Noise-Resistant Training for Sequence-to-Expression Models

Systematic Evaluation on CAGI5 Clinical Variant Benchmark

74	**+1.2%**	**-0.088**
Models	Best CAGI5 Improvement	Best Noise Correlation

Rank Stability

Distributional

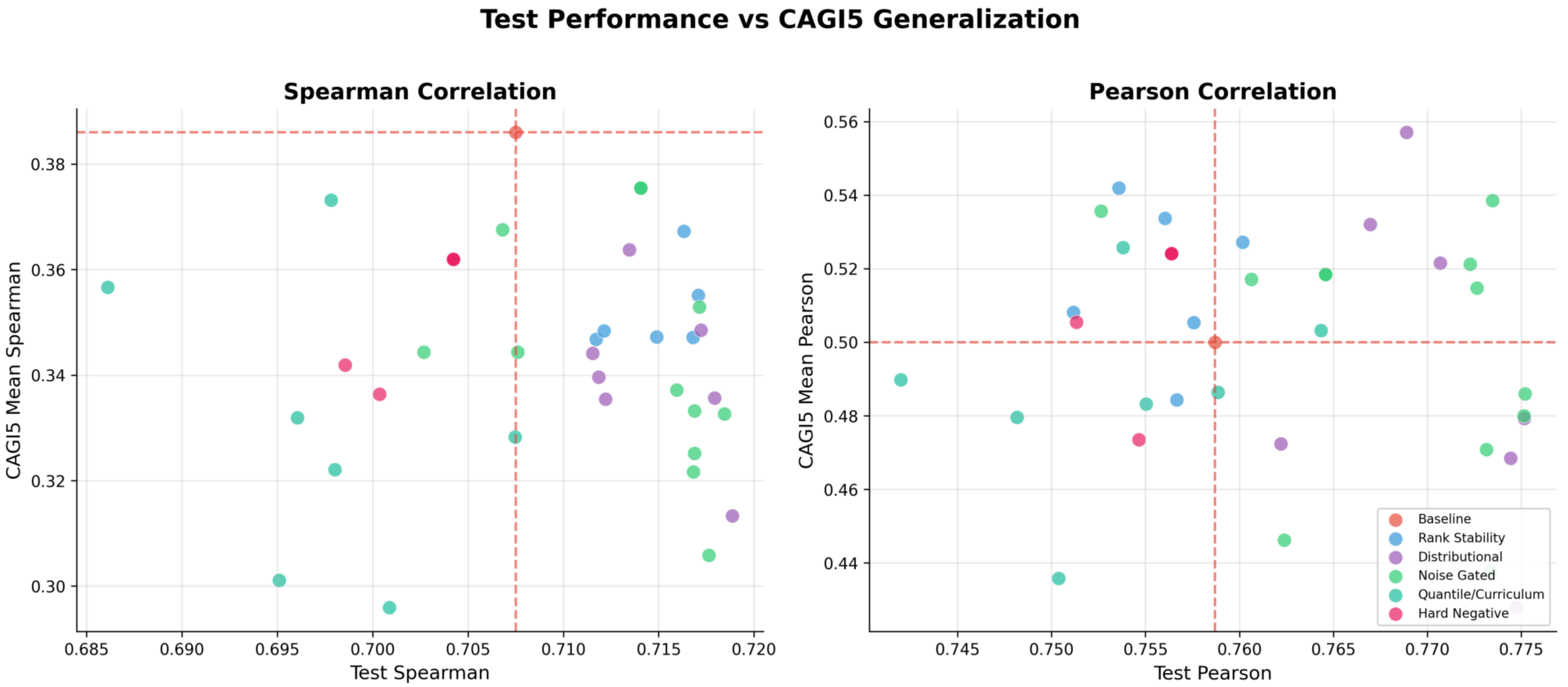
Noise Gated

Contrastive

Quantile/Curriculum

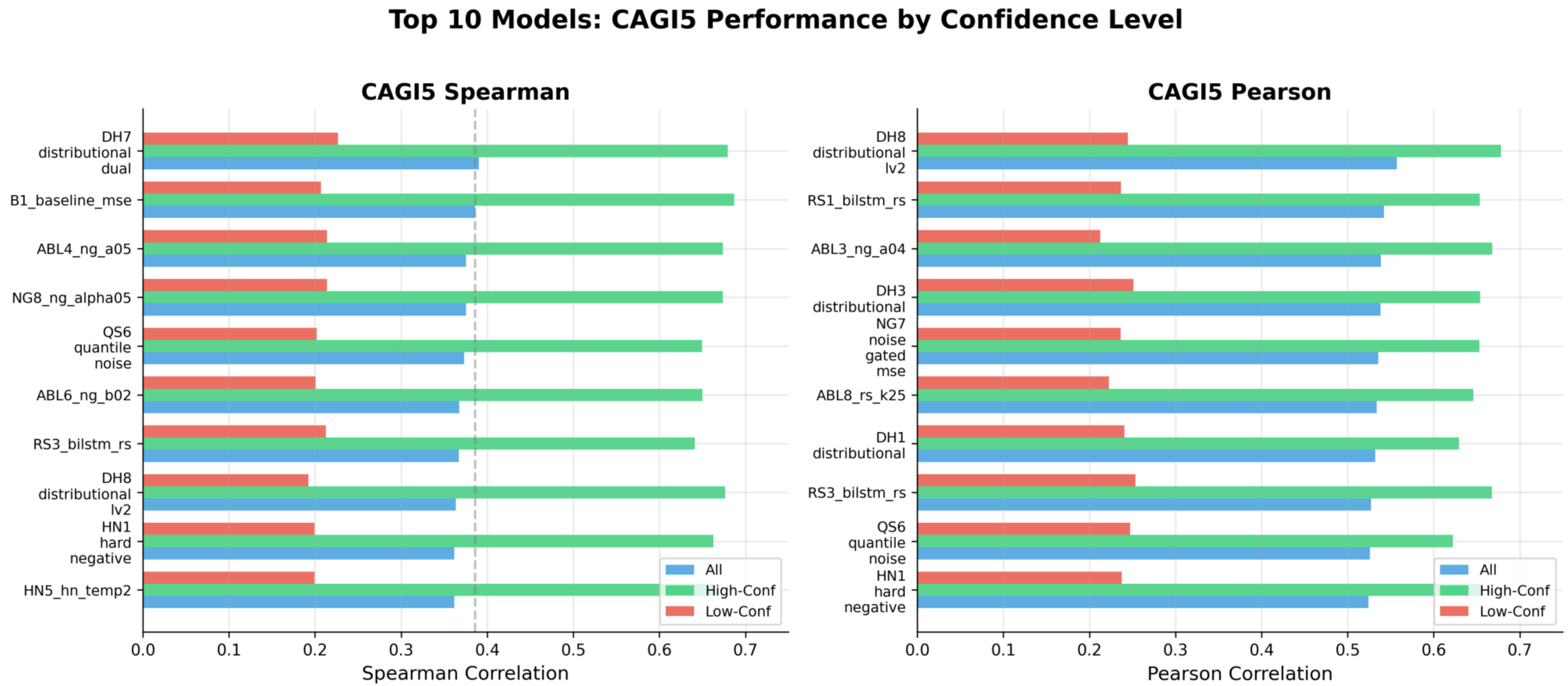
Hard Negative

Figure 1: Test Performance vs CAGI5 Generalization



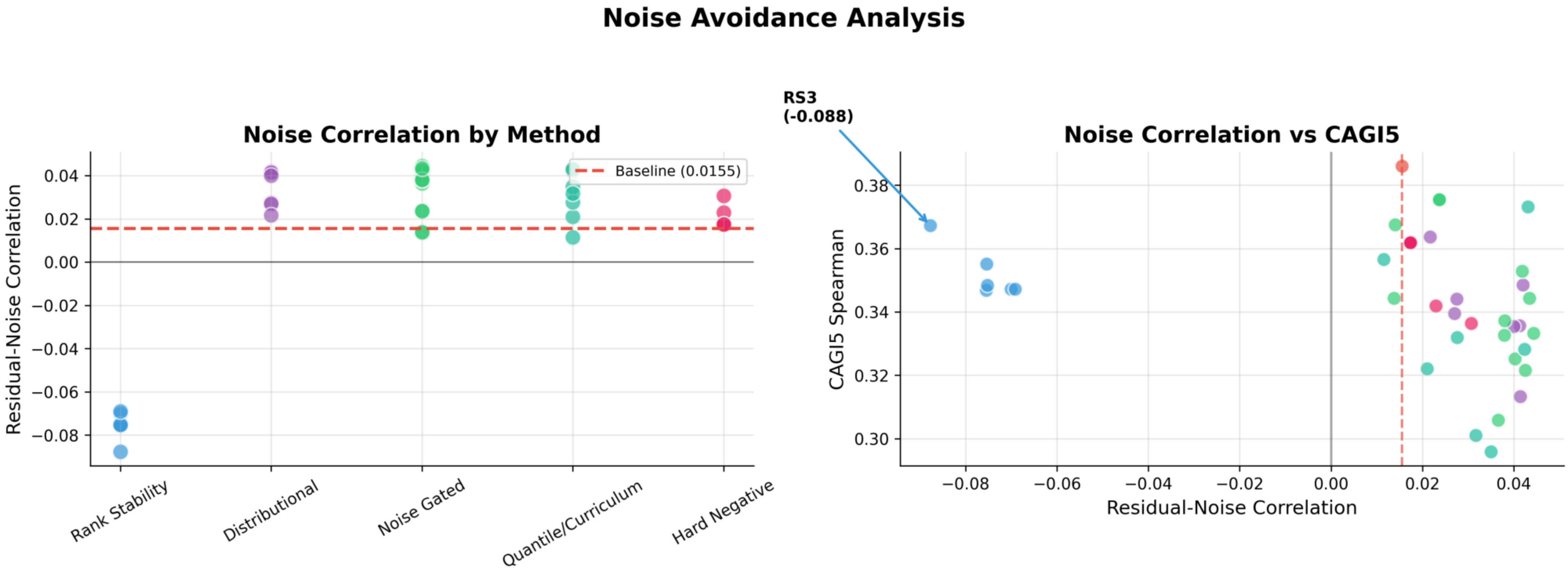
Both Spearman (left) and Pearson (right) correlations shown. Dashed red lines indicate baseline.
Key Finding: Higher test performance does not guarantee better CAGI5 generalization.

Figure 2: CAGI5 by Confidence Level



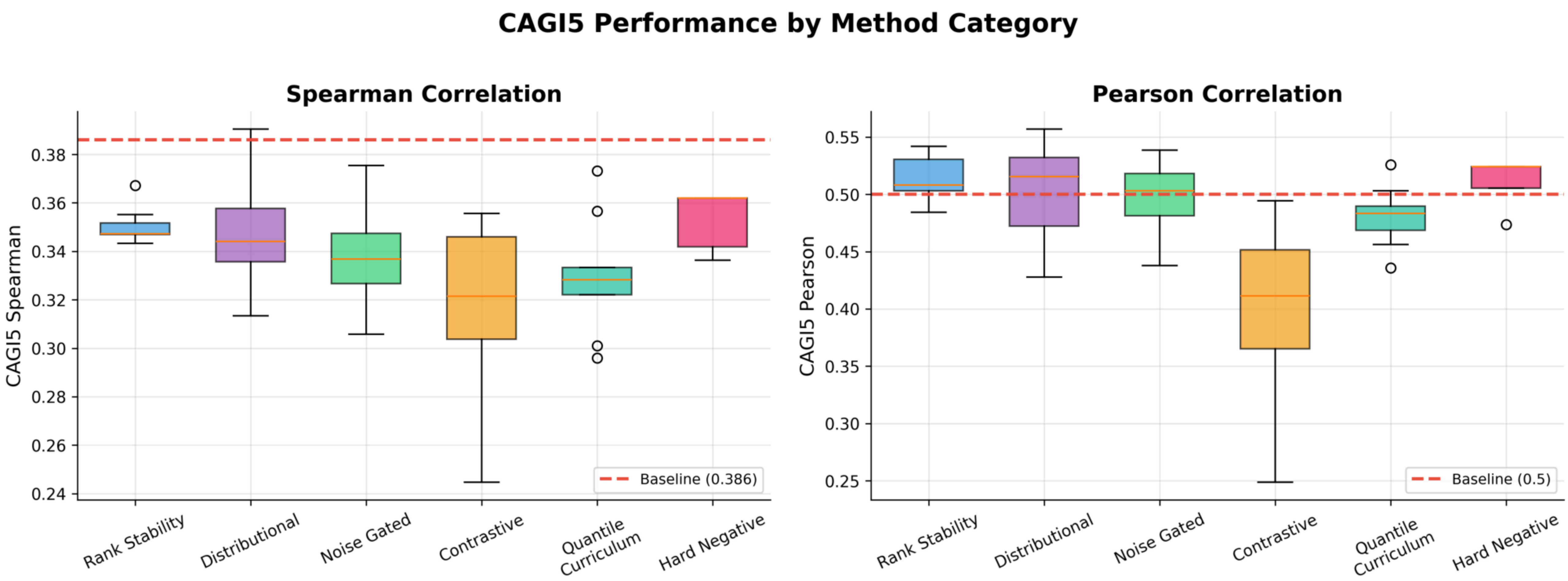
Top 10 models ranked by Spearman (left) and Pearson (right). Bars show All/High-Conf/Low-Conf performance.
Key Finding: DH7 achieves best overall with +10% improvement on low-confidence variants.

Figure 3: Noise Avoidance Analysis



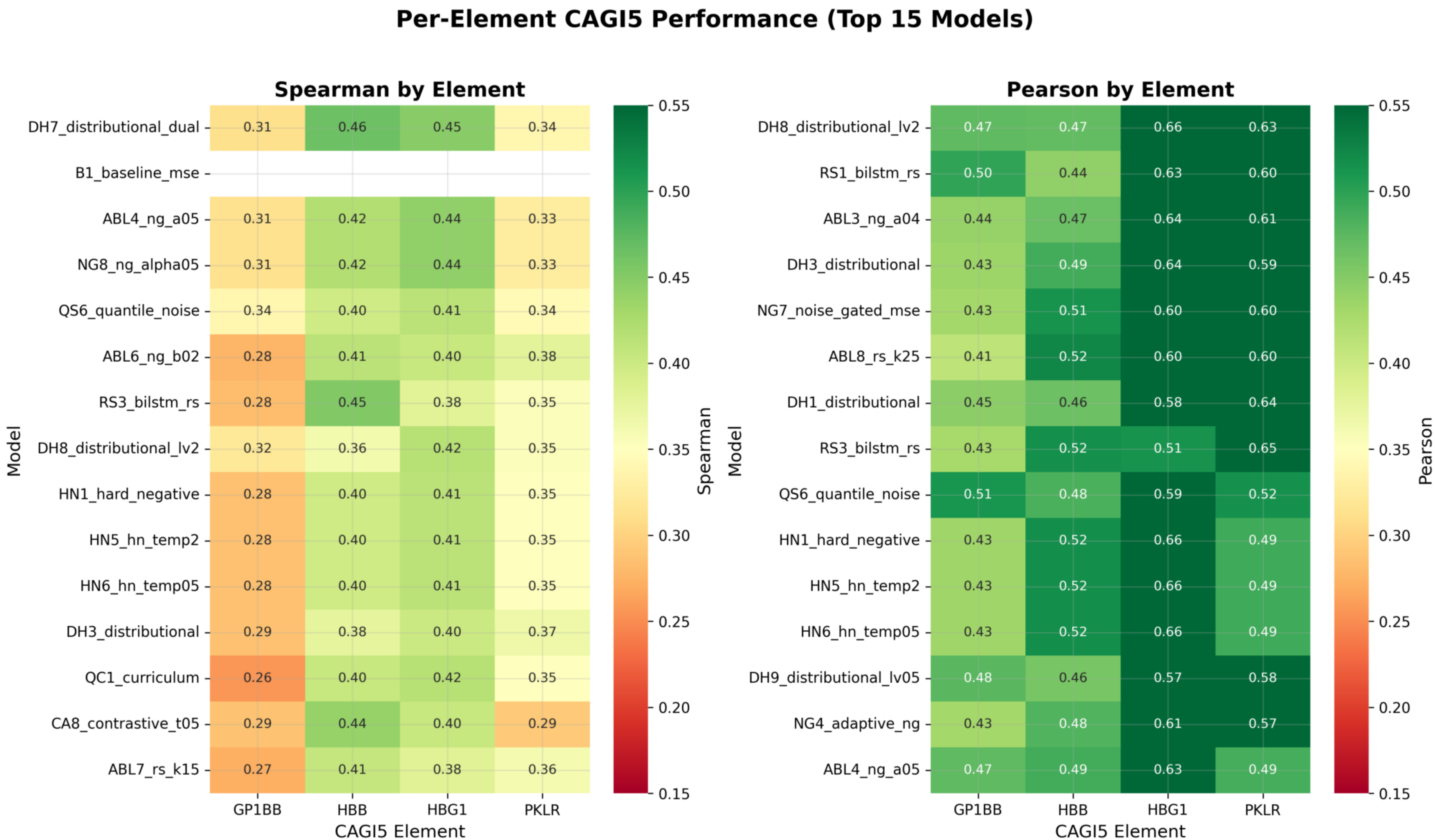
Left: Distribution by method. Right: Noise correlation vs CAGI5 Spearman.
Key Finding: RS3 is the ONLY model with negative noise correlation (-0.088).

Figure 4: Method Category Comparison



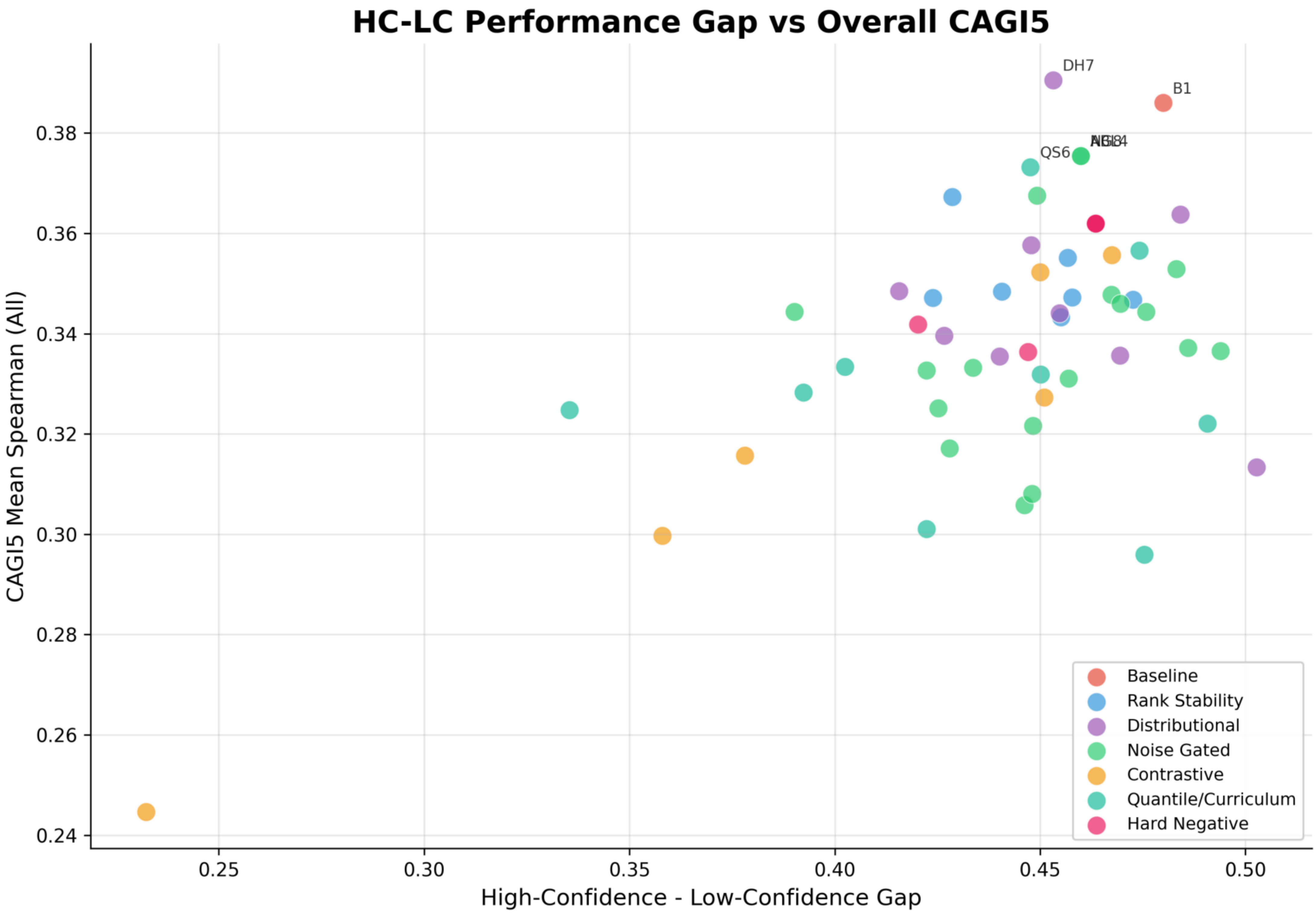
Box plots for both Spearman (left) and Pearson (right). Red line = baseline.
Key Finding: Distributional methods achieve highest and most consistent improvements.

Figure 5: Per-Element CAGI5 Performance



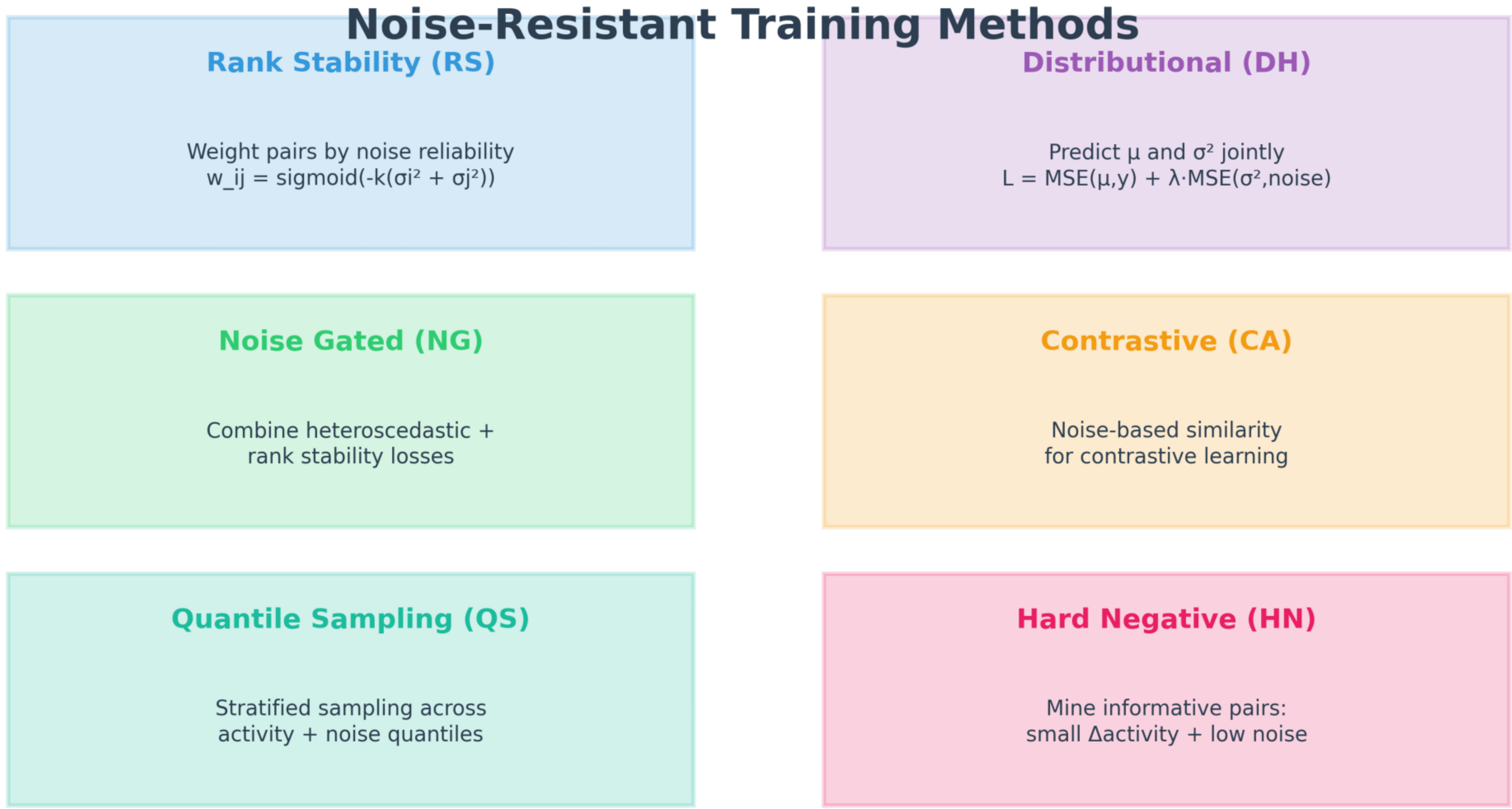
Heatmaps showing Spearman (left) and Pearson (right) for top 15 models across 4 CAGI5 elements.
Key Finding: GP1BB shows largest improvement (+108% for DH7). HBB consistently highest.

Figure 6: High-Confidence vs Low-Confidence Gap



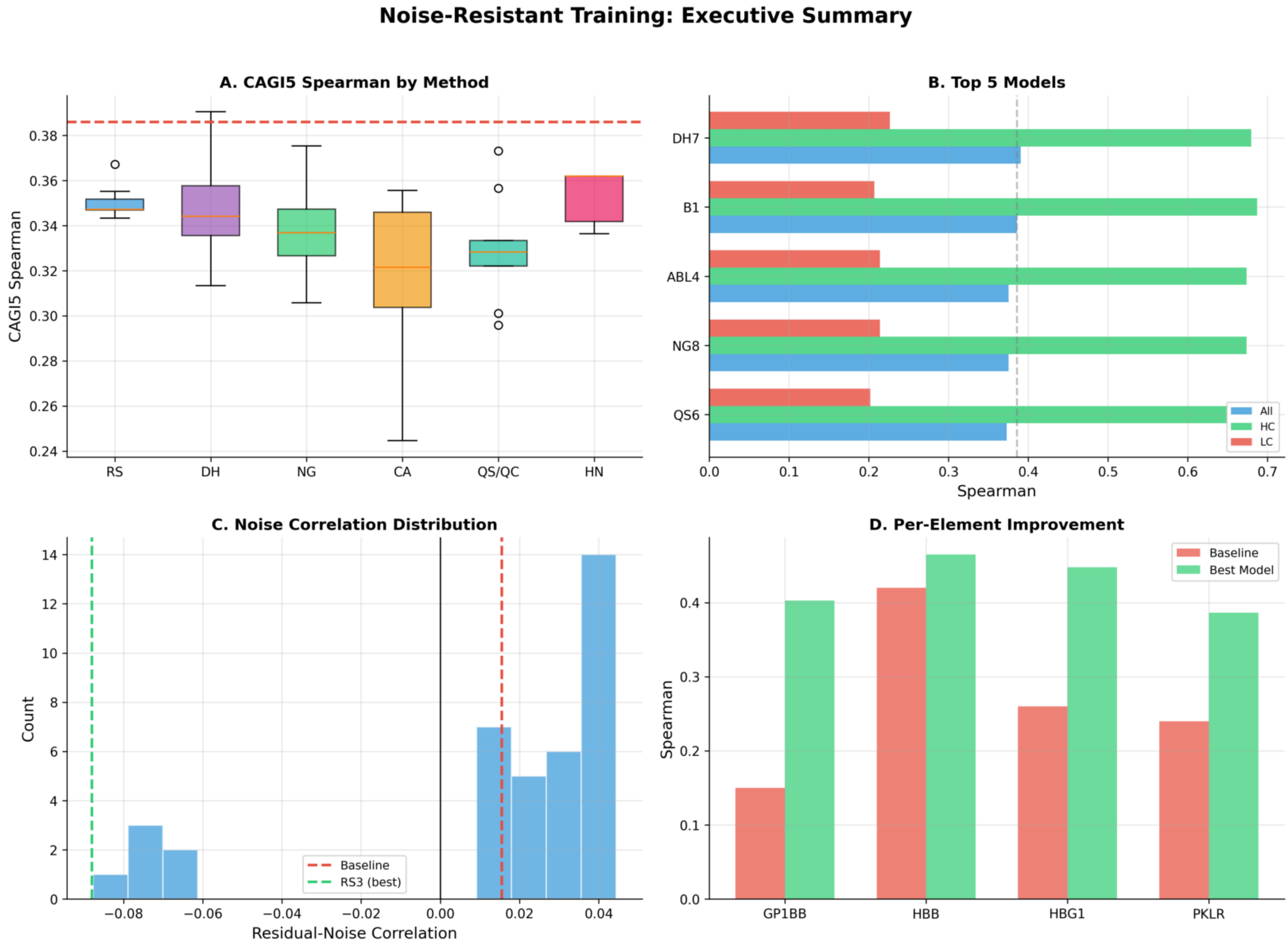
Each point represents a model. X-axis: HC-LC gap. Y-axis: Overall CAGI5.
Key Finding: DH7 achieves smallest gap among top performers, indicating balanced performance.

Figure 7: Methods Architecture



Six noise-resistant training strategies: Rank Stability, Distributional, Noise Gated, Contrastive, Quantile Sampling, and Hard Negative Mining.

Figure 8: Executive Summary



(A) Method comparison (B) Top 5 models (C) Noise correlation distribution (D) Per-element improvement. Key Results: DH7 best CAGI5 (+1.2%), RS3 only negative noise correlation, GP1BB +108%.