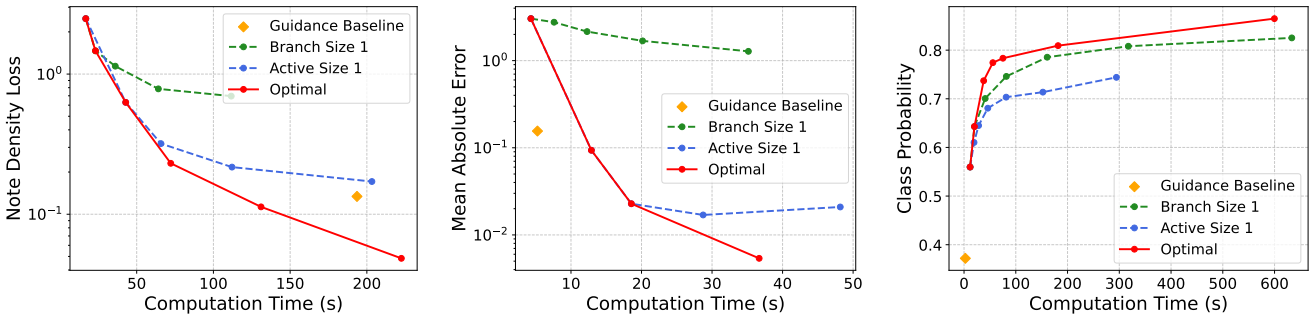
1 Optimization Effect and Scaling Behavior of SPS

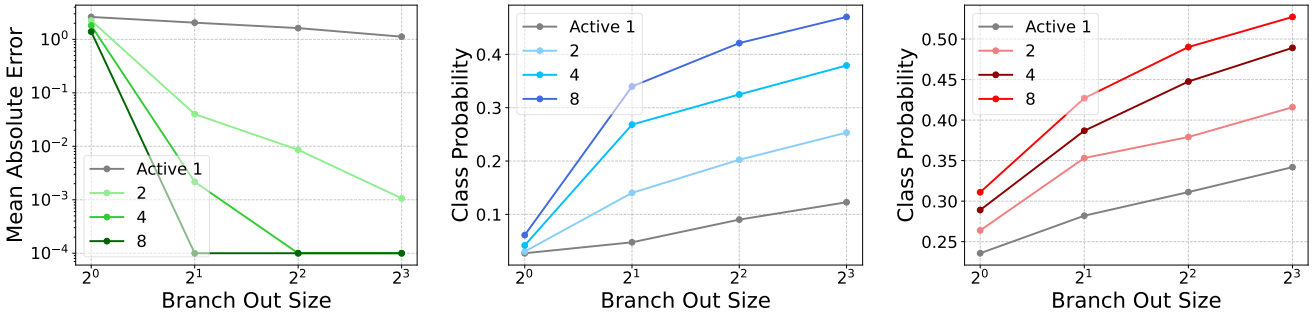
1.1 Overview



(a) Symbolic Music Generation (Loss ↓) (b) Small Molecule Generation (MAE ↓) (c) Enhancer DNA Design (Prob ↑)

Figure 1: **SPS outperforms the guidance baseline, with the optimization effect of the objective function following a scaling behavior with inference time.** We compare SPS with the strongest guidance baseline across tasks. We evaluate SPS with varying active set and branch-out sizes. **"Optimal"** refers to the combination of active set and branch out sample sizes that achieve the best performance within the same inference time. The optimal line shows that SPS outperforms the guidance baseline and shows an inference time scaling behavior.

1.2 Scaling Effect of Active and Branch Out Sizes



(a) SPS-SD for Small Molecule Generation (b) SPS-SC for Enhancer DNA Design (c) SPS-G for Enhancer DNA Design

Figure 2: **Scaling Behavior with Fixed Active or Branch-Out Size:** Increasing either the active size or the branch out size enhances performance. (a): The MAEs are capped at 10^{-4}; (b) & (c): The results are for Class 3 in DNA enhancer design with strength 20 for SPS-G.

2 A *v.s.* K with Fixed Computation Budget

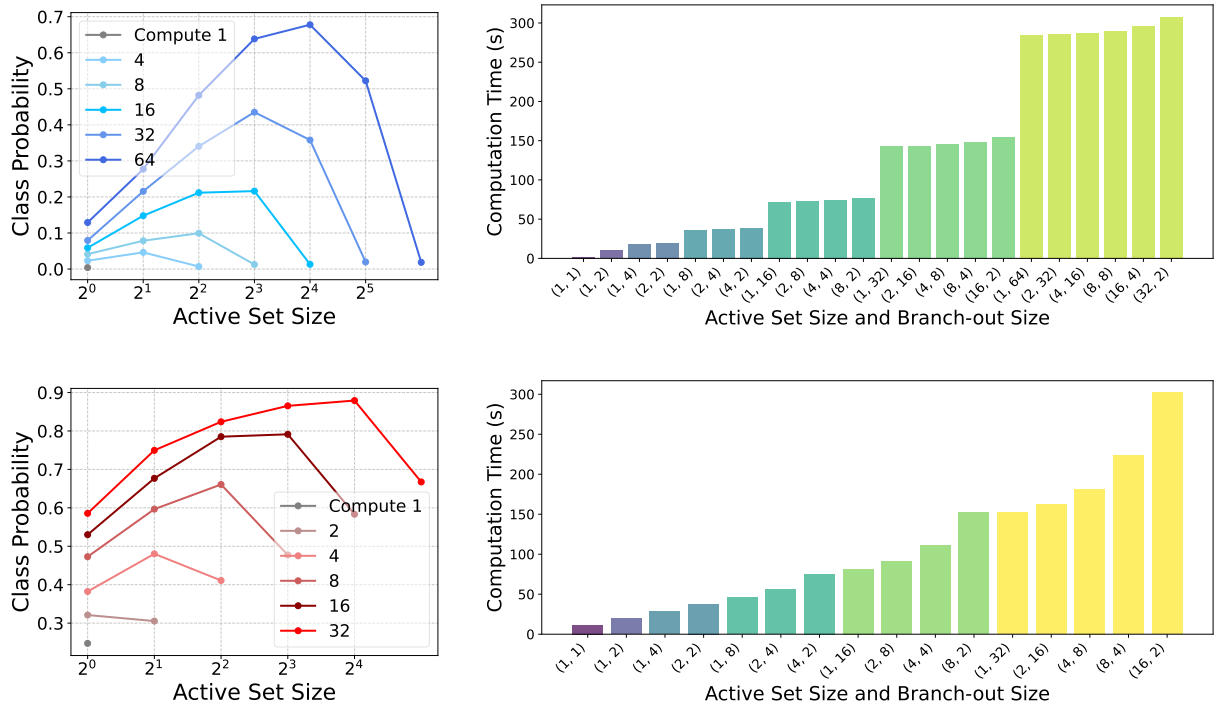
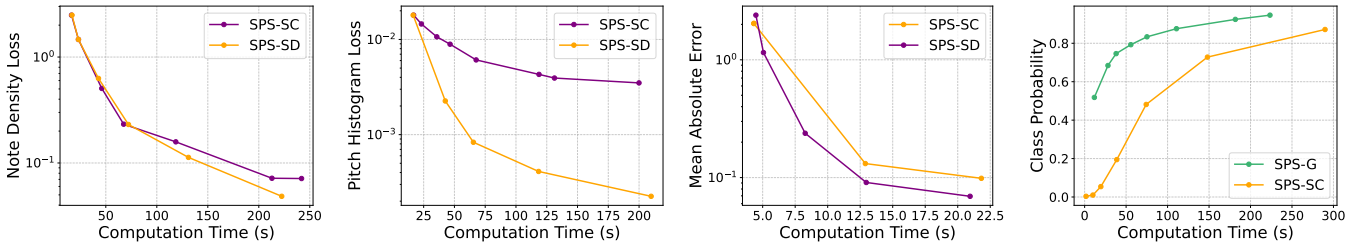


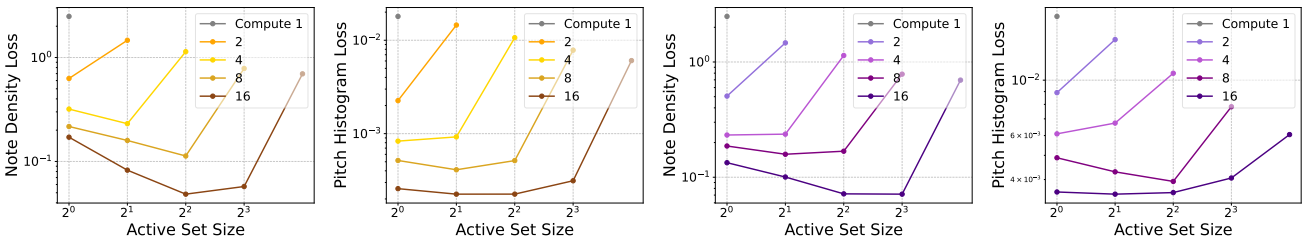
Figure 3: **Trade-off between Active Set Size A and Branch-out Size K with Fixed Computation:** SPS-SC (Top) and SPS-G (Bottom) vary (A, K) with fixed total compute $A * K$ for enhancer DNA design. Performance peaks when A and K are in the moderate range. Fixed $A * K$ can be viewed as a fixed computation budget where the actual inference time is shown on the right. The results are for Class 1 and $\gamma = 20$ for SPS-G.

3 More Results



(a) Note Density (Music) (b) Pitch Histogram (Music) (c) Molecular Design ($N_r^* = 5$) (d) Enhancer DNA (Class 4)

Figure 4: **Inference Time Scaling Behavior:** As the active set size and branch-out size increase, the optimization effect of the objective function scales with inference time. This trend is consistently observed across all algorithms and tasks. The inference time is measured with a batch size of 1 for music and 100 for molecule and DNA design. For DNA design, $\gamma = 20$ for SPS-G.



(a) SPS-SD for ND (b) SPS-SD for PH (c) SPS-SC for ND (d) SPS-SC for PH

Figure 5: Trade-off between Active Set Size A and Branch-out Size K on Music Generation.