

Experiment 4 — Figures Comparison

Compare commit_label vs commit_pred at Dice cutoffs 0.90 and 0.97. Grouped by Plan A, Plan C, and Plan B.

```
In [1]:  
from pathlib import Path  
from IPython.display import display, HTML  
  
def _repo_root():  
    p = Path().resolve()  
    for cand in [p] + list(p.parents):  
        if (cand / 'experiments' / 'scripts').is_dir():  
            return cand  
    return p  
  
ROOT = _repo_root()  
base = ROOT / 'experiments' / 'scripts' / 'experiment_4'  
paths = {  
    'commit_label_90': base / 'commit_label_90',  
    'commit_label_97': base / 'commit_label_97',  
    'commit_pred_90': base / 'commit_pred_90',  
    'commit_pred_97': base / 'commit_pred_97',  
}  
  
def plan_a_figures(root: Path) -> Path: return root / 'A' / 'results' / 'figures'  
def plan_b_figures(root: Path) -> Path: return root / 'B' / 'figures'  
  
def list_pngs(p: Path):  
    if not p.exists():  
        return {}  
    return {f.name: f for f in sorted(p.glob('*.*'))}  
  
def build_pairs(dir90: Path, dir97: Path, filter_fn=lambda n: True):  
    d90, d97 = list_pngs(dir90), list_pngs(dir97)  
    commons = sorted(n for n in (d90.keys() & d97.keys()) if filter_fn(n))  
    return [(n, str(d90[n]), str(d97[n])) for n in commons]  
  
def build_pairs_map(dir90: Path, dir97: Path, filter_fn=lambda n: True):  
    d90, d97 = list_pngs(dir90), list_pngs(dir97)  
    names = sorted(n for n in (d90.keys() & d97.keys()) if filter_fn(n))  
    return {n: (str(d90[n]), str(d97[n])) for n in names}  
  
def show_interleaved(heading, label_map, pred_map, col_titles=('cutoff 0.90', 'cutoff 0.97')):  
    display(HTML(f<h2 style="margin-top:18px">{heading}</h2>))  
    names = sorted(set(label_map.keys()) | set(pred_map.keys()))  
    if not names:  
        display(HTML('<p><em>No matching figures.</em></p>'))  
        return  
    rows = []  
    for name in names:  
        rows.append(f<tr><td colspan=2 style="text-align:center;font-weight:600;padding-top:14px">{name}</td></tr>)  
        if name in label_map:  
            p90, p97 = label_map[name]  
            rows.append(  
                f<tr><td style="width:50%;text-align:center"><div style="font-size:12px;color:#555">commit_label — {col_titles[0]}</div>  
                f<td style="width:50%;text-align:center"><div style="font-size:12px;color:#555">commit_label — {col_titles[1]}</div>  
            )  
        if name in pred_map:  
            p90, p97 = pred_map[name]  
            rows.append(  
                f<tr><td style="width:50%;text-align:center"><div style="font-size:12px;color:#555">commit_pred — {col_titles[0]}</div>  
                f<td style="width:50%;text-align:center"><div style="font-size:12px;color:#555">commit_pred — {col_titles[1]}</div>  
            )  
    html = "<table style='width:100%'>{}</table>".format("\n".join(rows))  
    display(HTML(html))  
  
is_plan_c = lambda n: n.startswith('eval_') or n.startswith('plan_c_')  
is_not_plan_c = lambda n: not is_plan_c(n)
```

Plan A (non-Plan C) — Interleaved Label/Pred

In [2]:

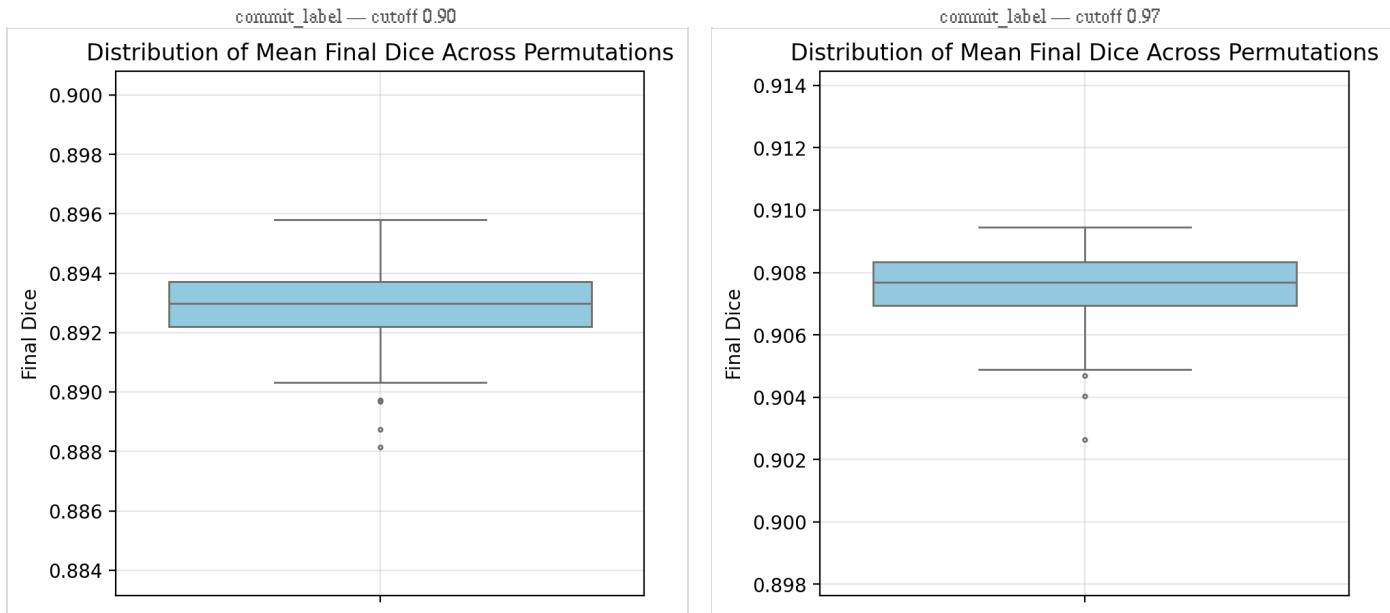
```

label_map = build_pairs_map(plan_a_figures(paths['commit_label_90']), plan_a_figures(paths['commit_label_97']), is_not_plan_c)
pred_map = build_pairs_map(plan_a_figures(paths['commit_pred_90']), plan_a_figures(paths['commit_pred_97']), is_not_plan_c)
show_interleaved('Plan A — Interleaved (commit_label & commit_pred)', label_map, pred_map)

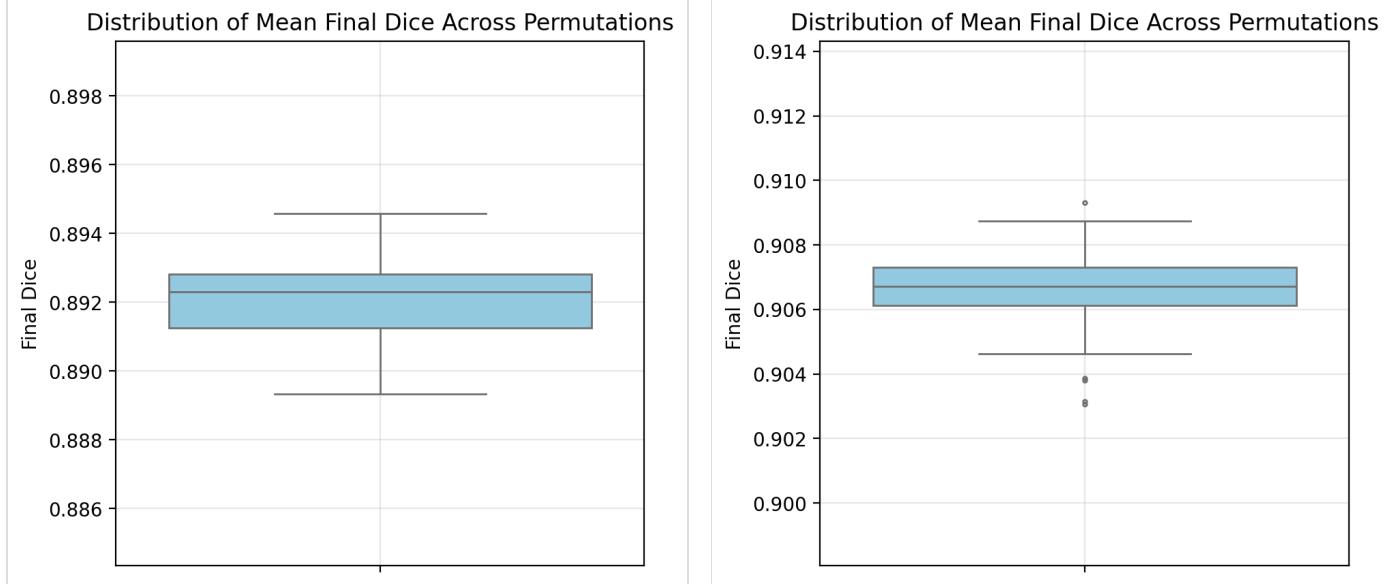
```

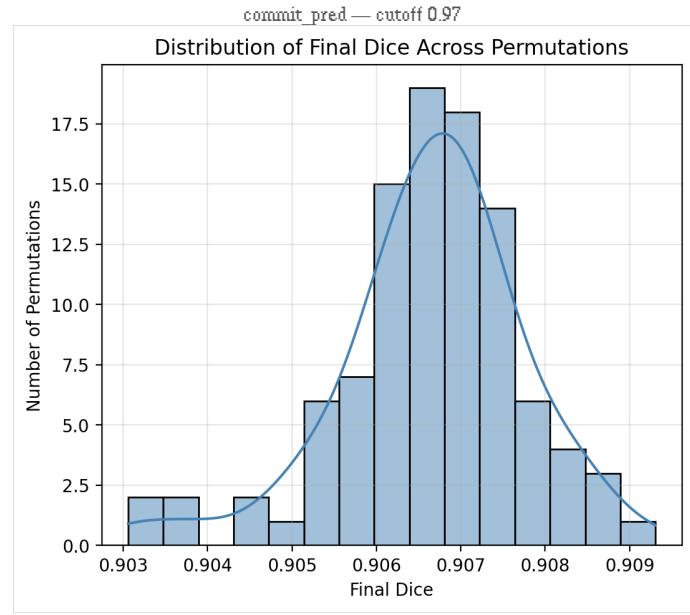
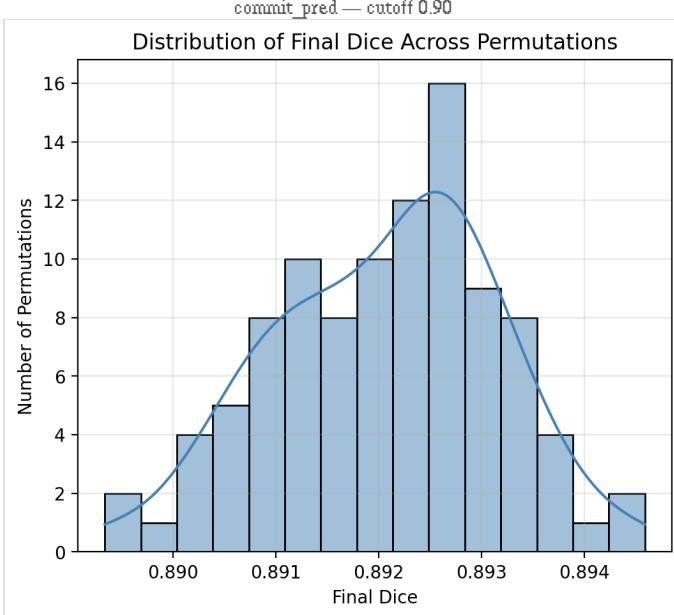
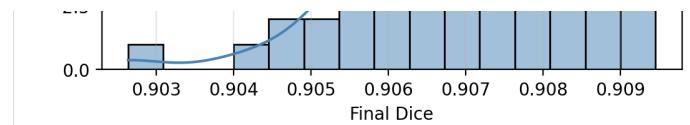
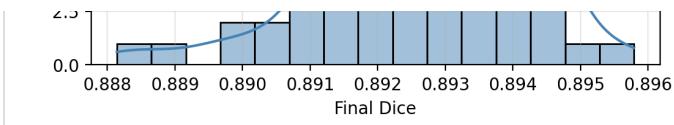
Plan A — Interleaved (commit_label & commit_pred)

`final_dice_mean_box.png`

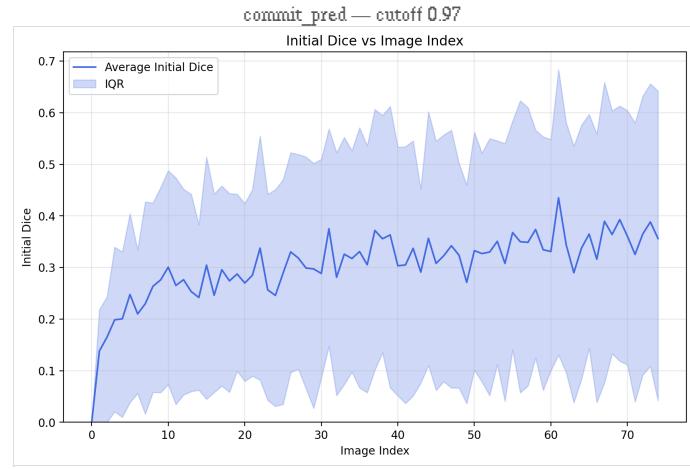
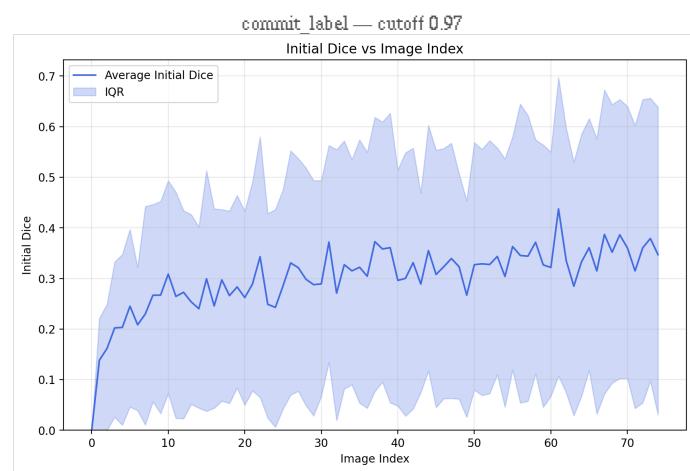
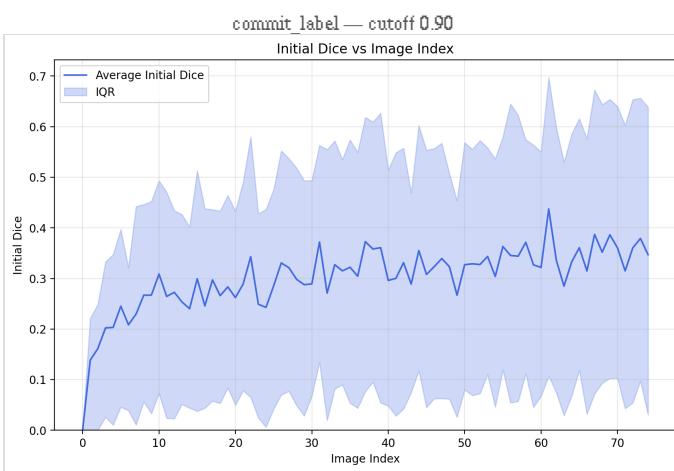


`final_dice_mean_hist.png`

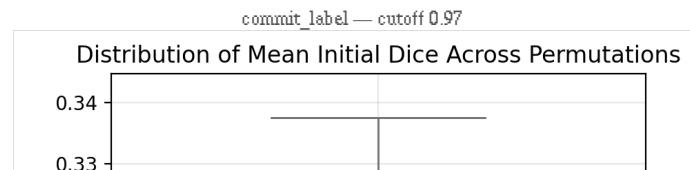
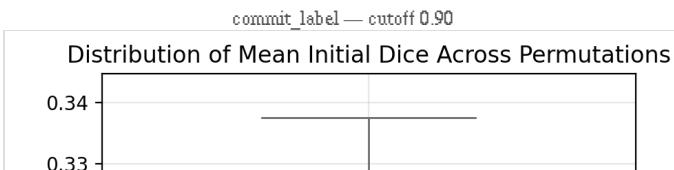


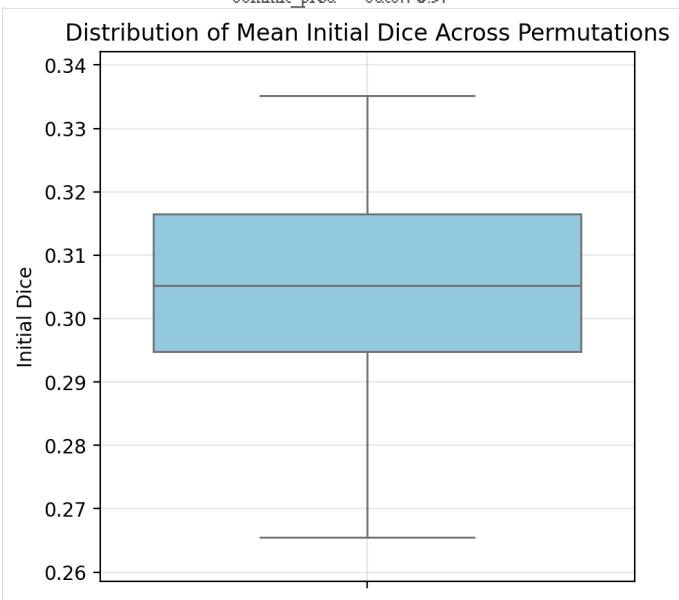
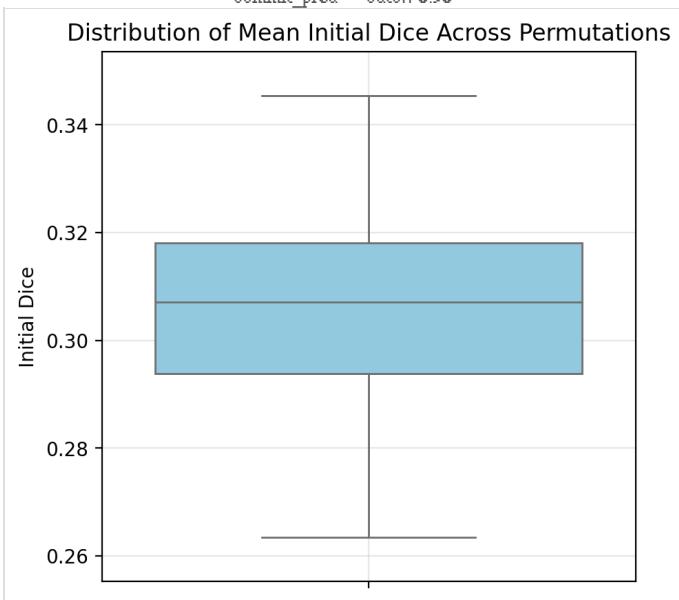
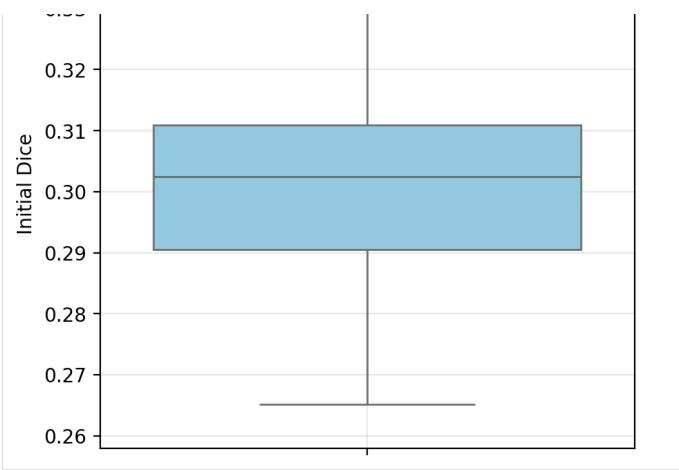
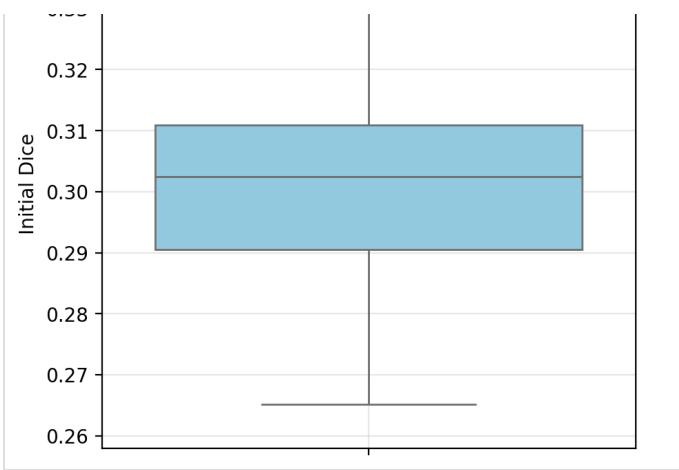


initial_dice_iqr.png

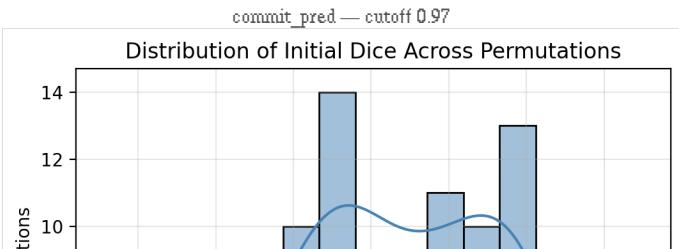
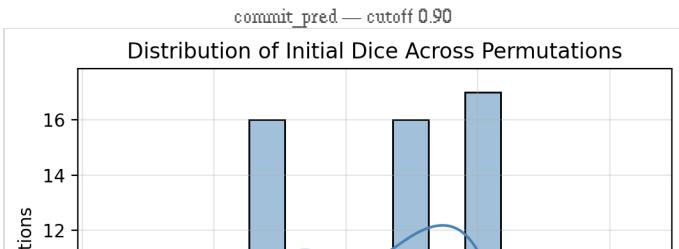
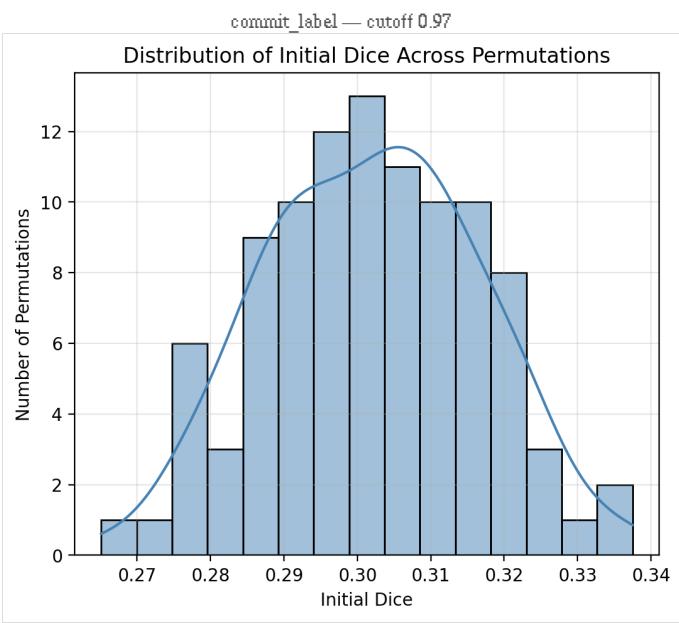
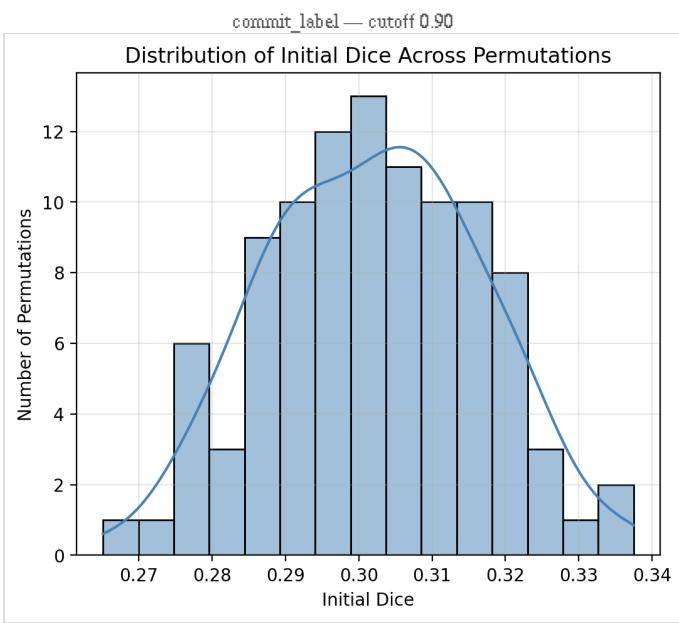


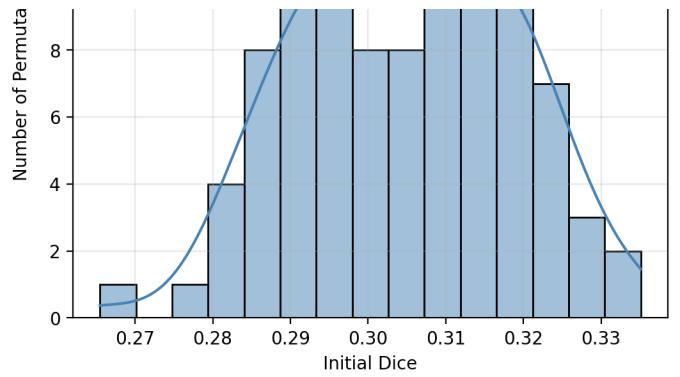
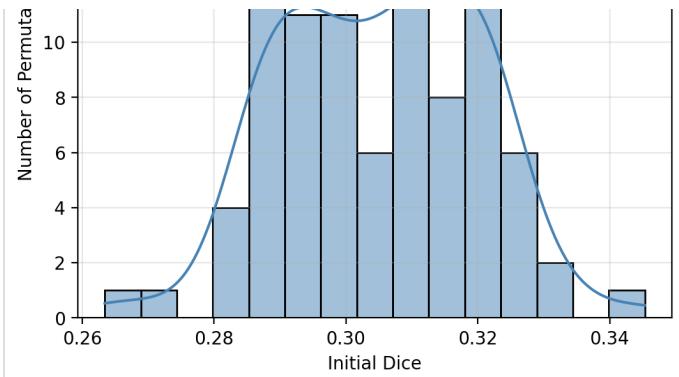
initial_dice_mean_box.png



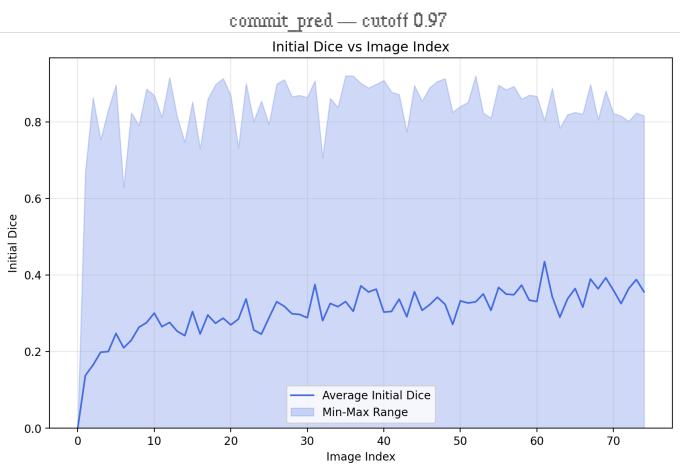
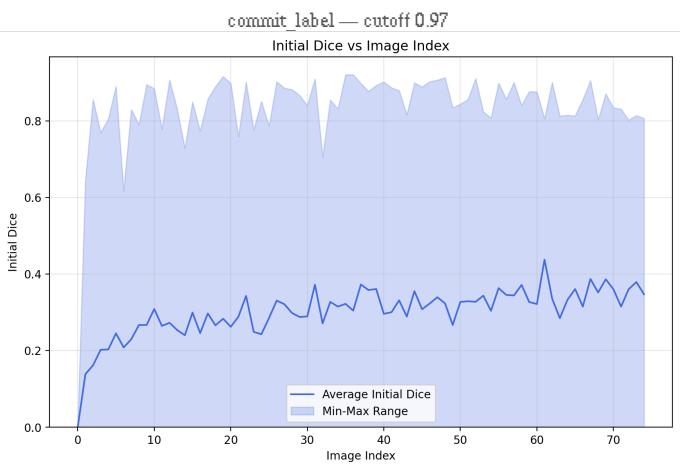
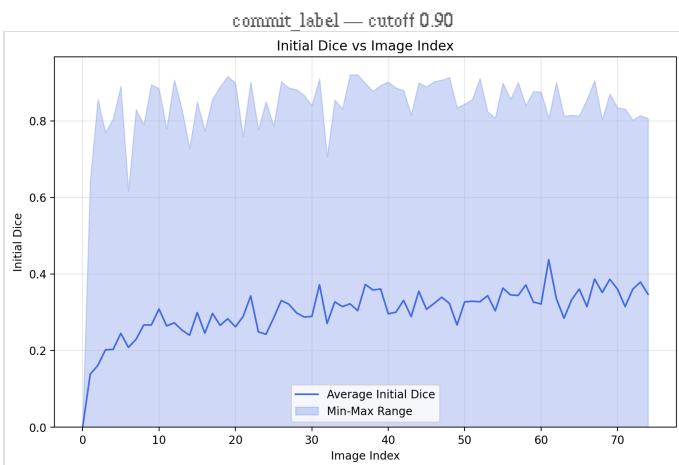


initial_dice_mean_hist.png

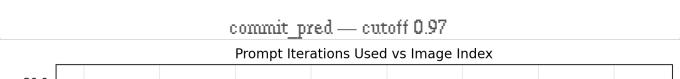
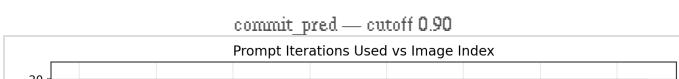
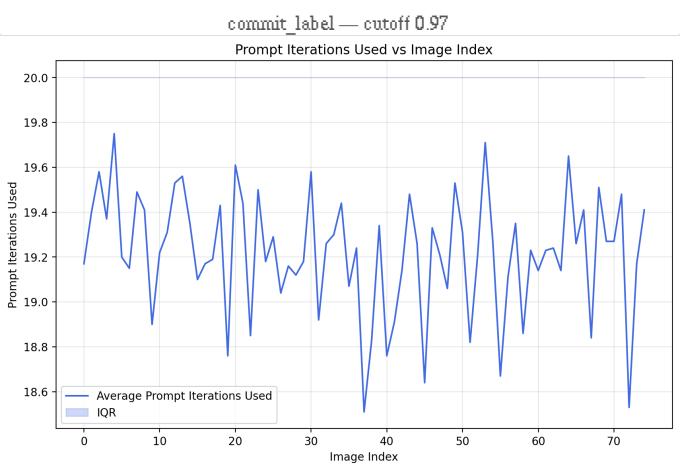
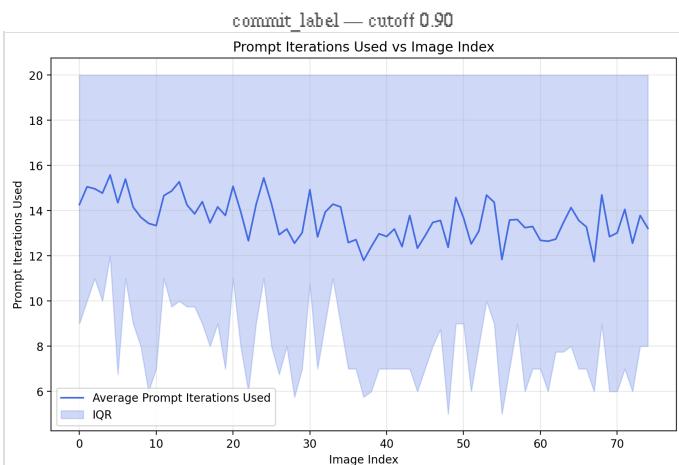


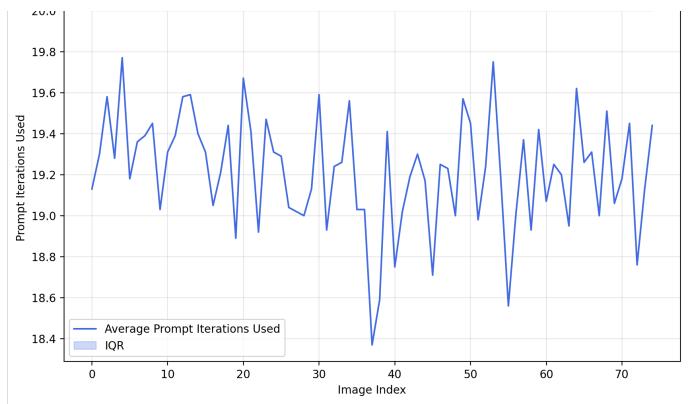
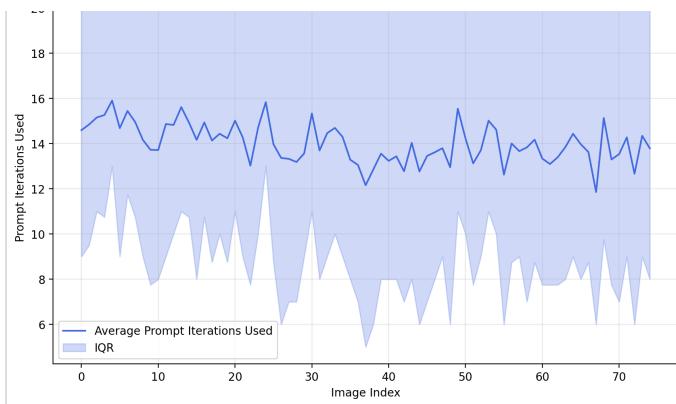


initial_dice_range.png

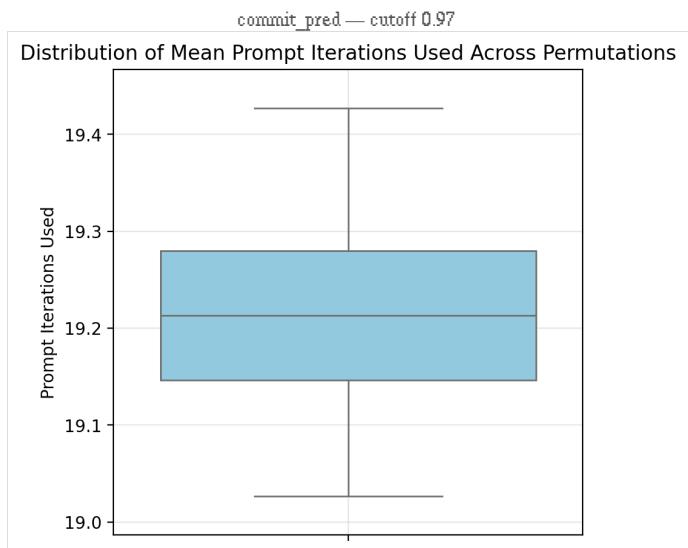
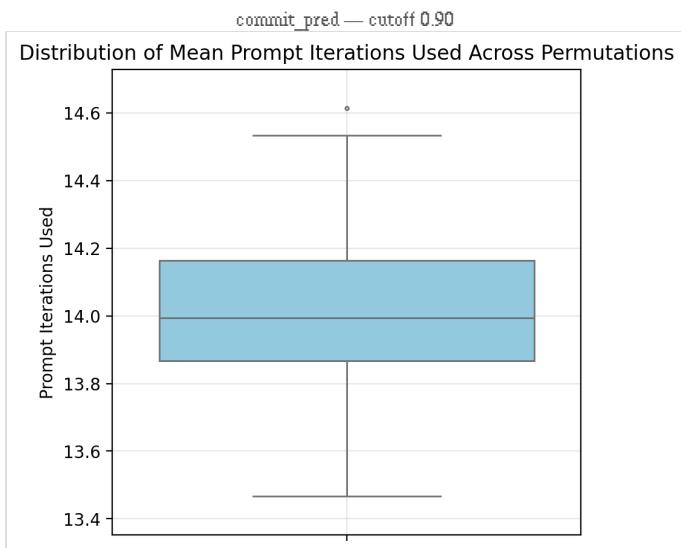
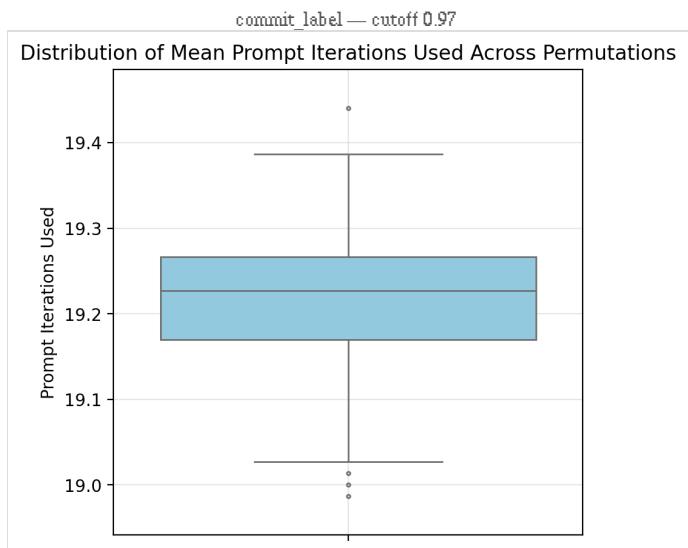
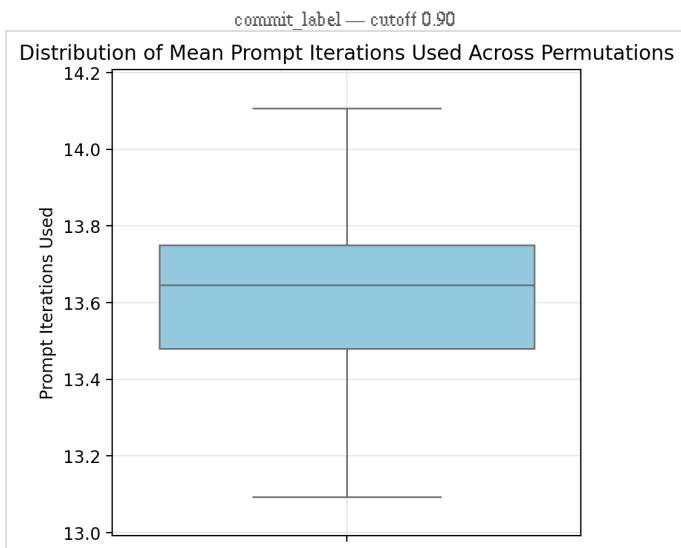


iterations_iqr.png

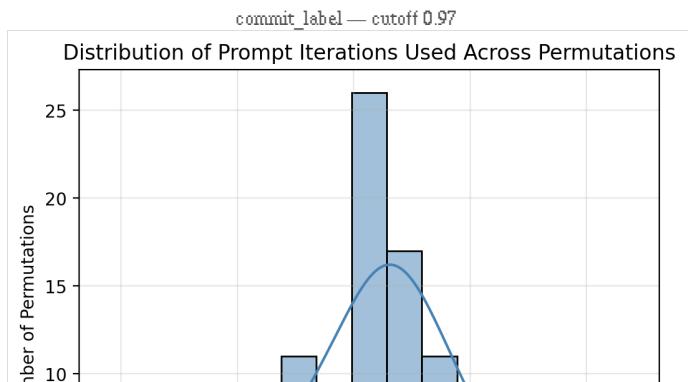
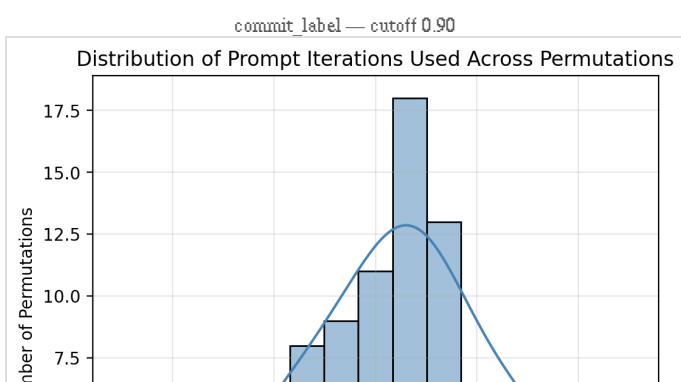


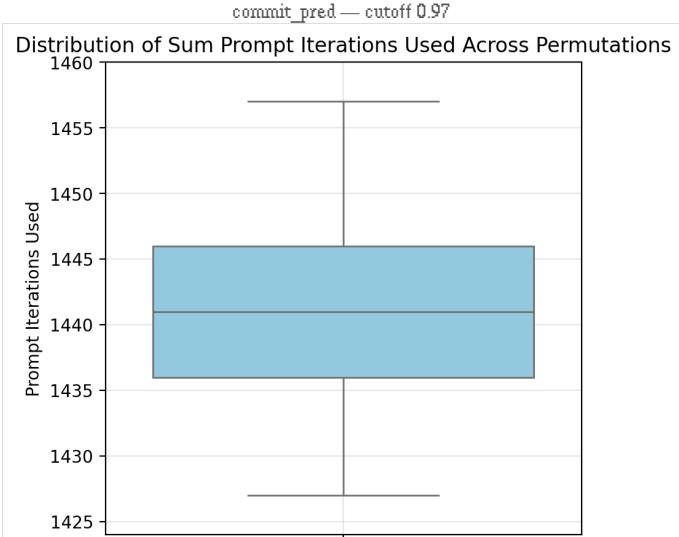
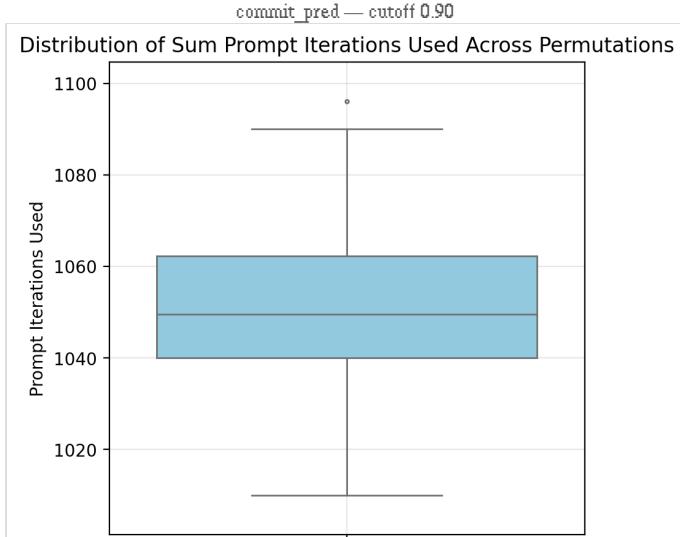
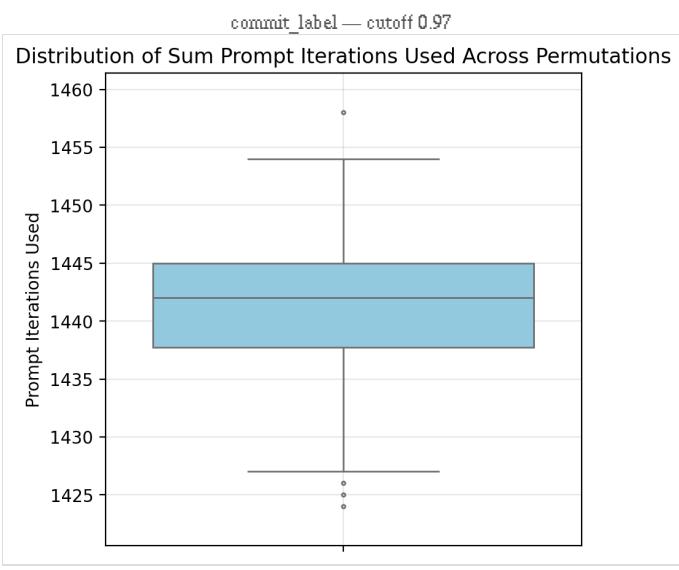
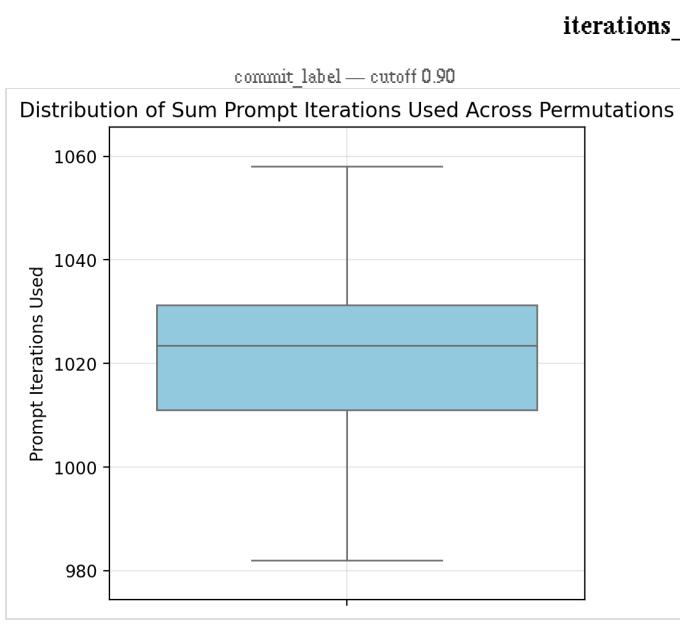
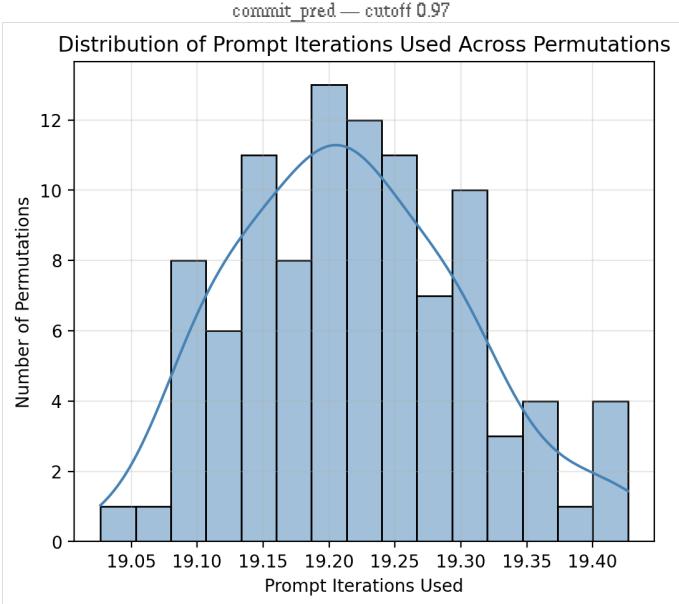
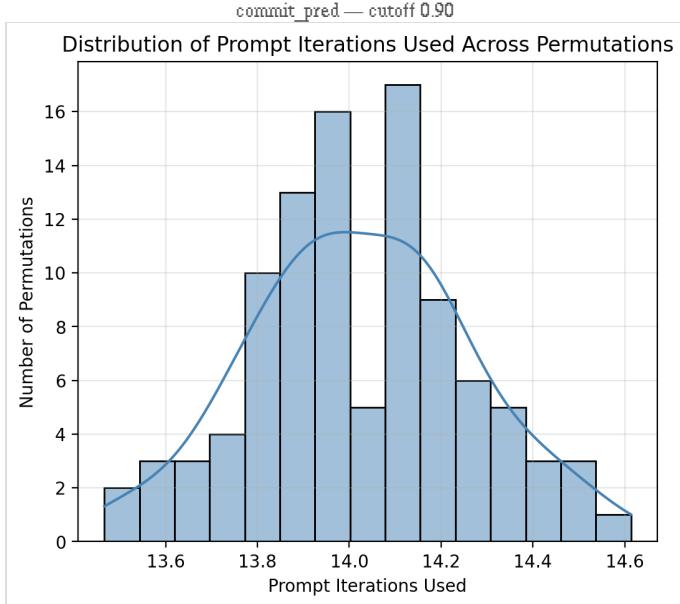
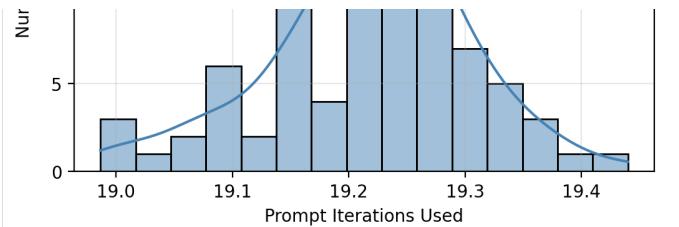
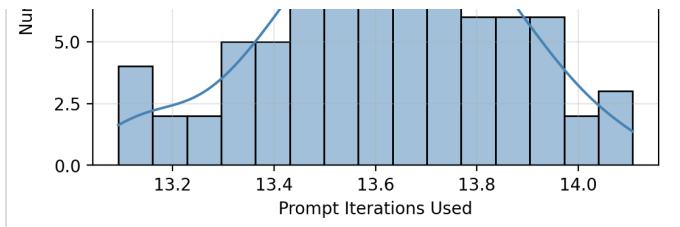


iterations_mean_box.png

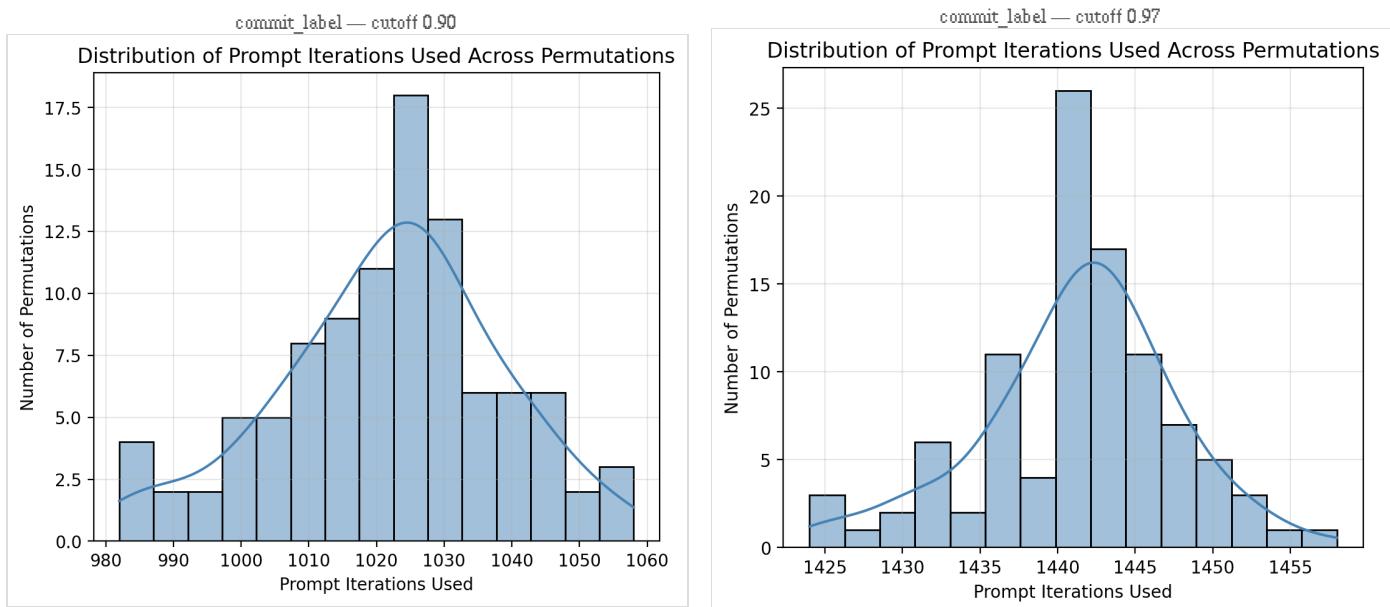


iterations_mean_hist.png

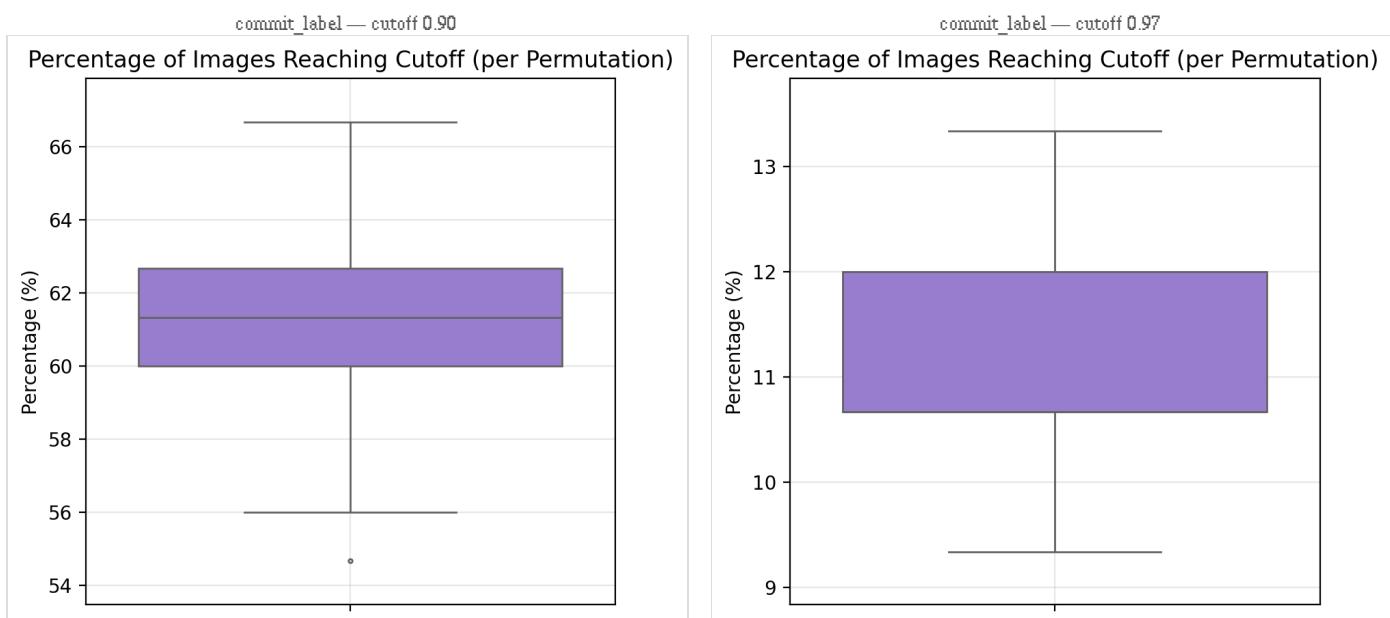


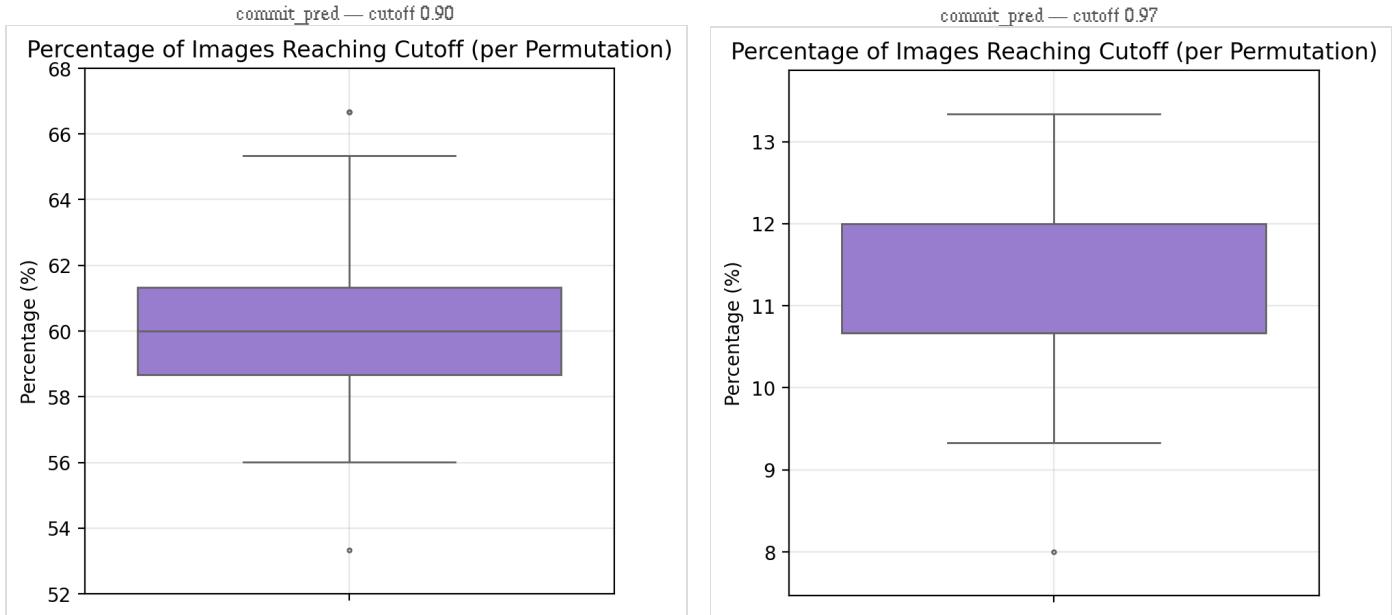


iterations_sum_hist.png

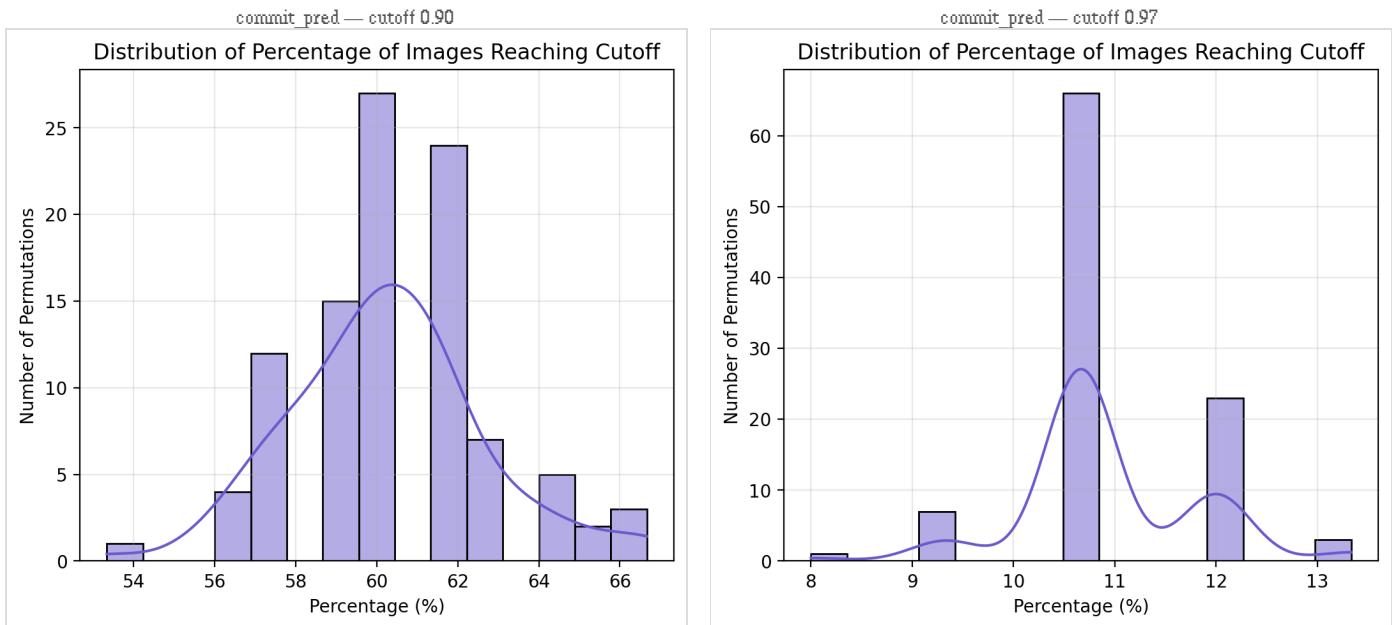
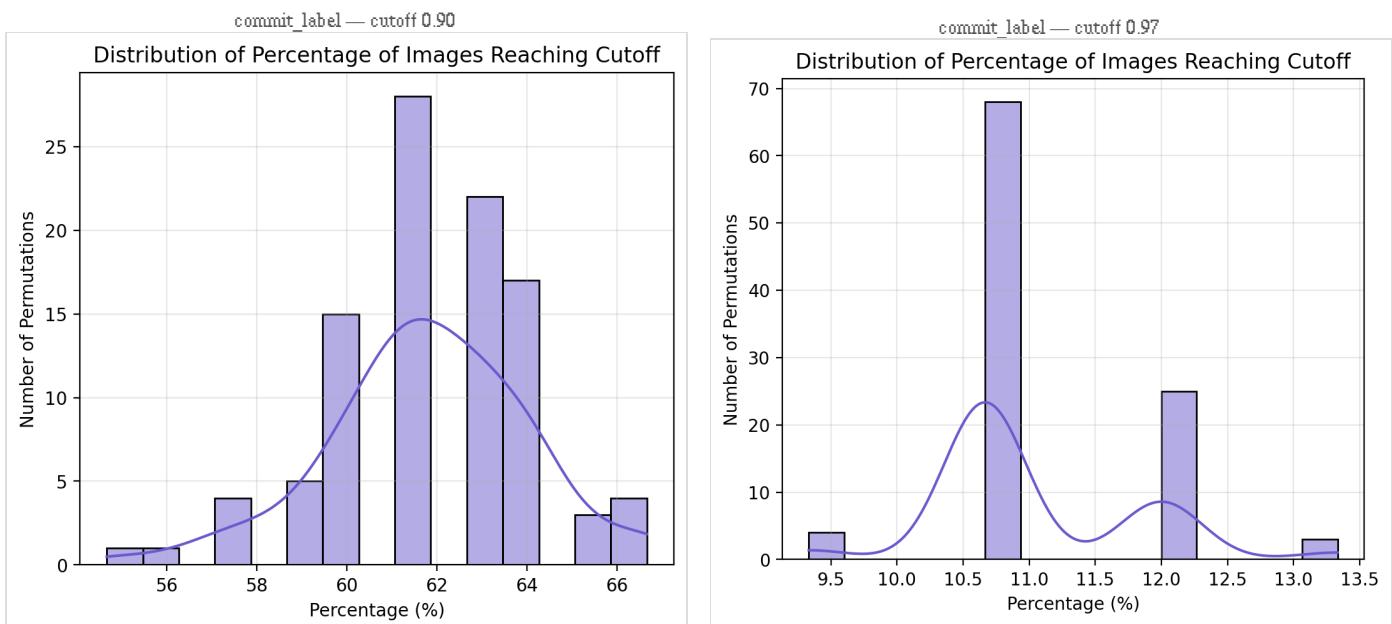


reached_cutoff_percentage_box.png





reached_cutoff_percentage_hist.png



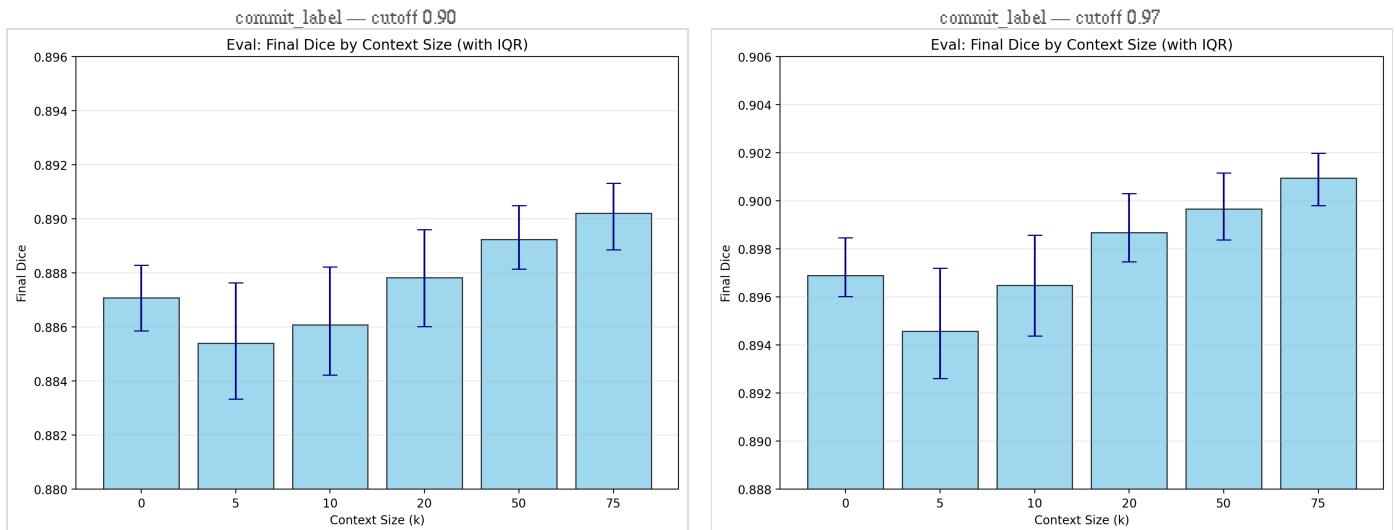
Plan C (eval * + plan_c *) — Interleaved Label/Pred

In [3]:

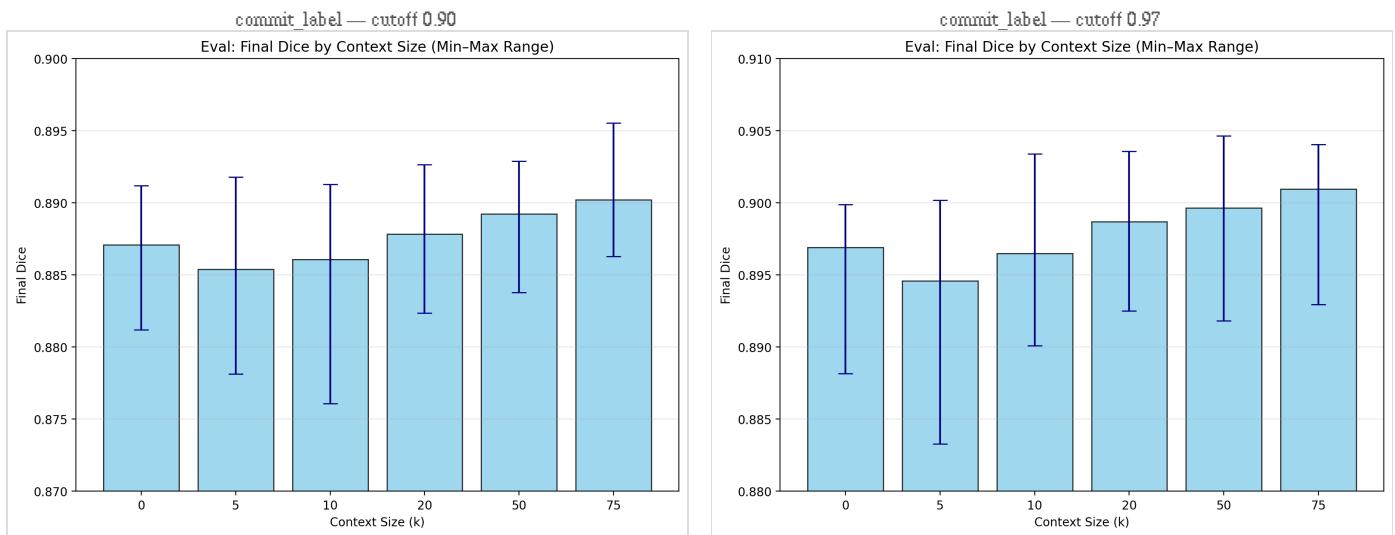
```
label_map = build_pairs_map(plan_a_figures(paths['commit_label_90']), plan_a_figures(paths['commit_label_97']), is_plan_c)
pred_map = build_pairs_map(plan_a_figures(paths['commit_pred_90']), plan_a_figures(paths['commit_pred_97']), is_plan_c)
show_interleaved('Plan C — Interleaved (commit_label & commit_pred)', label_map, pred_map)
```

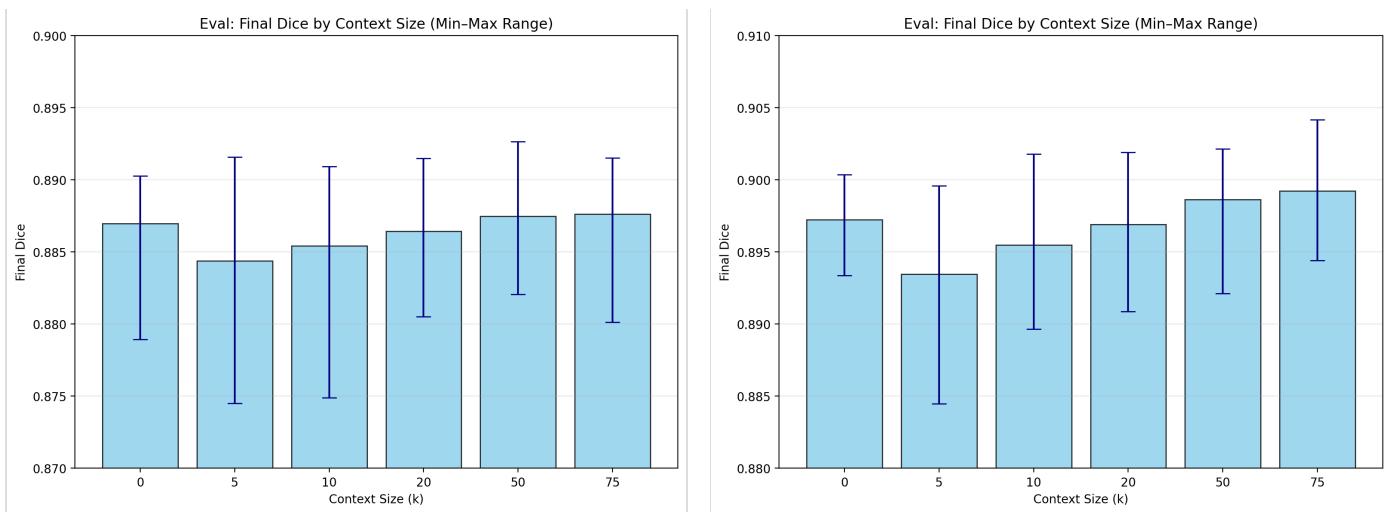
Plan C — Interleaved (commit_label & commit_pred)

eval_final_dice_iqr.png

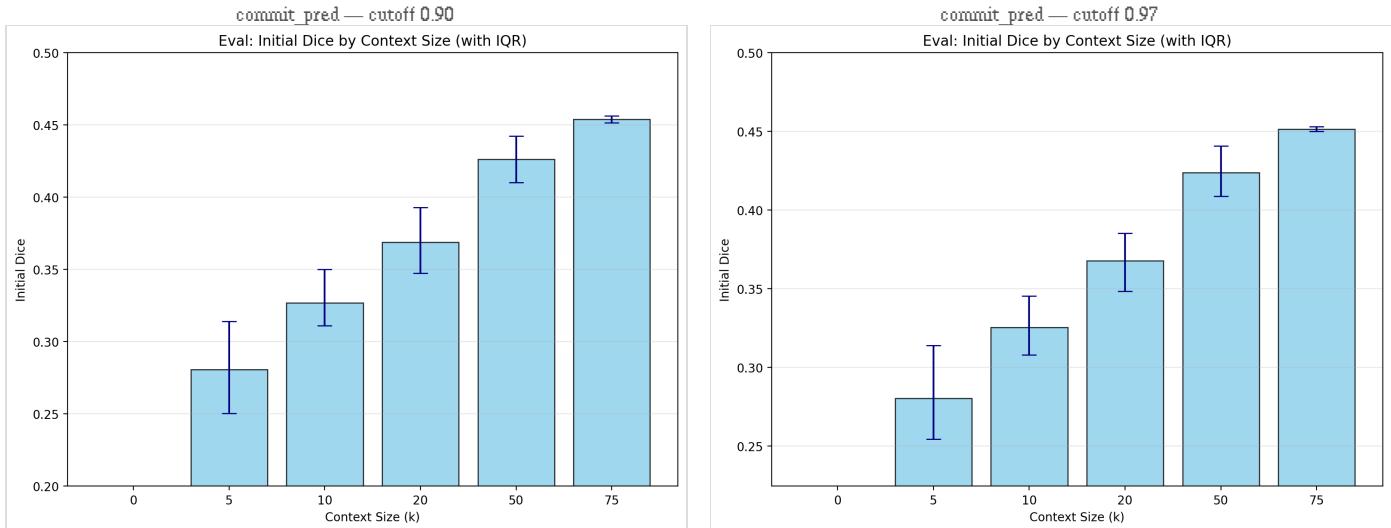
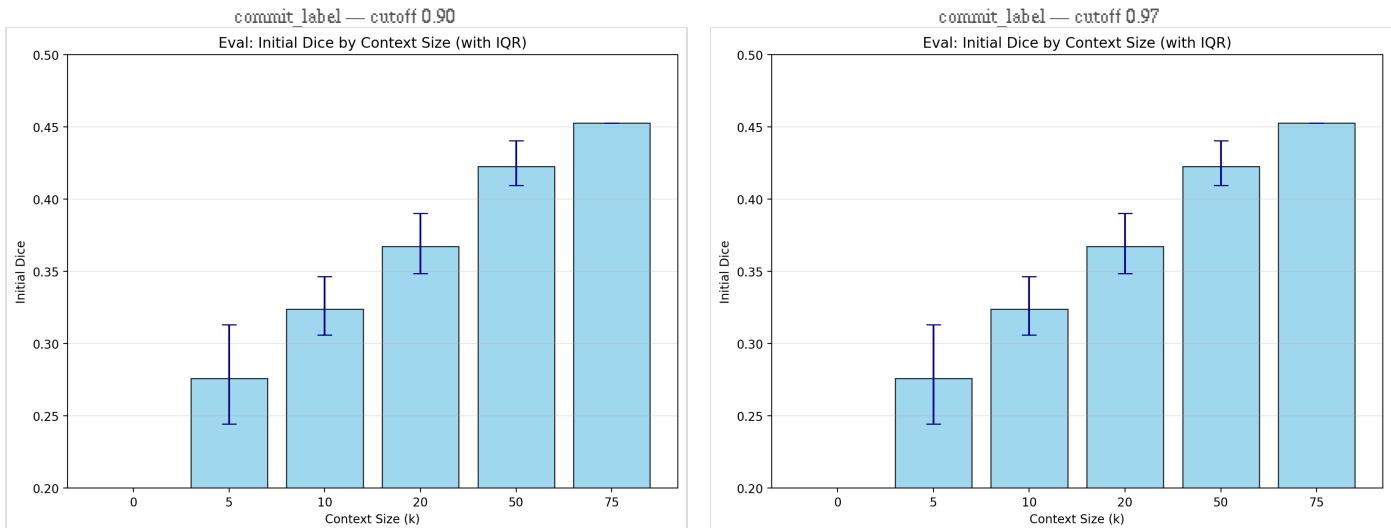


eval_final_dice_range.png

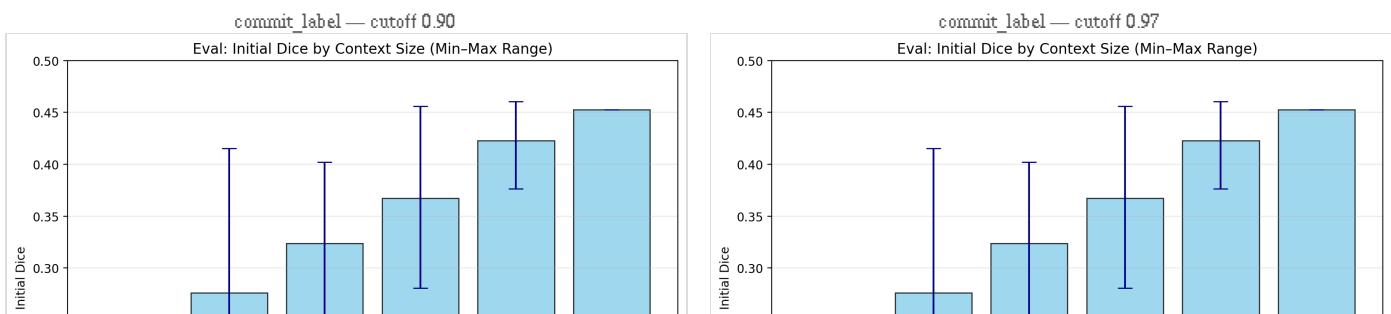


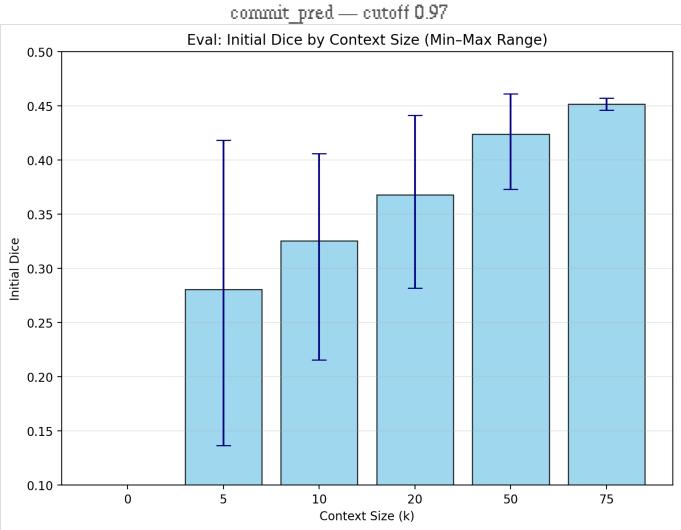
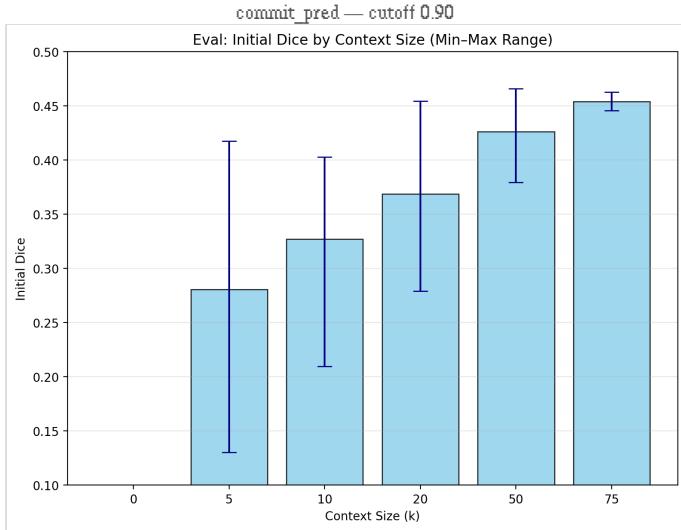
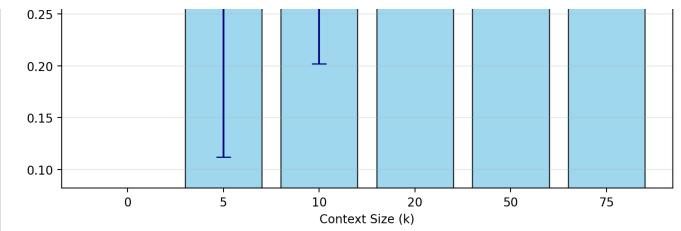
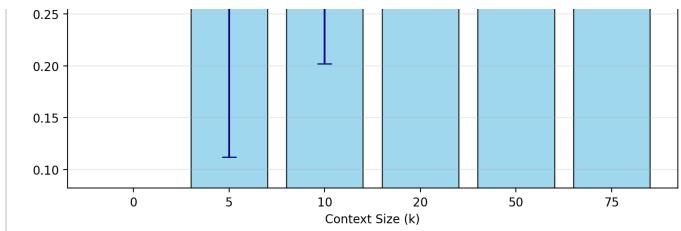


eval_initial_dice_iqr.png

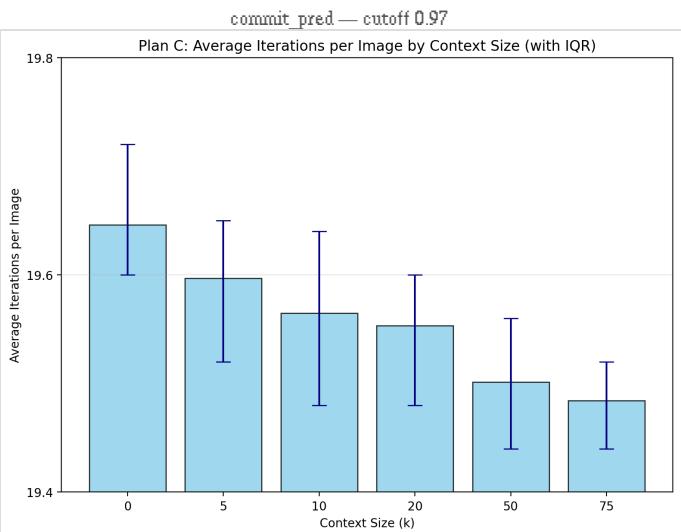
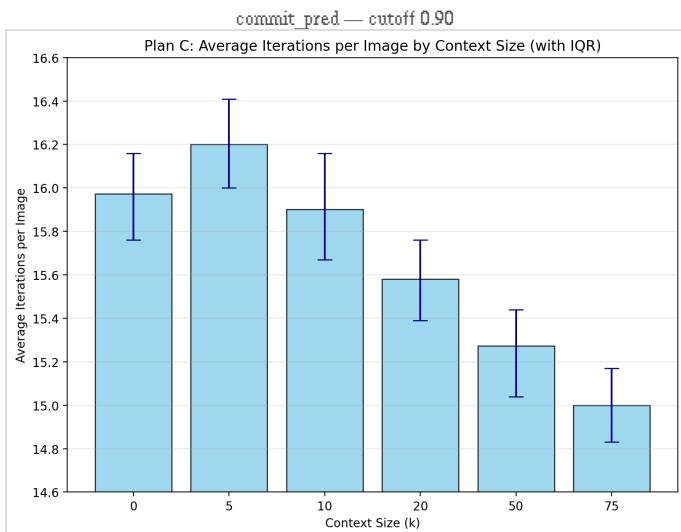
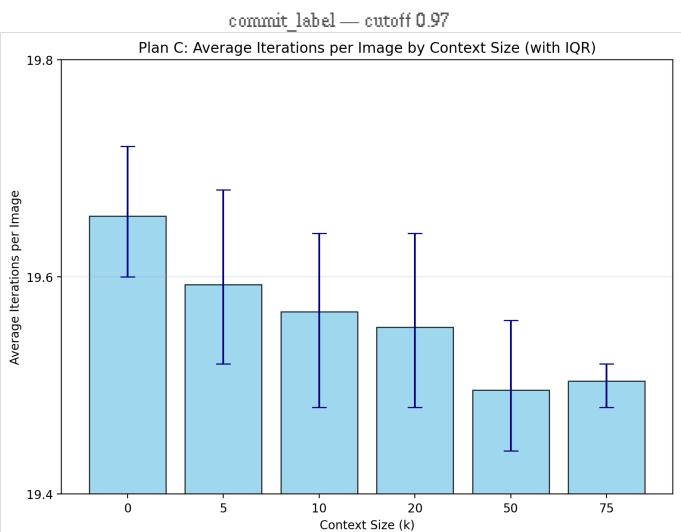
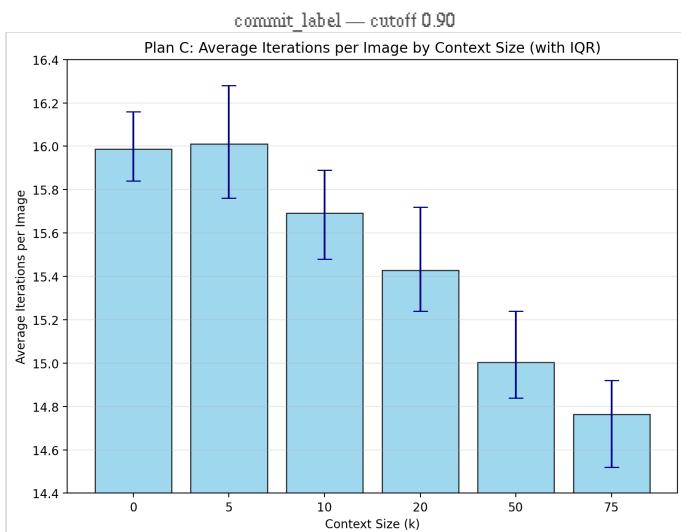


eval_initial_dice_range.png





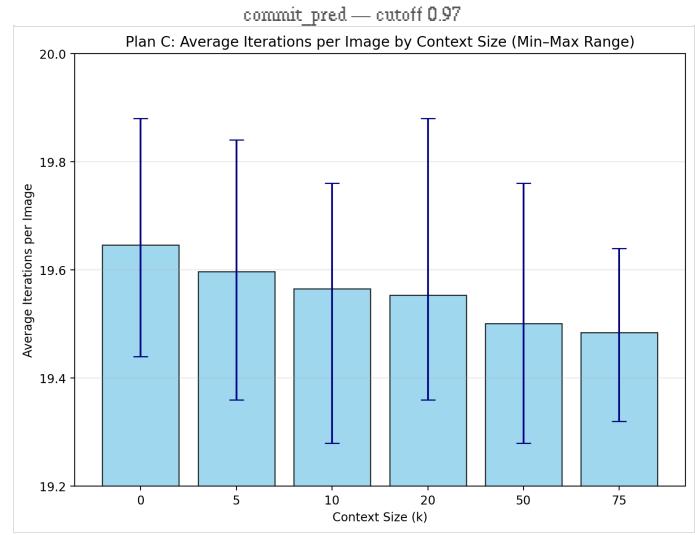
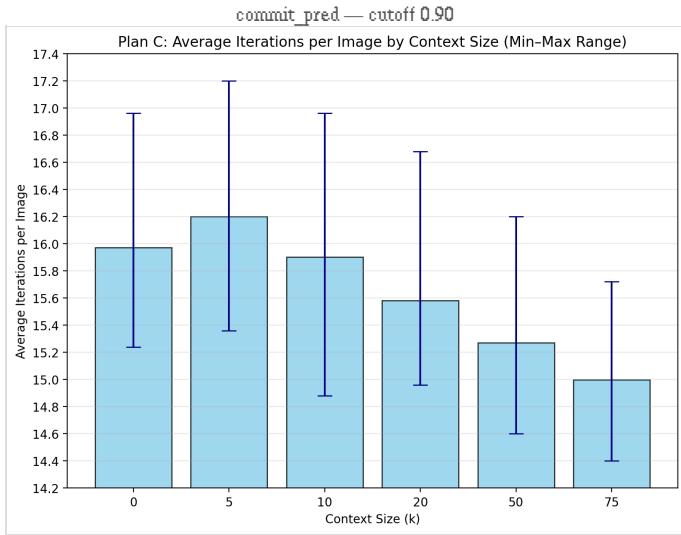
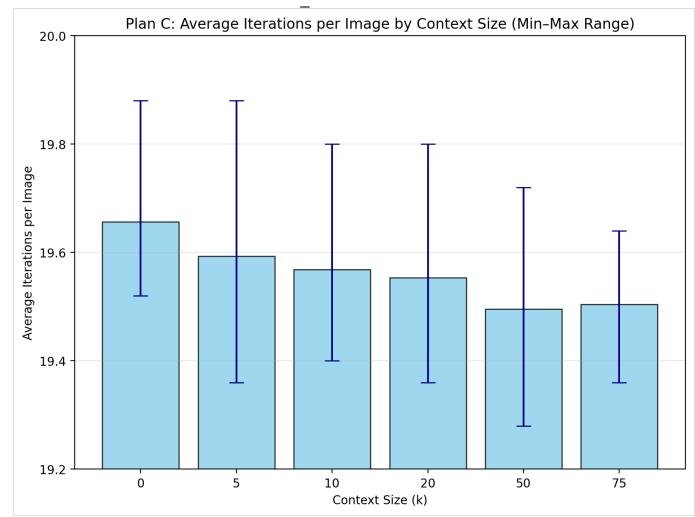
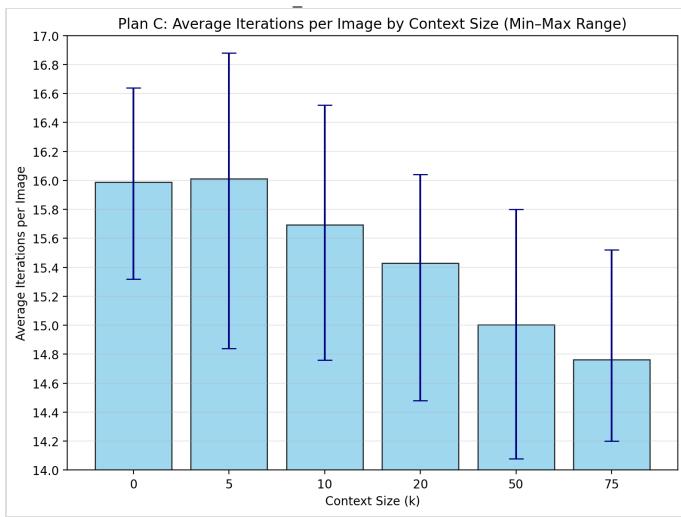
plan_c_avg_iterations_iqr.png



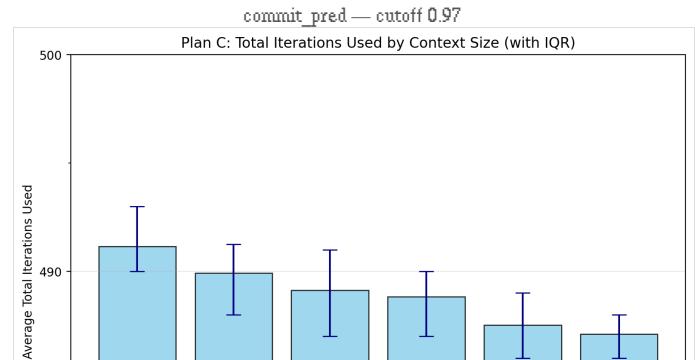
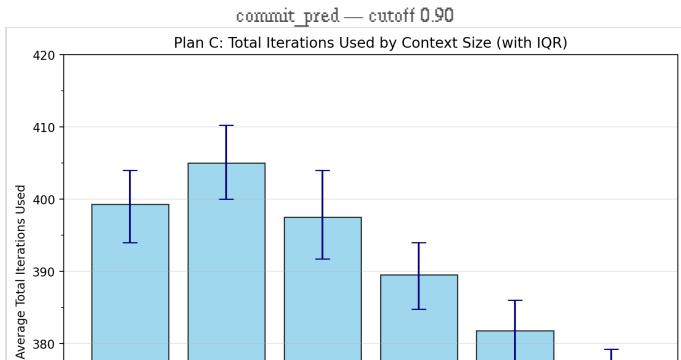
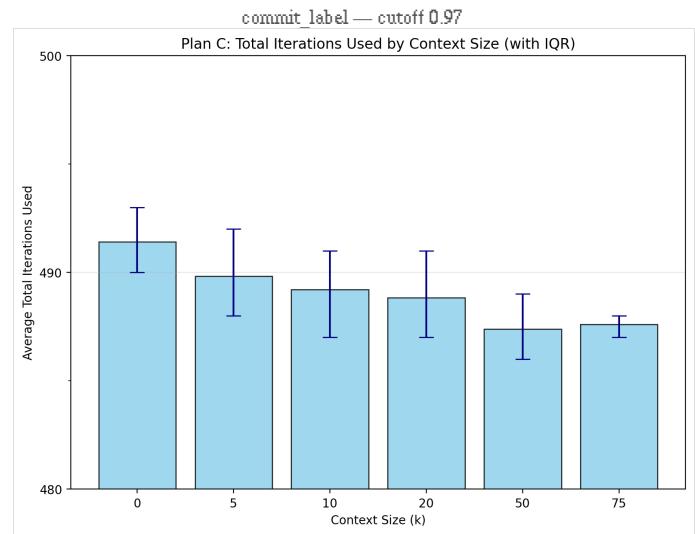
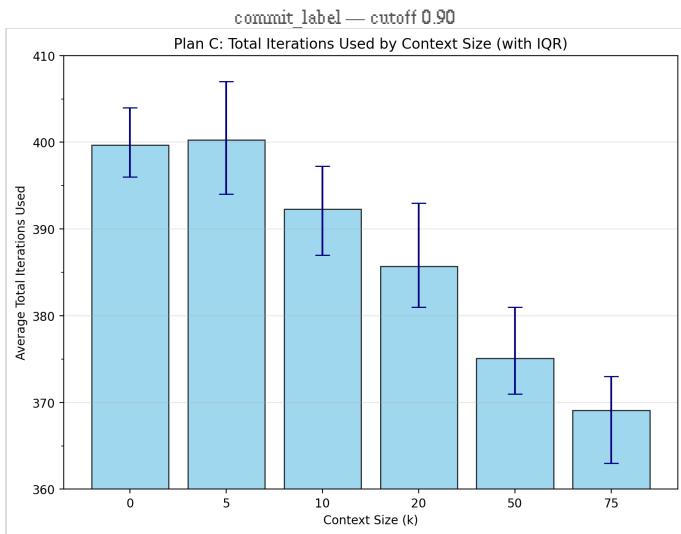
plan_c_avg_iterations_range.png

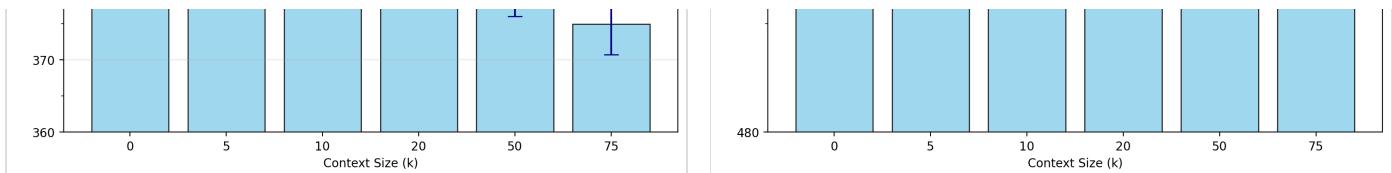
commit_label — cutoff 0.90

commit_label — cutoff 0.97



plan_c_total_iterations.png





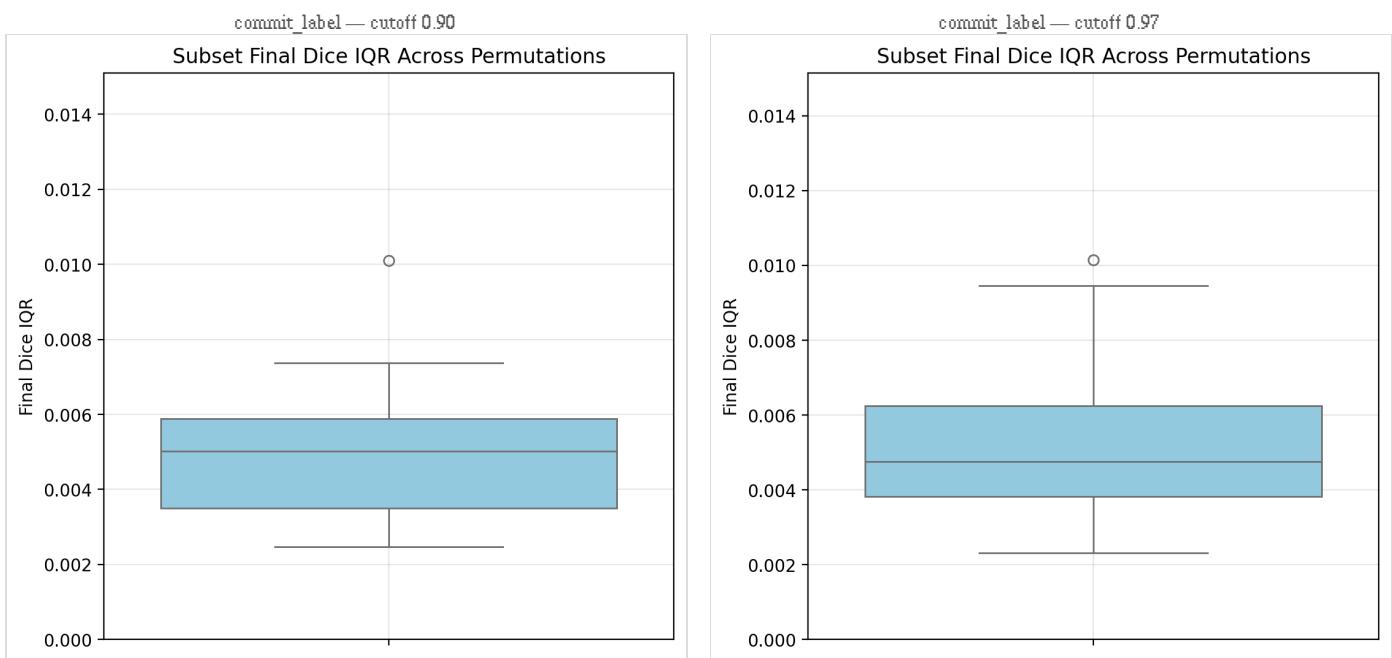
Plan B — Interleaved Label/Pred

In [4]:

```
label_map = build_pairs_map(plan_b_figures(paths['commit_label_90']), plan_b_figures(paths['commit_label_97']))
pred_map = build_pairs_map(plan_b_figures(paths['commit_pred_90']), plan_b_figures(paths['commit_pred_97']))
show_interleaved('Plan B — Interleaved (commit_label & commit_pred)', label_map, pred_map)
```

Plan B — Interleaved (commit_label & commit_pred)

`subset_final_dice_iqr.png`



`commit_label — cutoff 0.90`

`commit_label — cutoff 0.97`

Subset Final Dice IQR Across Permutations

`commit_pred — cutoff 0.90`

`commit_pred — cutoff 0.97`

Subset Final Dice IQR Across Permutations

`subset_final_dice_range.png`

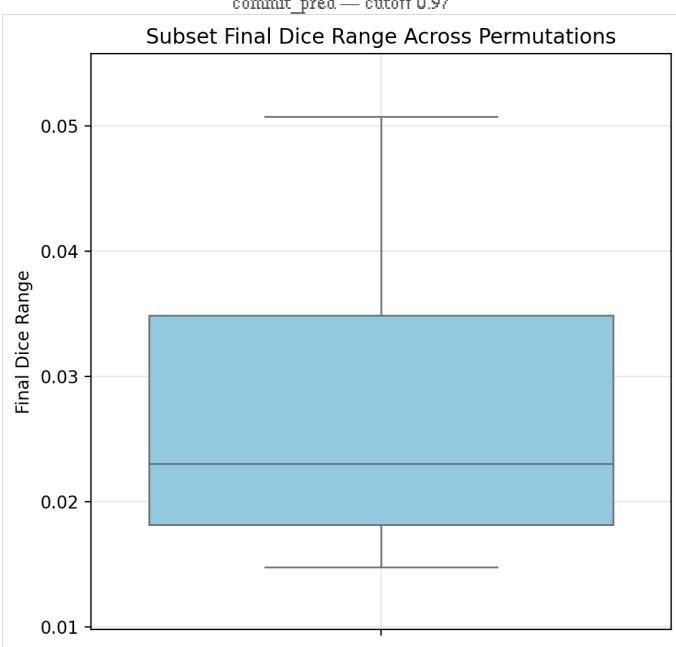
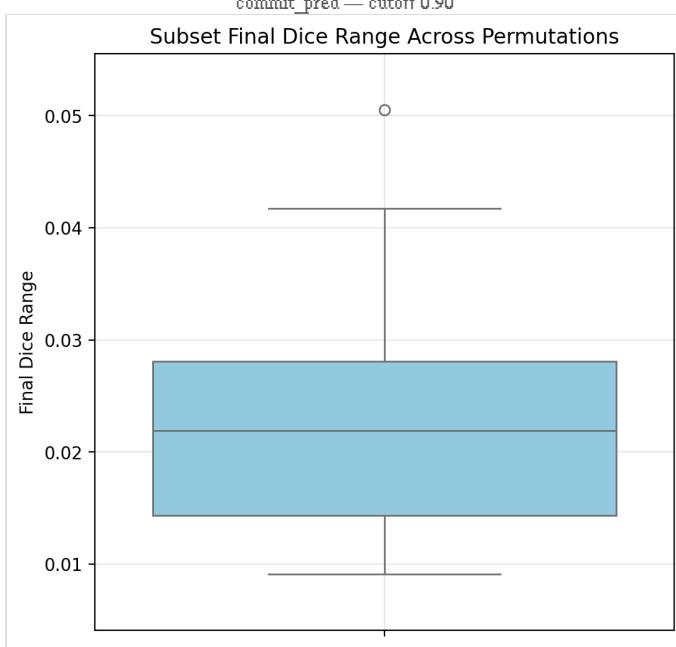
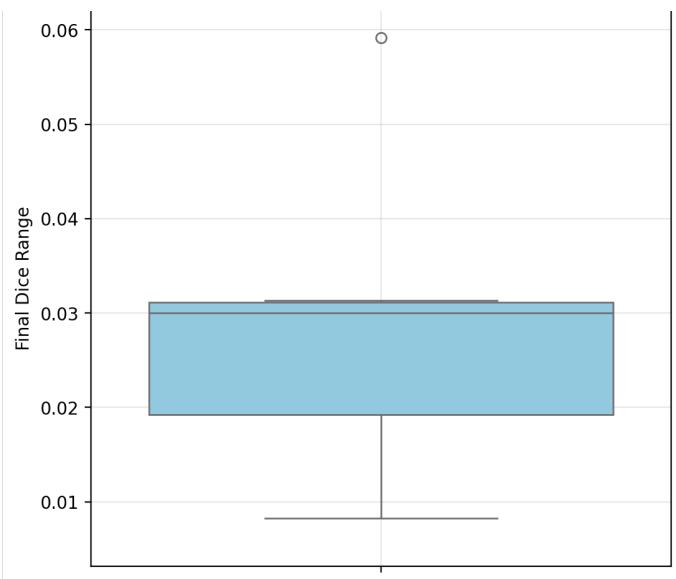
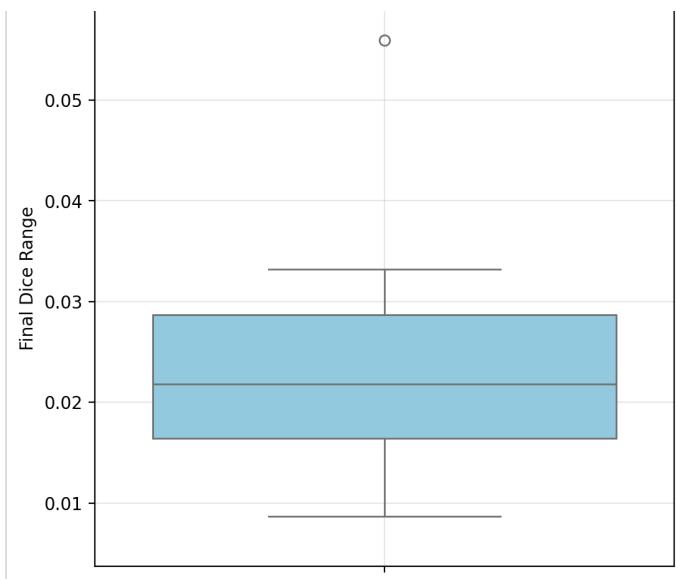
`commit_label — cutoff 0.90`

`commit_label — cutoff 0.97`

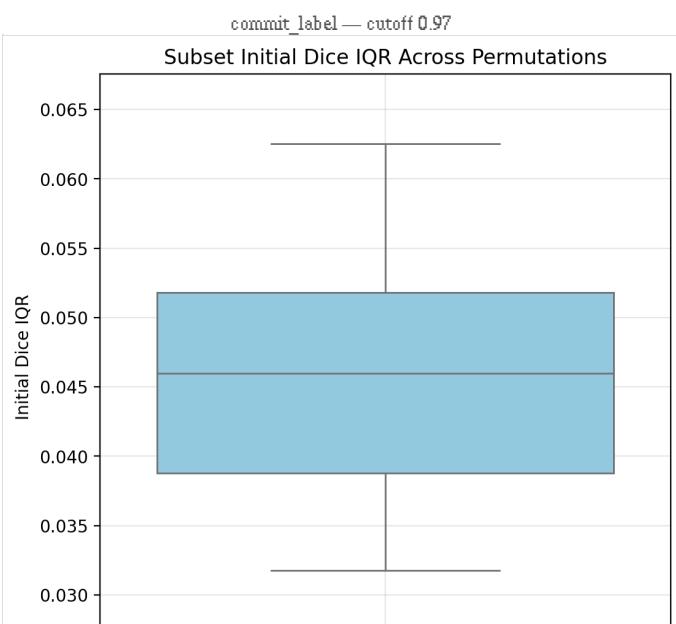
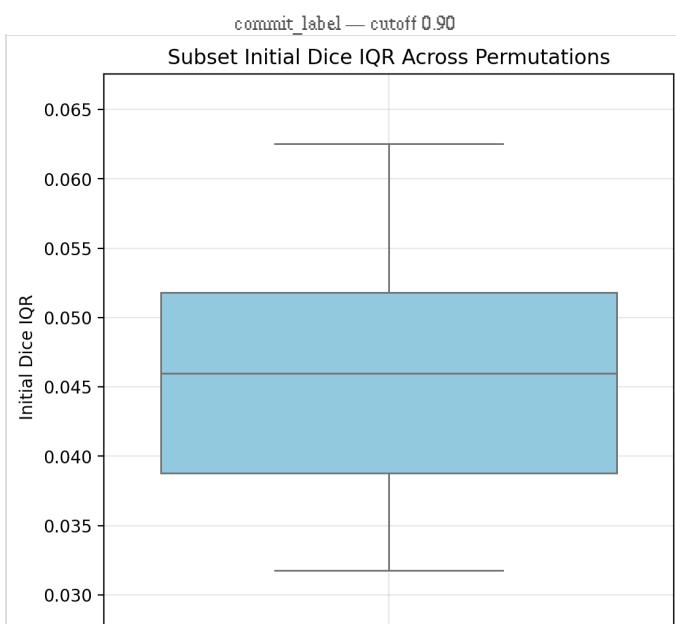
Subset Final Dice Range Across Permutations

Subset Final Dice Range Across Permutations

0.06

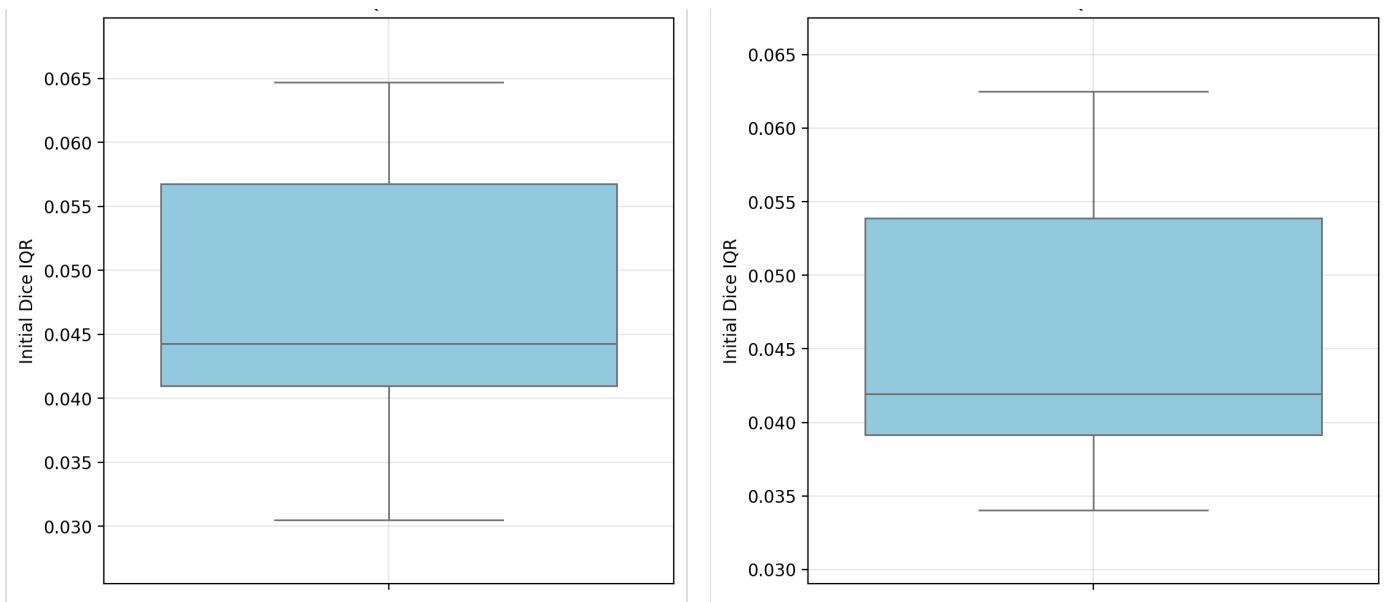


subset_initial_dice_iqr.png

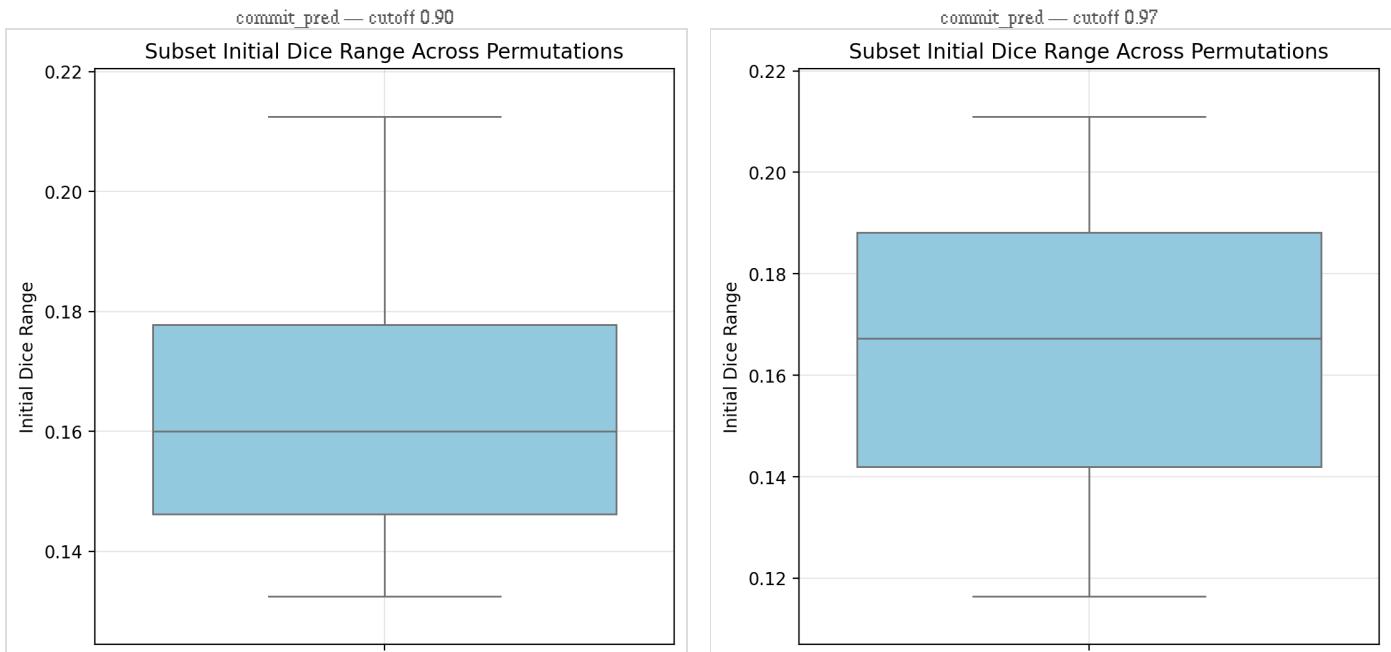
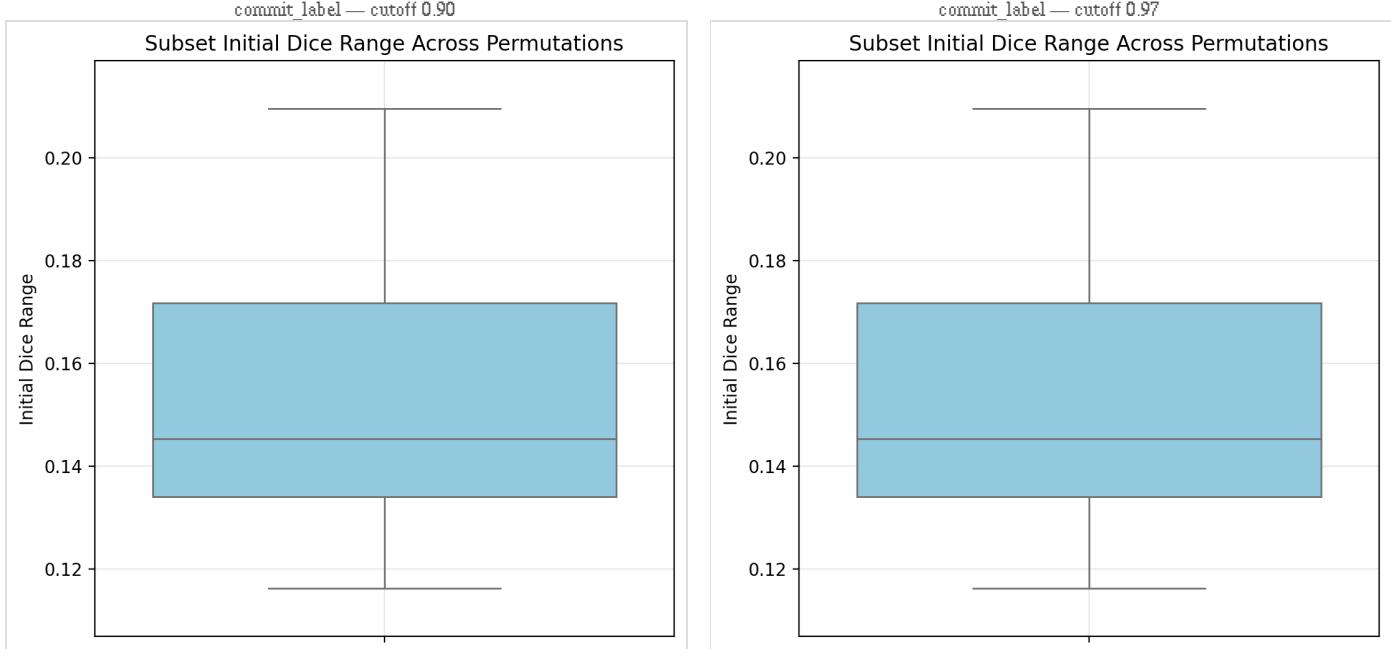


commit_pred — cutoff 0.90
Subset Initial Dice IQR Across Permutations

commit_pred — cutoff 0.97
Subset Initial Dice IQR Across Permutations

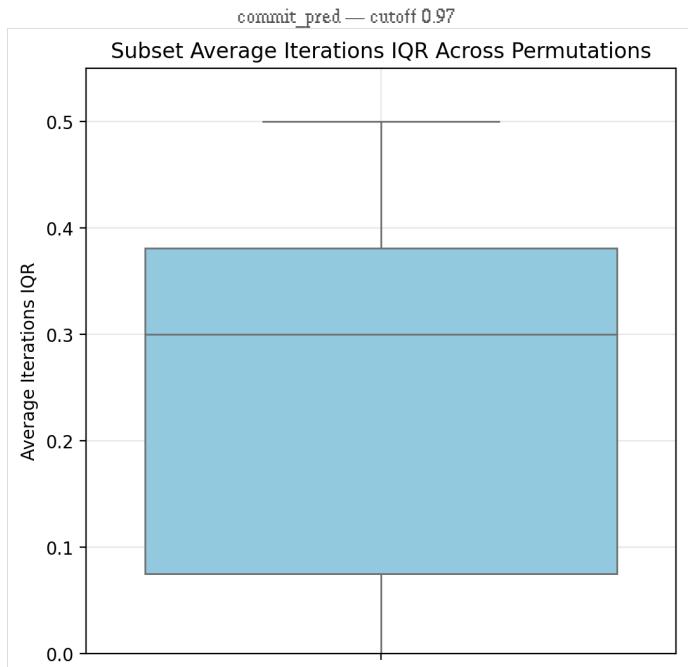
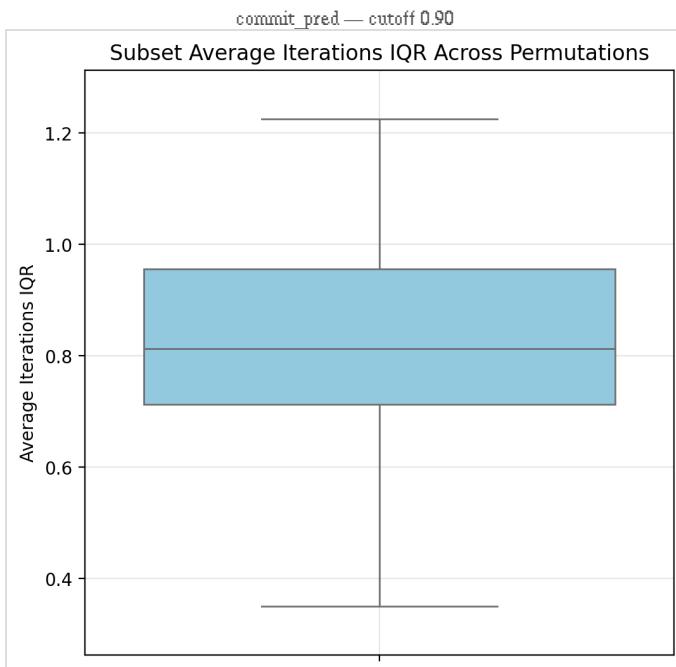
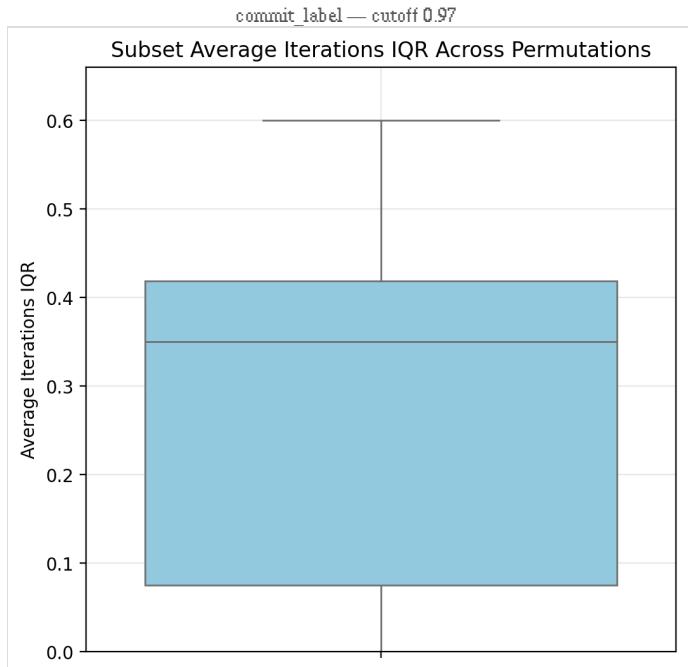
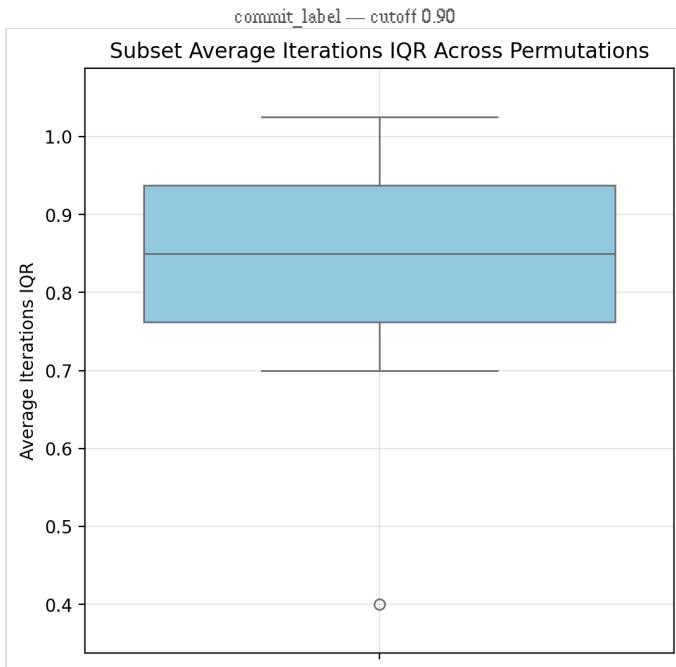


subset_initial_dice_range.png

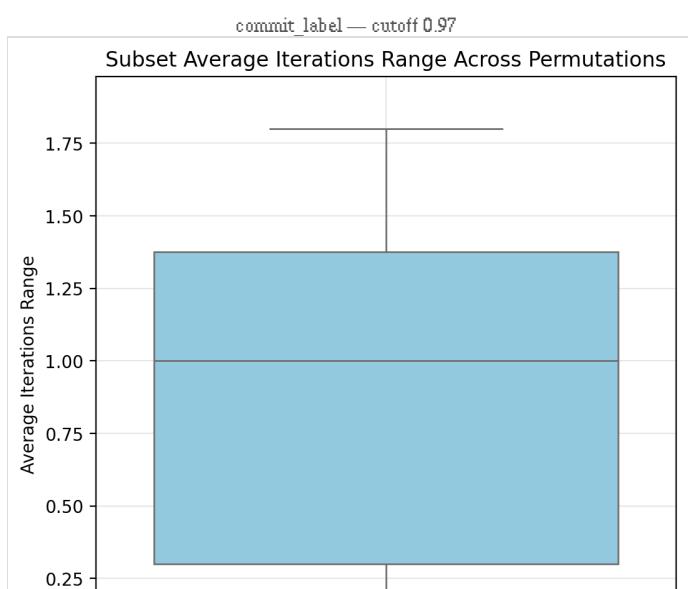
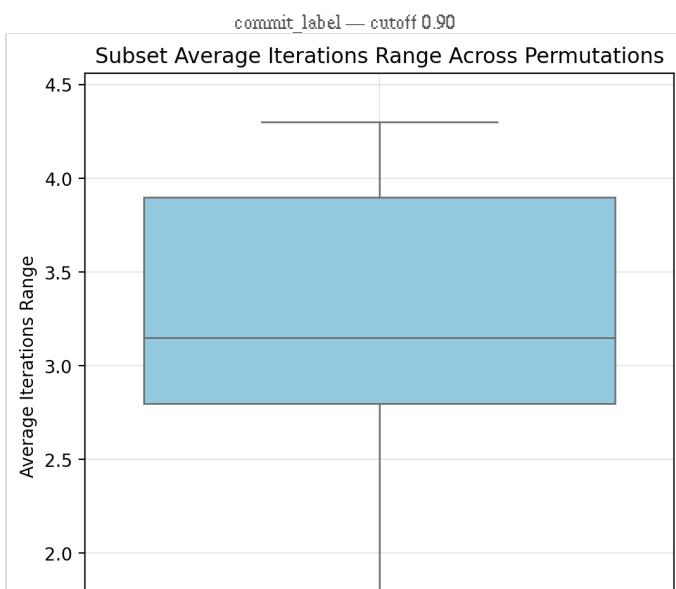


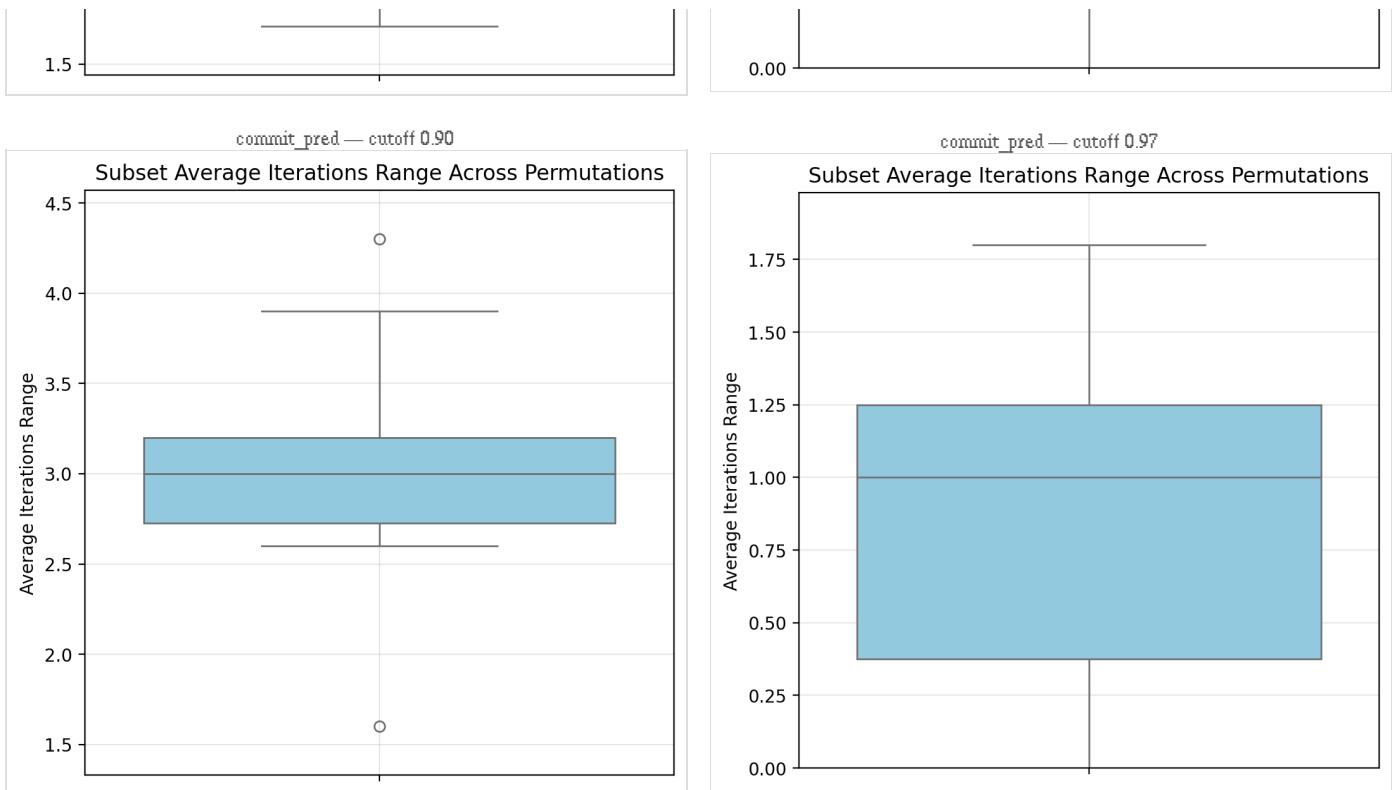
subset_iterations_iqr.png

subset_iterations_iqr.png

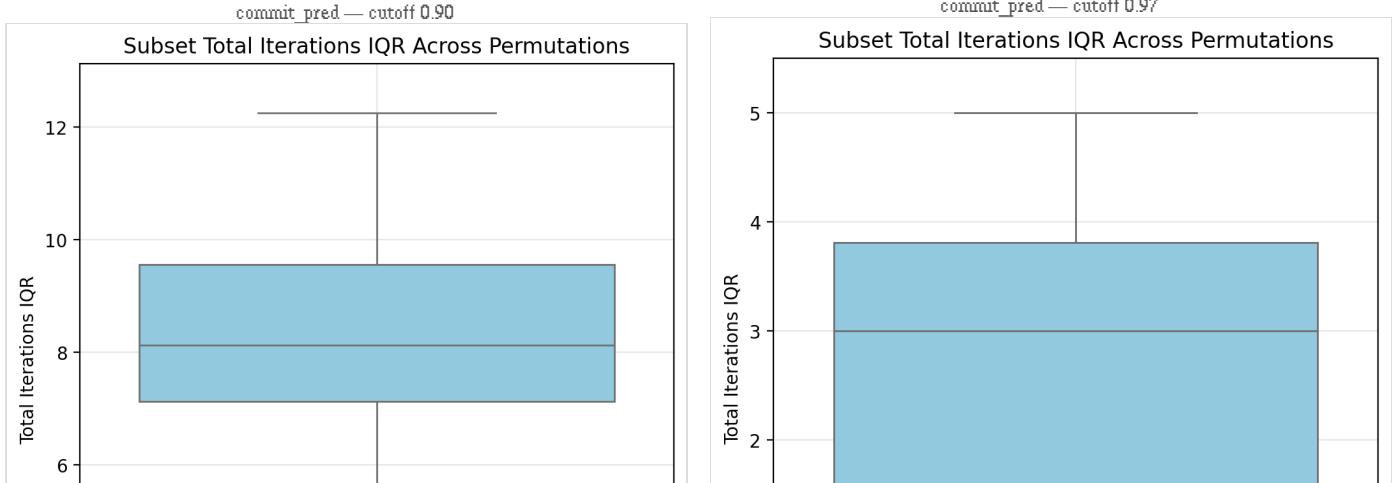
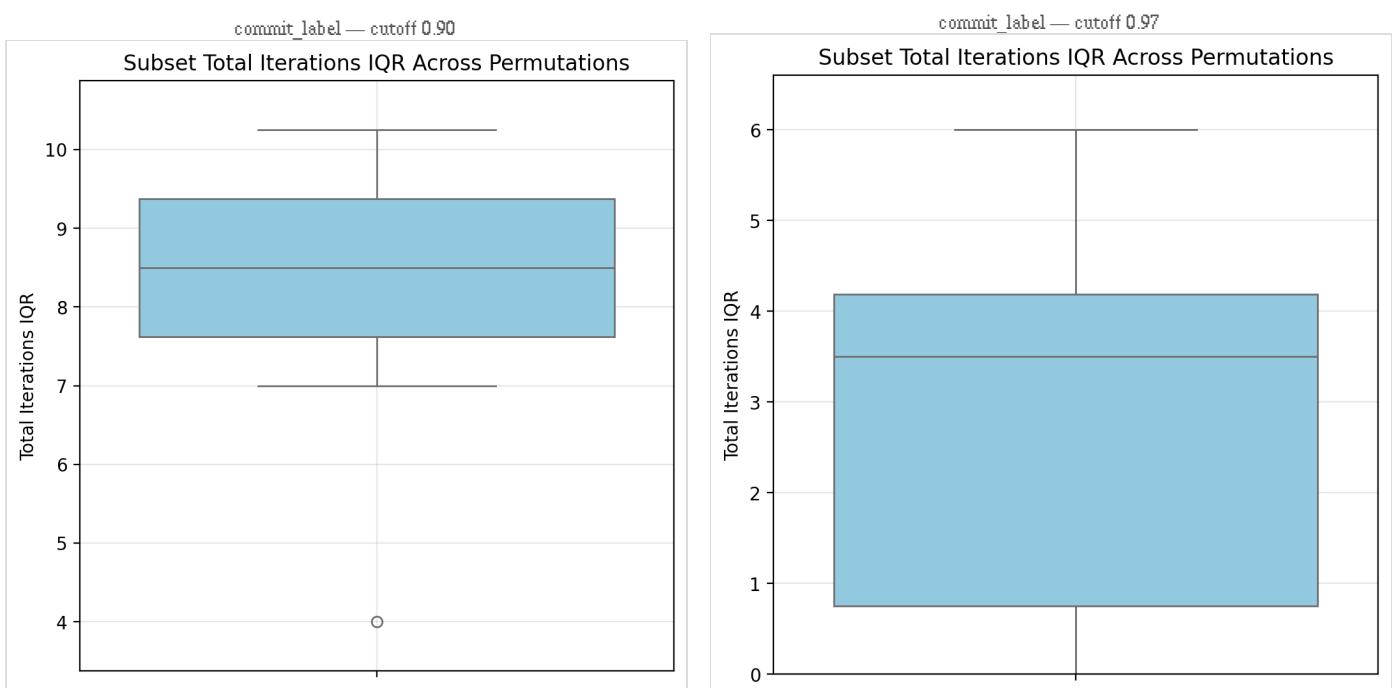


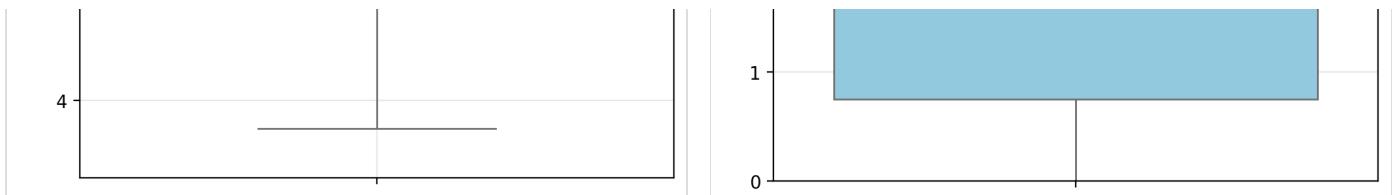
subset_iterations_range.png





subset_iterations_total_iqr.png





subset_iterations_total_range.png

