G&C Training Loss Distribution
Student Dim: 100, Seed: 0, Seq Len: 5

Total samples: 100000
Below ε_train: 690 (0.7%)
Mean: 1.0768
Median: 0.9071
Std: 0.8882

ε_train = 0.0100