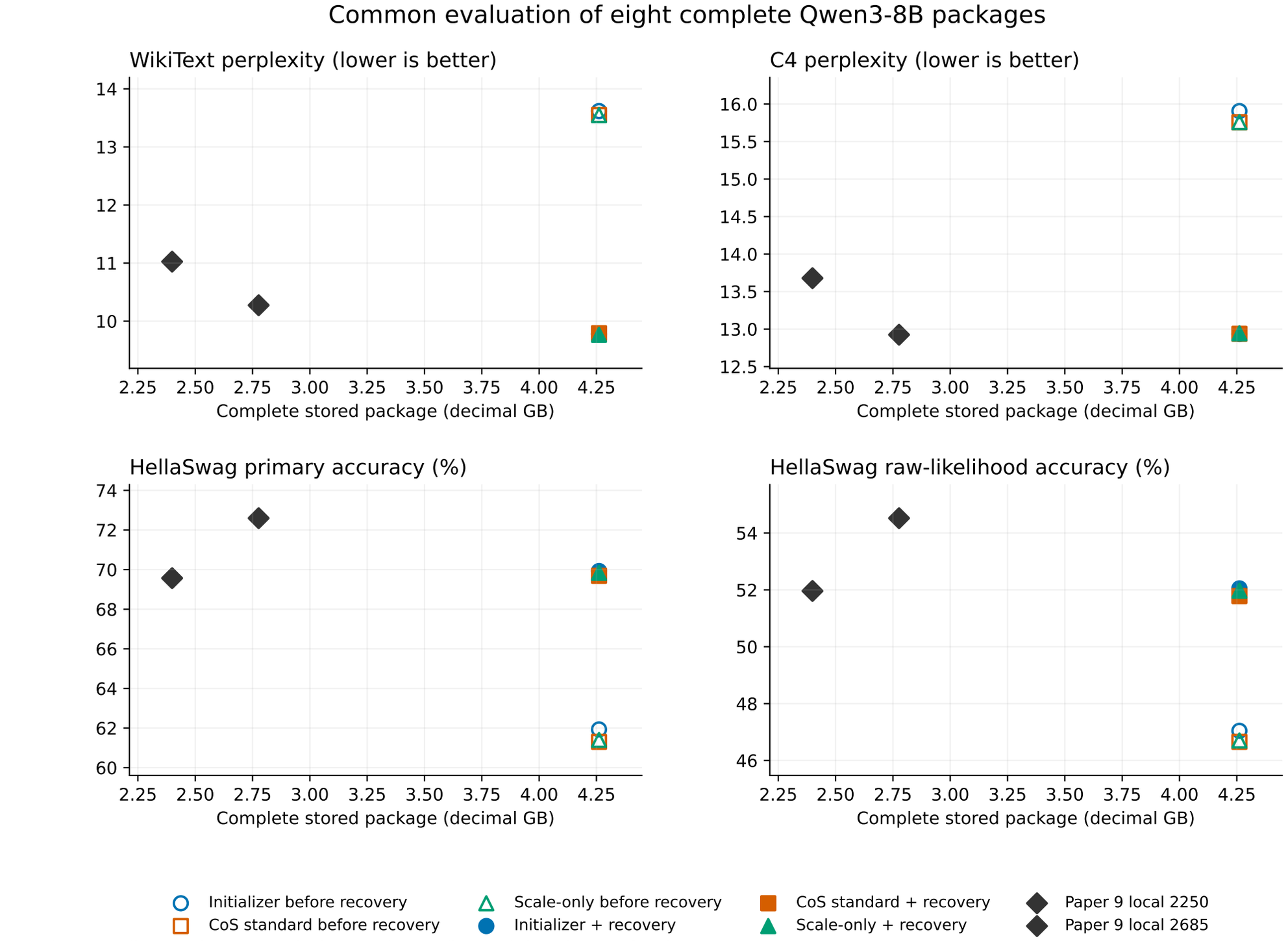
Common evaluation of eight complete Qwen3-8B packages
WikiText perplexity (lower is better)
C4 perplexity (lower is better)
HellaSwag primary accuracy (%)
HellaSwag raw-likelihood accuracy (%)
Complete stored package (decimal GB)
Initializer before recovery
CoS standard before recovery
Scale-only before recovery
Initializer + recovery
CoS standard + recovery
Scale-only + recovery
Paper 9 local 2250
Paper 9 local 2685
Exact byte coordinates; close CoS endpoints overlap. No jitter or inferred Pareto frontier.
CoS retains BF16 vocabulary; Paper 9 compresses it. Data, body rates and recovery differ.
Exposed panels; point estimates are not equivalence tests or a published-recipe reproduction.