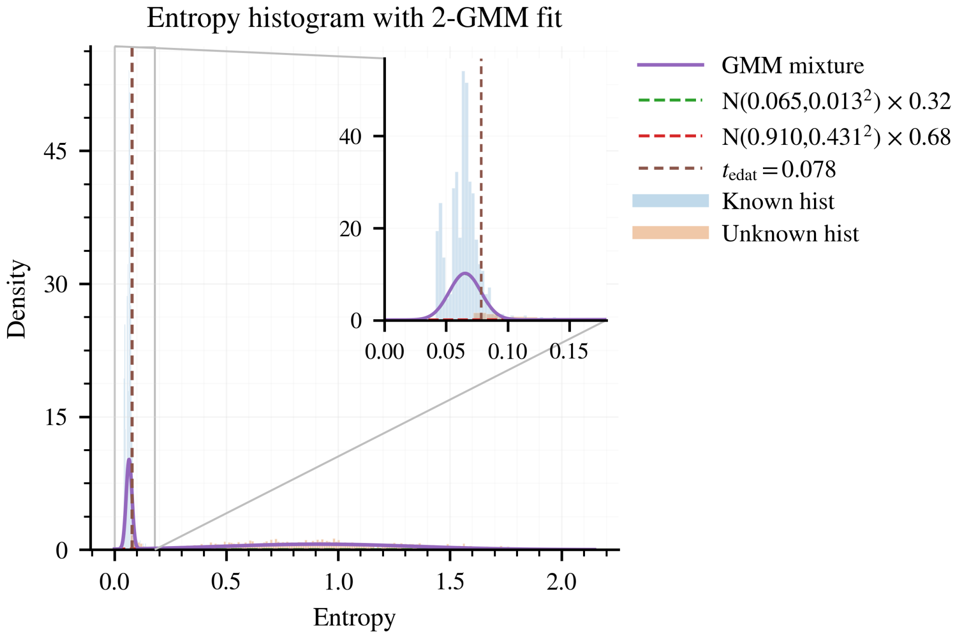


Fig. 1: Entropy histogram with 2-GMM.

A. Visual fit of the mixture.

The histogram with the fitted 2-GMM (Fig. 1) shows two modes with parameters

$$\widehat{\pi}_1 = 0.32, \quad \widehat{\mu}_1 = 0.065, \quad \widehat{\sigma}_1 = 0.013; \qquad \widehat{\pi}_2 = 0.68, \quad \widehat{\mu}_2 = 0.910, \quad \widehat{\sigma}_2 = 0.431.$$

We mark the adaptive decision threshold $t_{\mathrm{EDAT}} = 0.078$ on the plot; it lies near the density valley between the modes.

B. Model selection.

We keep the main text concise and report the supporting details here. Let $\{h_i\}_{i=1}^{n}$ be the entropies. We fit a K-component Gaussian mixture

$$p(h) = \sum_{j=1}^{K} \pi_j \, \mathcal{N}\big(h \mid \mu_j, \sigma_j^2\big), \qquad \pi_j \geq 0, \ \sum_j \pi_j = 1,$$

by maximum likelihood (EM). Information criteria are computed as

$$\mathrm{AIC} = 2p - 2\ln\hat{\mathcal{L}}, \qquad \mathrm{BIC} = p\ln n - 2\ln\hat{\mathcal{L}},$$

where $\hat{\mathcal{L}}$ is the maximized likelihood and p is the number of free parameters in the K-GMM. For univariate mixtures with component-specific variances, $p = (K-1)$ (weights) $+ K$ (means) $+ K$ (variances) $= 3K - 1$. Both criteria are "smaller-is-better." AIC estimates out-of-sample prediction error (via an unbiased estimate of expected Kullback–Leibler risk), while BIC approximates the log model evidence with a stronger complexity penalty.

Table I: Model selection for Gaussian mixtures on entropy (smaller is better).

K	$p = 3K - 1$	AIC	BIC
1	2	5432.524	5444.820
2	5	**147.918**	**197.101**
3	8	236.291	267.030

The large drops from $K = 1$ to $K = 2$ and the subsequent increases at $K = 3$ indicate that a two-component mixture is the most parsimonious and best-supported specification under both AIC and BIC.

C. Q–Q diagnostics.

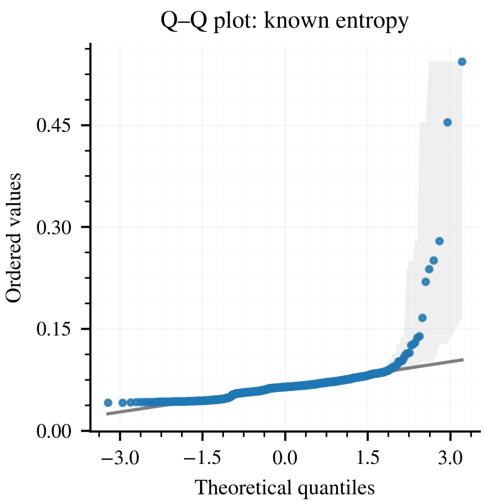
A Q–Q plot maps empirical sample quantiles to the theoretical quantiles of a reference distribution. Let $y_{(1)} \le \cdots \le y_{(n)}$ be the ordered entropy values and $p_i = (i - 0.5)/n$ the plotting positions. For a fitted normal $\mathcal{N}(\mu, \sigma^2)$, define the theoretical quantiles $q_i = \mu + \sigma\, \Phi^{-1}(p_i)$, where Φ^{-1} is the inverse CDF of the standard normal. If the sample is approximately Gaussian up to location/scale, the points $\{(q_i,\ y_{(i)})\}$ fall close to a straight line; equivalently, after standardization $z_{(i)} = (y_{(i)} - \mu)/\sigma$, the Q–Q reference line is $z_{(i)} \approx \Phi^{-1}(p_i)$ (slope ≈ 1, intercept ≈ 0). Systematic curvature at upper quantiles indicates heavier-than-normal right tails, while an S-shape indicates skewness.

Known–class entropy quantiles align with the normal reference line in the central region with a mild right-tail deviation; unknown–class entropy exhibits a heavier upper tail. These patterns are consistent with a finite mixture rather than a single Gaussian for the *pooled* entropy distribution.

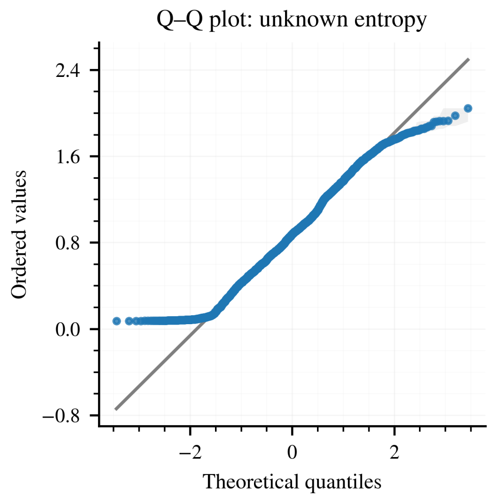
Fitted mixture. We model the overall entropy density by a two-component Gaussian mixture

$$f(y) = \sum_{k=1}^{2} \pi_k \, \frac{1}{\sigma_k} \, \phi\!\left(\frac{y - \mu_k}{\sigma_k}\right),$$

where ϕ is the standard normal pdf and π_k are the mixing weights. The fitted components are $\mathcal{N}(0.065,\ 0.013^2)$ with weight $\pi_1 = 0.32$ and $\mathcal{N}(0.910,\ 0.431^2)$ with weight $\pi_2 = 0.68$. Our EDAT threshold is $t_{\mathrm{EDAT}} = 0.078$, lying near the inter–modal valley, which aligns with the Q–Q diagnostics.



(a) Known-class Q–Q.



(b) Unknown-class Q–Q.

Fig. 2: Q–Q diagnostics for entropy. Points lying along the 45° reference line indicate approximate normality (up to location/scale); tail curvature reveals heavier tails.