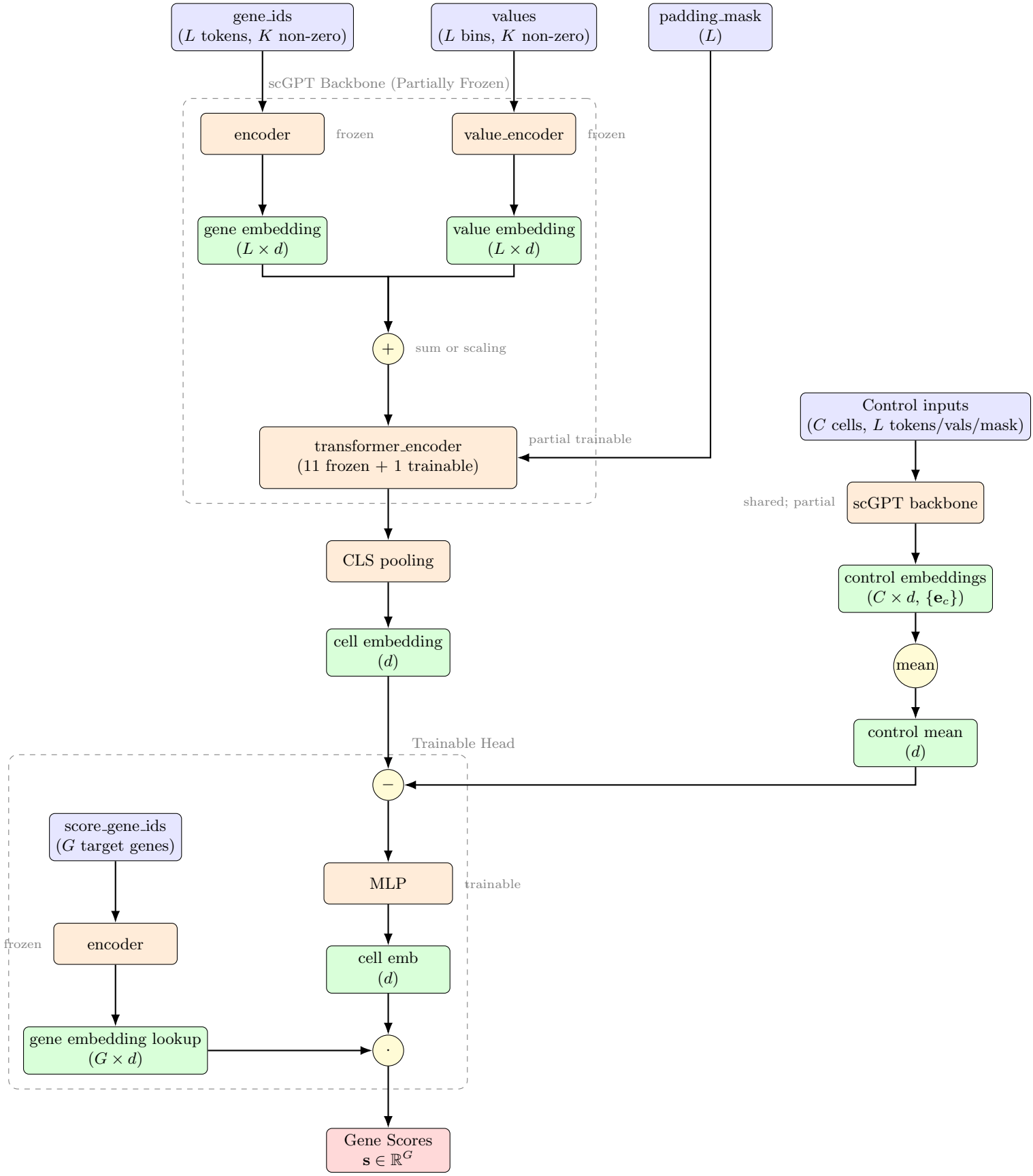

Figure 1: Per-cell scGPT Discriminative Pipeline for Perturbation Gene Prediction. For a single cell with K non-zero genes, inputs are padded to length L. Token IDs (`gene_ids`), binned expression values, and a padding mask are fed to the scGPT backbone. The `encoder` and `value_encoder` produce gene/value embeddings, which are fused by element-wise addition (default) or scaling (`input_emb_style=scaling`), then passed through `transformer_encoder` to produce a CLS-based cell embedding. A control branch runs the same backbone and the control mean is subtracted from the perturbed embedding. The GeneScore head applies a projection MLP to the cell embedding before the dot product with the gene embedding lookup. The `score_gene_ids` input is length G, the full set of candidate target genes, and its order defines the score output columns. The backbone is partially frozen in this setup, with only the last transformer layer(s) and encoder norm optionally trainable. Here, G is the total number of genes in the dataset, C is the number of control cells per perturbed cell, and d is the embedding size.