.base:Model is not converging. Current: 121037.8544

WARNING:hmmlearn.base:Model is not converging. Current: 121067.3278
WARNING:hmmlearn.base:Model is not converging. Current: 148829.2751
WARNING:hmmlearn.base:Model is not converging. Current: 121048.8812
WARNING:hmmlearn.base:Model is not converging. Current: 211697.1890
WARNING:hmmlearn.base:Model is not converging. Current: 211697.1889
WARNING:hmmlearn.base:Model is not converging. Current: 211697.1889
WARNING:hmmlearn.base:Model is not converging. Current: 164179.6602
WARNING:hmmlearn.base:Some rows of transmat_ have zero sum because n
WARNING:hmmlearn.base:Some rows of transmat_ have zero sum because n
WARNING:hmmlearn.base:Model is not converging. Current: 149990.3761
WARNING:hmmlearn.base:Model is not converging. Current: 151871.5969
WARNING:hmmlearn.base:Model is not converging. Current: 153532.7390
WARNING:hmmlearn.base:Model is not converging. Current: 101672.6750
WARNING:hmmlearn.base:Model is not converging. Current: 179470.3710
WARNING:hmmlearn.base:Model is not converging. Current: 181064.6747
WARNING:hmmlearn.base:Model is not converging. Current: 204866.0101
WARNING:hmmlearn.base:Model is not converging. Current: 204851.2278
WARNING:hmmlearn.base:Model is not converging. Current: 205257.7280
WARNING:hmmlearn.base:Model is not converging. Current: 181064.6679
WARNING:hmmlearn.base:Model is not converging. Current: 235882.9449
WARNING:hmmlearn.base:Model is not converging. Current: 235882.9451
WARNING:hmmlearn.base:Model is not converging. Current: 235882.9449
WARNING:hmmlearn.base:Model is not converging. Current: 165816.2216
WARNING:hmmlearn.base:Model is not converging. Current: 235882.9449
WARNING:hmmlearn.base:Model is not converging. Current: 235882.9449
WARNING:hmmlearn.base:Model is not converging. Current: 187599.3885
WARNING:hmmlearn.base:Model is not converging. Current: 187599.4705
WARNING:hmmlearn.base:Model is not converging. Current: 102989.2978
WARNING:hmmlearn.base:Model is not converging. Current: 103055.7731
WARNING:hmmlearn.base:Model is not converging. Current: 187599.4492
WARNING:hmmlearn.base:Model is not converging. Current: 21145.53184
WARNING:hmmlearn.base:Model is not converging. Current: 198511.0179
WARNING:hmmlearn.base:Model is not converging. Current: 198510.8736
WARNING:hmmlearn.base:Model is not converging. Current: 195094.6404
WARNING:hmmlearn.base:Model is not converging. Current: 198510.8787
WARNING:hmmlearn.base:Model is not converging. Current: 198511.0608
WARNING:hmmlearn.base:Model is not converging. Current: 229268.7192
WARNING:hmmlearn.base:Model is not converging. Current: 199036.6094
WARNING:hmmlearn.base:Model is not converging. Current: 199036.6093
WARNING:hmmlearn.base:Model is not converging. Current: 229268.7191
WARNING:hmmlearn.base:Model is not converging. Current: 199169.2500
WARNING:hmmlearn.base:Model is not converging. Current: 199036.6092
WARNING:hmmlearn.base:Model is not converging. Current: 202803.5616
WARNING:hmmlearn.base:Model is not converging. Current: 202803.5114
WARNING:hmmlearn.base:Model is not converging. Current: 71925.35344
WARNING:hmmlearn.base:Model is not converging. Current: 202803.5023
WARNING:hmmlearn.base:Model is not converging. Current: 79512.98121
WARNING:hmmlearn.base:Model is not converging. Current: 79512.98200
WARNING:hmmlearn.base:Model is not converging. Current: 79512.98270
WARNING:hmmlearn.base:Model is not converging. Current: 79561.12539
WARNING:hmmlearn.base:Model is not converging. Current: 79512.98314
WARNING:hmmlearn.base:Model is not converging. Current: 72722.04691
WARNING:hmmlearn.base:Model is not converging. Current: 72722.04609
WARNING:hmmlearn.base:Model is not converging. Current: 72722.04549
WARNING:hmmlearn.base:Model is not converging. Current: 72754.53473
WARNING:hmmlearn.base:Model is not converging. Current: 72722.04682
WARNING:hmmlearn.base:Model is not converging. Current: 72722.04645

[illegible]

WARNING:hmmlearn.base:Model is not converging. Current: 116713.2373

WARNING:hmmlearn.base:Model is not converging. Current: 88080.62420

WARNING:hmmlearn.base:Model is not converging. Current: 143002.5015

WARNING:hmmlearn.base:Model is not converging. Current: 143002.5015

WARNING:hmmlearn.base:Model is not converging. Current: 143002.5012

WARNING:hmmlearn.base:Model is not converging. Current: 143002.5015

WARNING:hmmlearn.base:Model is not converging. Current: 116713.2373

WARNING:hmmlearn.base:Model is not converging. Current: 141190.6780

WARNING:hmmlearn.base:Model is not converging. Current: 141190.6782

WARNING:hmmlearn.base:Model is not converging. Current: 144787.6003

WARNING:hmmlearn.base:Model is not converging. Current: 141190.6781

WARNING:hmmlearn.base:Model is not converging. Current: 141190.6779

WARNING:hmmlearn.base:Model is not converging. Current: 179244.0395

WARNING:hmmlearn.base:Model is not converging. Current: 179244.0401

WARNING:hmmlearn.base:Model is not converging. Current: 179244.0398

WARNING:hmmlearn.base:Model is not converging. Current: 179244.0395

WARNING:hmmlearn.base:Model is not converging. Current: 179244.0400

WARNING:hmmlearn.base:Model is not converging. Current: 179244.0395

WARNING:hmmlearn.base:Model is not converging. Current: 136212.4849

WARNING:hmmlearn.base:Model is not converging. Current: 136181.4199

WARNING:hmmlearn.base:Model is not converging. Current: 136181.4198

WARNING:hmmlearn.base:Model is not converging. Current: 110570.9698

WARNING:hmmlearn.base:Model is not converging. Current: 136212.4851

WARNING:hmmlearn.base:Model is not converging. Current: 136181.4198

WARNING:hmmlearn.base:Model is not converging. Current: 142179.4625

WARNING:hmmlearn.base:Model is not converging. Current: 142179.4622

WARNING:hmmlearn.base:Model is not converging. Current: 142179.4626

WARNING:hmmlearn.base:Model is not converging. Current: 142179.4619

WARNING:hmmlearn.base:Model is not converging. Current: 117199.4430

WARNING:hmmlearn.base:Model is not converging. Current: 142179.4624

WARNING:hmmlearn.base:Model is not converging. Current: 154219.0985

WARNING:hmmlearn.base:Model is not converging. Current: 161338.7509

WARNING:hmmlearn.base:Model is not converging. Current: 178570.5892

WARNING:hmmlearn.base:Model is not converging. Current: 178570.5893

WARNING:hmmlearn.base:Model is not converging. Current: 120869.9582

WARNING:hmmlearn.base:Model is not converging. Current: 98120.76731

WARNING:hmmlearn.base:Model is not converging. Current: 177797.5933

WARNING:hmmlearn.base:Model is not converging. Current: 98120.77858

WARNING:hmmlearn.base:Model is not converging. Current: 98120.80028

WARNING:hmmlearn.base:Model is not converging. Current: 177797.4996

WARNING:hmmlearn.base:Model is not converging. Current: 177797.5871

WARNING:hmmlearn.base:Model is not converging. Current: 85931.63397

WARNING:hmmlearn.base:Model is not converging. Current: 161430.6324

WARNING:hmmlearn.base:Model is not converging. Current: 161430.6324

WARNING:hmmlearn.base:Model is not converging. Current: 85933.77532

WARNING:hmmlearn.base:Model is not converging. Current: 161430.6326

WARNING:hmmlearn.base:Some rows of transmat_ have zero sum because n

WARNING:hmmlearn.base:Model is not converging. Current: 149208.1398

WARNING:hmmlearn.base:Model is not converging. Current: 145913.7012

WARNING:hmmlearn.base:Model is not converging. Current: 171431.5595

WARNING:hmmlearn.base:Model is not converging. Current: 199089.6109

WARNING:hmmlearn.base:Model is not converging. Current: 199089.6108

WARNING:hmmlearn.base:Model is not converging. Current: 199089.6113

WARNING:hmmlearn.base:Model is not converging. Current: 178469.9897

WARNING:hmmlearn.base:Model is not converging. Current: 199089.6103

WARNING:hmmlearn.base:Model is not converging. Current: 100233.7503

WARNING:hmmlearn.base:Model is not converging. Current: 171238.7950

WARNING:hmmlearn.base:Model is not converging. Current: 171238.7903

WARNING:hmmlearn.base:Model is not converging. Current: 100233.9748
WARNING:hmmlearn.base:Model is not converging. Current: 171238.7923
WARNING:hmmlearn.base:Model is not converging. Current: 158429.9239
WARNING:hmmlearn.base:Model is not converging. Current: 205591.2242
WARNING:hmmlearn.base:Model is not converging. Current: 205591.2239
WARNING:hmmlearn.base:Model is not converging. Current: 205591.2243
WARNING:hmmlearn.base:Model is not converging. Current: 205591.2237
WARNING:hmmlearn.base:Model is not converging. Current: 85445.16572
WARNING:hmmlearn.base:Model is not converging. Current: 85445.16590
WARNING:hmmlearn.base:Model is not converging. Current: 85445.16417
WARNING:hmmlearn.base:Model is not converging. Current: 85445.16541
WARNING:hmmlearn.base:Model is not converging. Current: 85445.16592
WARNING:hmmlearn.base:Model is not converging. Current: 85445.16469
WARNING:hmmlearn.base:Model is not converging. Current: 41374.57199
WARNING:hmmlearn.base:Model is not converging. Current: 70783.25941
WARNING:hmmlearn.base:Model is not converging. Current: 70783.25949
WARNING:hmmlearn.base:Model is not converging. Current: 81394.33530
WARNING:hmmlearn.base:Model is not converging. Current: 83806.08937
WARNING:hmmlearn.base:Model is not converging. Current: 55350.85937
WARNING:hmmlearn.base:Model is not converging. Current: 81635.70189
WARNING:hmmlearn.base:Model is not converging. Current: 81635.70265
WARNING:hmmlearn.base:Model is not converging. Current: 81635.70213
WARNING:hmmlearn.base:Model is not converging. Current: 66508.51873
WARNING:hmmlearn.base:Model is not converging. Current: 81635.70206
WARNING:hmmlearn.base:Model is not converging. Current: 53293.28896
WARNING:hmmlearn.base:Model is not converging. Current: 79043.04878
WARNING:hmmlearn.base:Model is not converging. Current: 79043.04779
WARNING:hmmlearn.base:Model is not converging. Current: 79043.04854
WARNING:hmmlearn.base:Model is not converging. Current: 79043.04878
WARNING:hmmlearn.base:Model is not converging. Current: 79043.04815
WARNING:hmmlearn.base:Model is not converging. Current: 85331.33039
WARNING:hmmlearn.base:Model is not converging. Current: 86079.75983
WARNING:hmmlearn.base:Model is not converging. Current: 85331.33039
WARNING:hmmlearn.base:Model is not converging. Current: 86079.75942
WARNING:hmmlearn.base:Model is not converging. Current: 86079.75997
WARNING:hmmlearn.base:Model is not converging. Current: 87105.95881
WARNING:hmmlearn.base:Model is not converging. Current: 87105.95842
WARNING:hmmlearn.base:Model is not converging. Current: 64119.35330
WARNING:hmmlearn.base:Model is not converging. Current: 105624.1536
WARNING:hmmlearn.base:Model is not converging. Current: 89225.72759
WARNING:hmmlearn.base:Model is not converging. Current: 89223.97445
WARNING:hmmlearn.base:Model is not converging. Current: 89225.73047
WARNING:hmmlearn.base:Model is not converging. Current: 89223.97466
WARNING:hmmlearn.base:Model is not converging. Current: 89225.72786
WARNING:hmmlearn.base:Model is not converging. Current: 89223.97329
WARNING:hmmlearn.base:Model is not converging. Current: 87499.66661
WARNING:hmmlearn.base:Model is not converging. Current: 87499.66658
WARNING:hmmlearn.base:Model is not converging. Current: 87499.66679
WARNING:hmmlearn.base:Model is not converging. Current: 87501.50689
WARNING:hmmlearn.base:Model is not converging. Current: 87501.50706
WARNING:hmmlearn.base:Model is not converging. Current: 87499.66651
WARNING:hmmlearn.base:Model is not converging. Current: 91307.22256
WARNING:hmmlearn.base:Model is not converging. Current: 91307.22257
WARNING:hmmlearn.base:Model is not converging. Current: 91307.22256
WARNING:hmmlearn.base:Model is not converging. Current: 44271.26631
WARNING:hmmlearn.base:Model is not converging. Current: 92150.50025
WARNING:hmmlearn.base:Model is not converging. Current: 91307.22256
WARNING:hmmlearn.base:Model is not converging. Current: 115532.8094

[illegible]

[illegible]

Robustness summary (mean over strength and reps):

		mean_delta	precision	recall
delay	noise			
Default	High	24.950	0.333	0.046
	Low	20.476	1.000	0.411
	Medium	18.640	1.000	0.402
Long	High	19.581	0.556	0.217
	Low	23.237	1.000	0.747
	Medium	19.073	0.889	0.585
Short	High	11.444	0.333	0.033
	Low	22.745	1.000	0.533

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

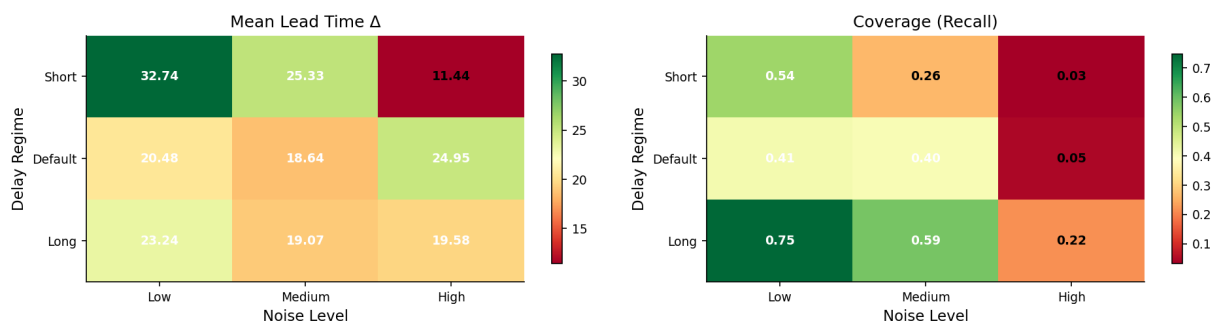
[illegible]

[illegible]

[illegible]

[illegible]

Robustness: Mean Lead Time Δ and Coverage across Delay Regime $\times$ Noise Level [v7]



```
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
```

[illegible]

[illegible]

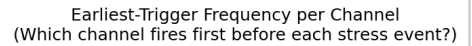
[illegible]

[illegible]

[illegible]

[illegible]

WARNING:malicious.tone_manager.tttttttt. genetic family still not

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Detection Performance Summary – v7
(Threshold-Based Early Detection with Coverage and Precision-Recall)

Detector	Mean Δ	95% CI	Precision	Coverage	Mean Δ (early)	N(τ)	N(early)
Model	+14.95	[+12.15, +17.95]	100.0%	43.2%	+14.95	20	20
Adaptive	+18.62	[+14.77, +23.54]	100.0%	52.6%	+18.62	26	26
Multi-Trig.	+13.15	[+10.23, +16.62]	100.0%	28.1%	+13.15	13	13
Imbalance	-24.84	[-30.44, -19.29]	54.9%	78.7%	+4.01	122	67
Volatility	-32.02	[-38.26, -25.77]	45.5%	43.3%	+1.55	88	40

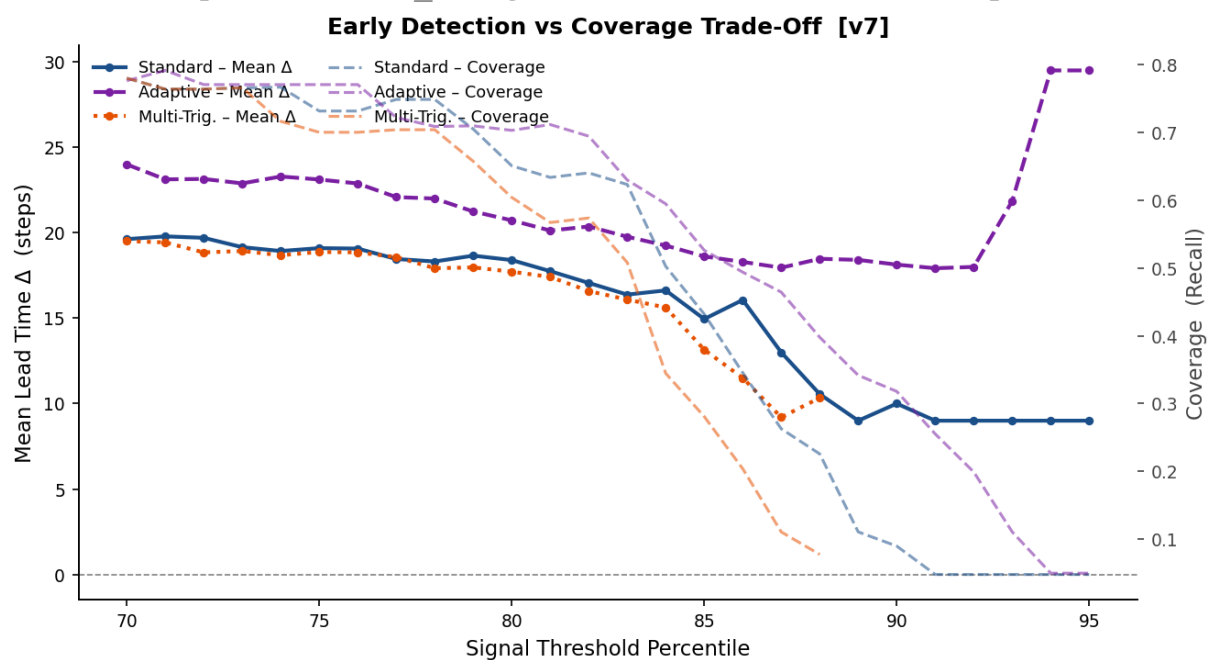
[illegible]

[illegible]

[illegible]

[illegible]

WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not
WARNING:matplotlib.font_manager.findfont: Generic family 'serif' not



Experiment complete. Outputs saved as fig1_*.pdf ... fig9_*.pdf

