Results Overview

• Designs generated: 1

• Designs passing all filters: 0

• MMCIF files of final designs are in:

output/final_ranked_designs/final_1_designs

• Metrics and sequences of 1 final designs are in:

output/final_ranked_designs/final_designs_metrics_1.csv

• Metrics of all designs are in:

output/final_ranked_designs/all_designs_metrics.csv

You can rerun filtering (very quick), using this command with changed parameters:

```
-- boltzgen run input_spec.yaml --steps filtering --config filtering budget=60 alpha=0.05
```

You can also rerun filtering in a jupyter notebook if you want using `filter.ipynb`

What was run to produce this in the Filter task:

1. Filtering: each design is evaluated against mandatory thresholds.
2. Ranking: for every metric we compute its rank, then scale it by the metric's inverse-importance weight.

Designs with fewer passed filters are automatically penalised because the ranking key is the pair (num_filters_passed, metric). The overall quality score is the worst (maximum) of these scaled ranks.

The 10 best designs form the Top set.

3. Diversity: a lazy-greedy algorithm selects 1 designs that jointly maximise quality and minimise sequence similarity (sequence-identity distance). The trade-off is controlled by α = 0.001:

• α = 0 → 100 % quality focus (same as Top set)

• α = 1 → 100 % diversity focus (ignores quality)

• Quality – composite of metrics such as iPTM, salt-bridges, ΔSASA, etc. Each metric has an "inverse importance" weight (see "Sorting Criteria" table). A larger weight divides the rank by a bigger number and therefore down-weights that metric.

• Diversity – 1 − sequence identity between designs.

We use α = 0.001, meaning 100 % emphasis on quality, 0 % on diversity.

CSV Column Reference

Column	Description
id	filename to retrieve the file
design_sequence	designed amino acids (may be subset of chain)
designed_chain_sequence	full sequence of the chain containing designed residues (recommended for synthesis)
num_design	number of designed residues
secondary_rank	intermediate rank from the sorting procedure
design_ptm	predicted TM score for intra-design contacts (higher = better)
design_iptm	predicted TM score for design–target contacts (higher = better)
design_to_target_iptm	same as design_iptm but for multi-chain designs
min_design_to_target_pae	minimum PAE between design & target (lower = better)
plip_saltbridge	number of salt-bridge interactions
plip_hbonds	number of hydrogen-bond interactions
plip_hydrophobic	number of hydrophobic interactions
delta_sasa_original	ΔSASA when binder present vs absent
delta_sasa_refolded	same as above but on the refolded structure

Filtering & Sorting Criteria

This section first lists threshold filters that every design must pass (Filtering Criteria table)then explains how designs are ranked (Sorting Criteria table) weighted by their inverse importance.

Filtering Criteria

feature	lower_is_better	threshold	Pass
has_x	True	0.0	1
filter_rmsd	True	2.5	0
filter_rmsd_design	True	2.5	0
designfolding-filter_rmsd	True	2.5	0
ALA_fraction	True	0.3	1
GLY_fraction	True	0.2	1
GLU_fraction	True	0.2	1
LEU_fraction	True	0.3	1
VAL_fraction	True	0.2	1

Sorting Criteria

Metric	Inverse Importance
design_to_target_iptm	1.1
design_ptm	1.1
neg_min_design_to_target_pae	1.1
affinity_probability_binary1	1
plip_hbonds_refolded	2
plip_saltbridge_refolded	2
delta_sasa_refolded	2

Results Summary

Aggregate statistics

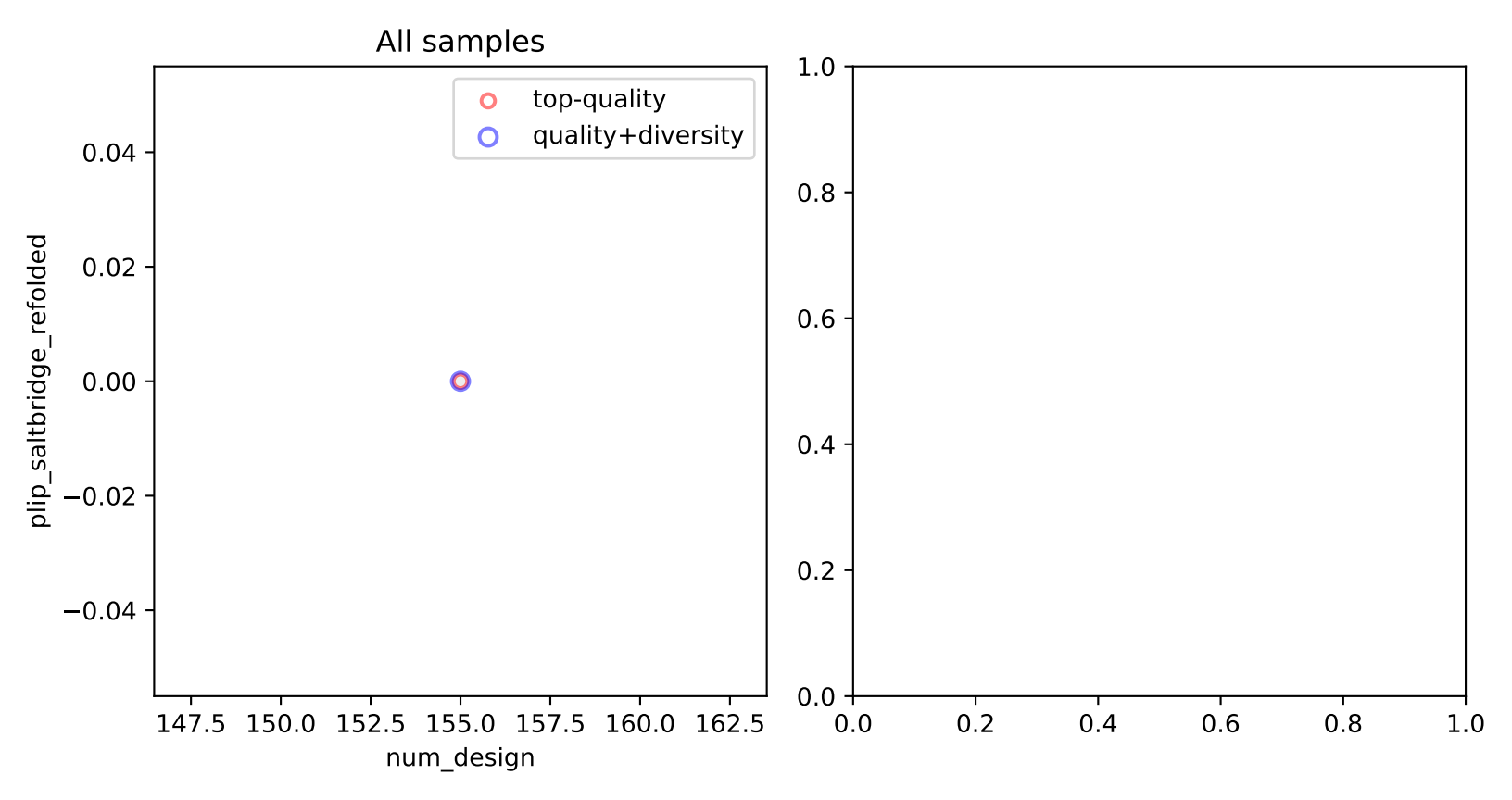
Quick numeric overview comparing (i) ALL incoming designs, (ii) the top-10 highest-quality designs, and (iii) the 1 quality+diversity designs produced by the lazy-greedy selection.

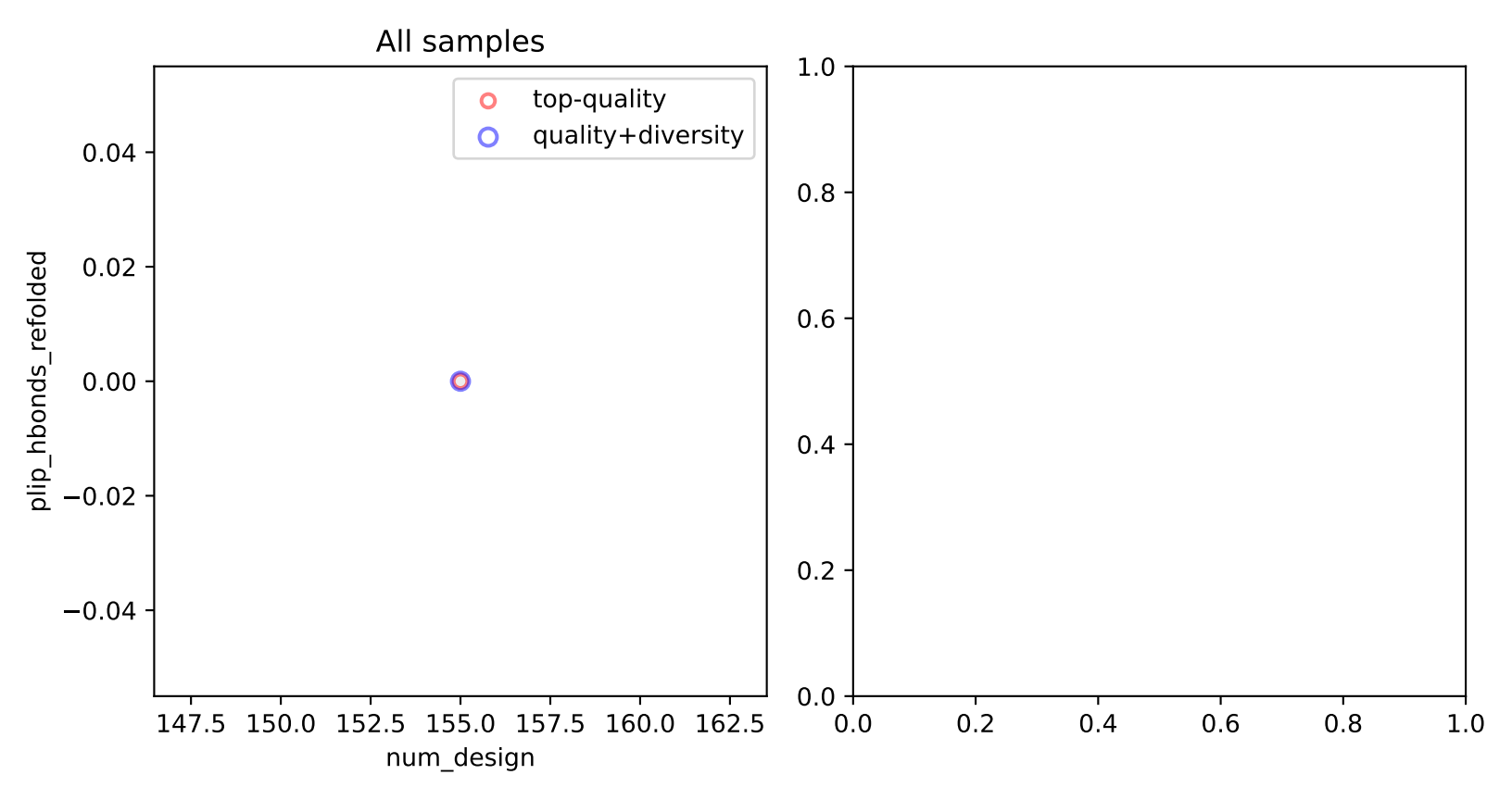
Metric	Mean	Mean top 10	Mean top 1 diverse
Num designs	1	-	-
num_design	155.000	155.000	155.000
filter_rmsd	16.738	16.738	16.738
affinity_probability_binary1	0.506	0.506	0.506
designfolding-filter_rmsd	14.524	14.524	14.524
design_iptm	0.255	0.255	0.255
design_to_target_iptm	0.255	0.255	0.255
min_design_to_target_pae	11.863	11.863	11.863
delta_sasa_refolded	374.507	374.507	374.507
plip_saltbridge_refolded	0.000	0.000	0.000
plip_hbonds_refolded	0.000	0.000	0.000

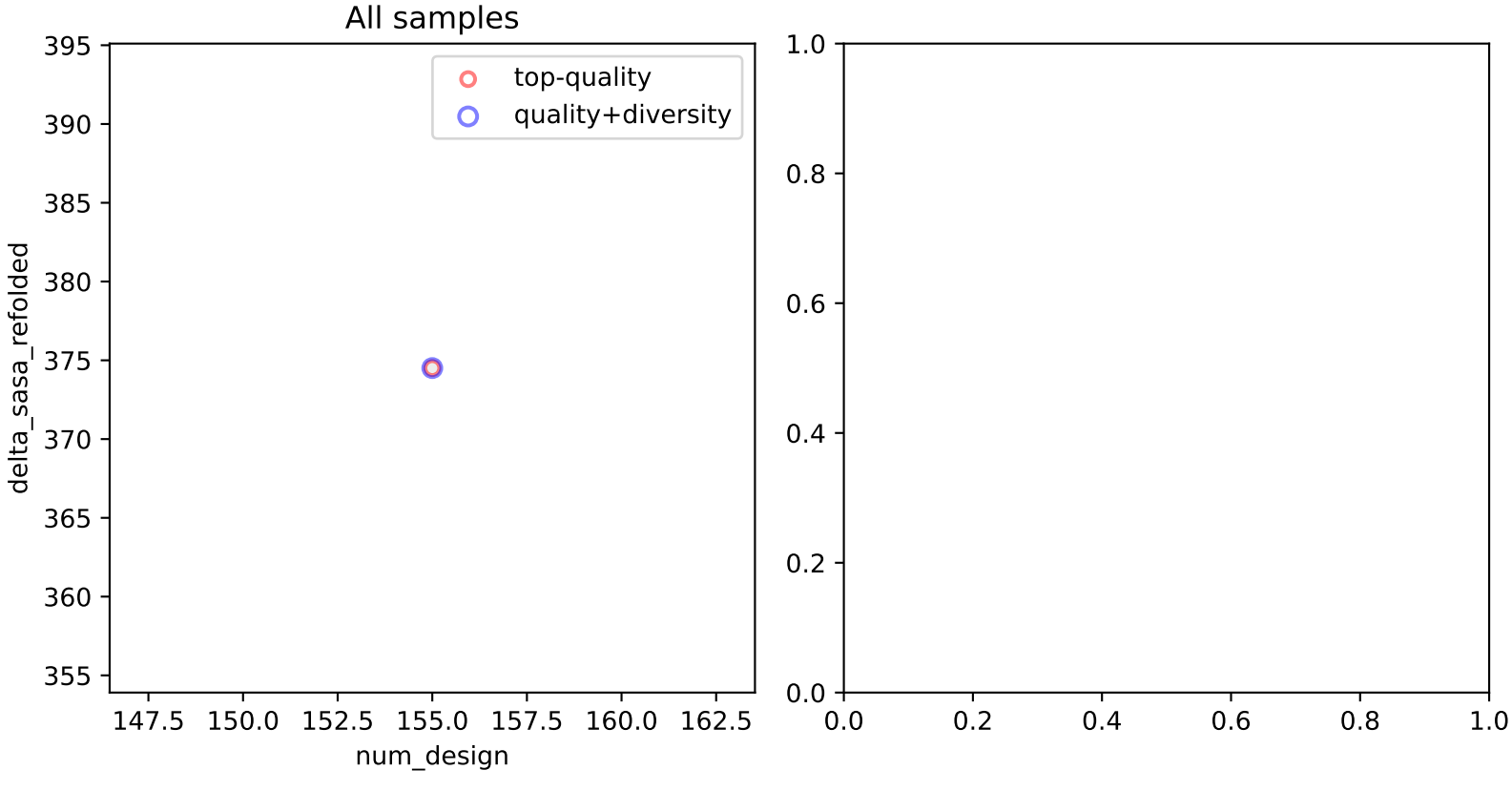
Scatter Plots

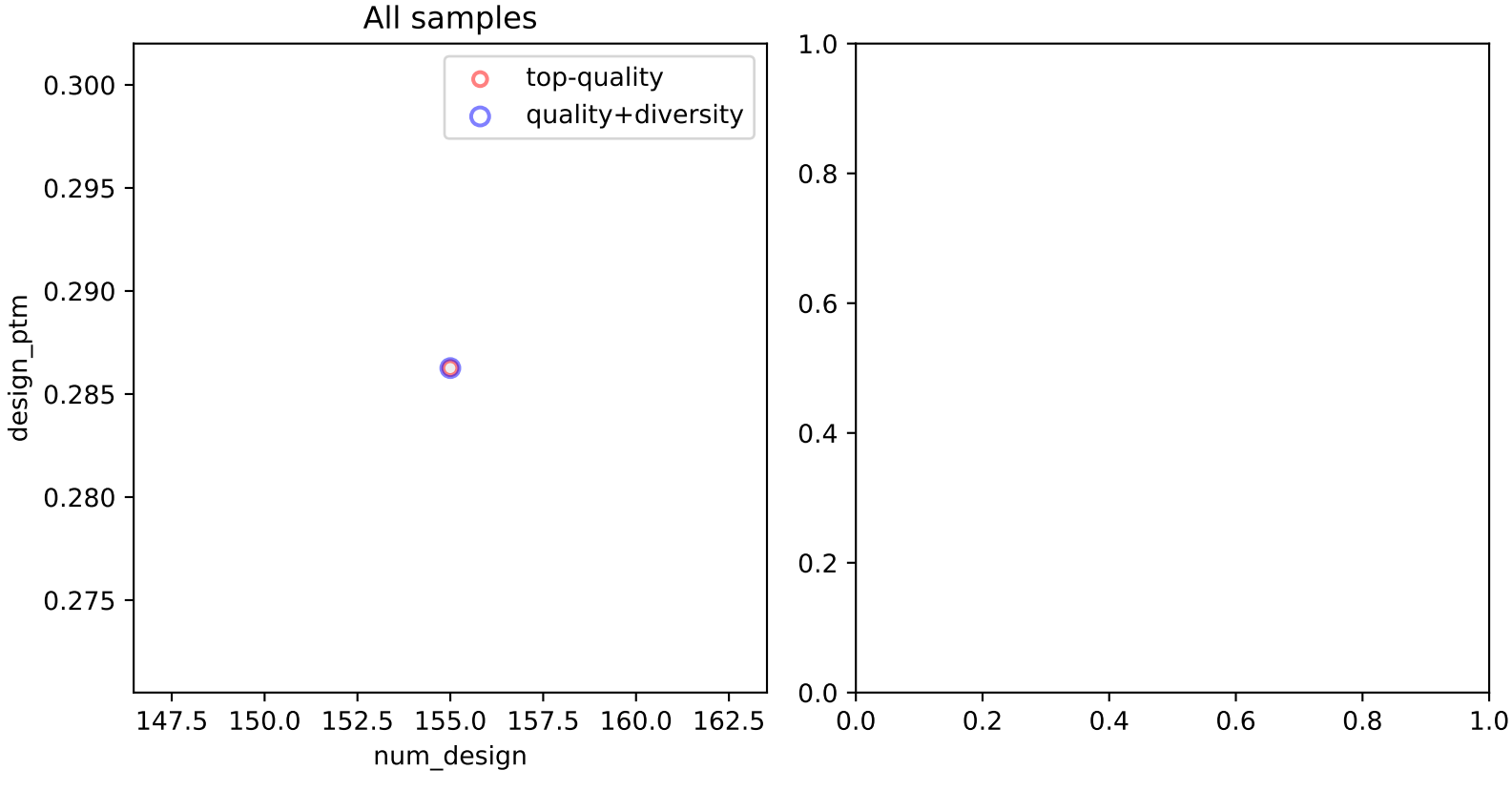
Metric relationships

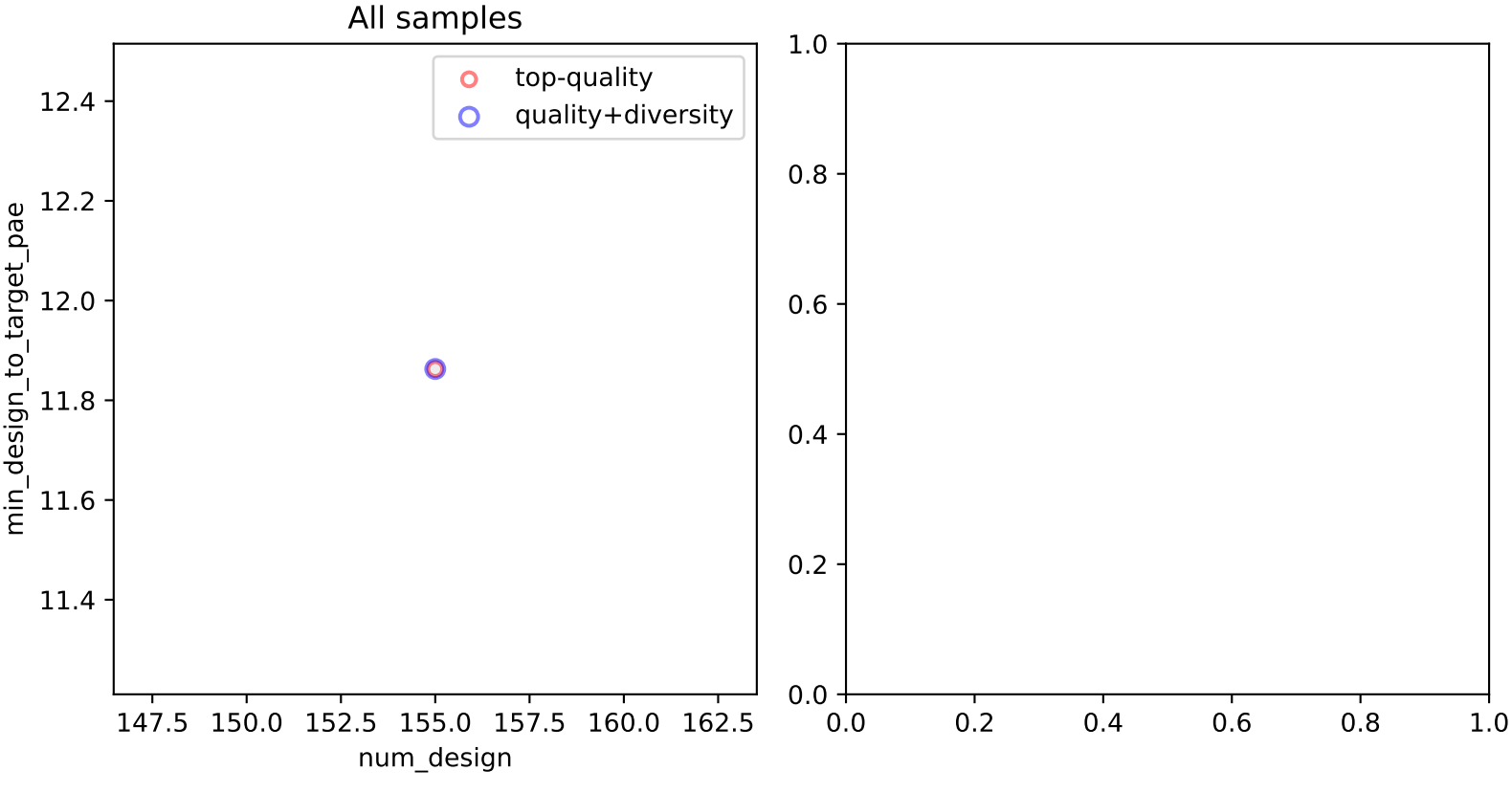
Each scatter page contains two panels:
• Left – all designs (grey) with overlays of Top (red) and Diverse (blue).
• Right – same but limited to designs passing the RMSD filter.

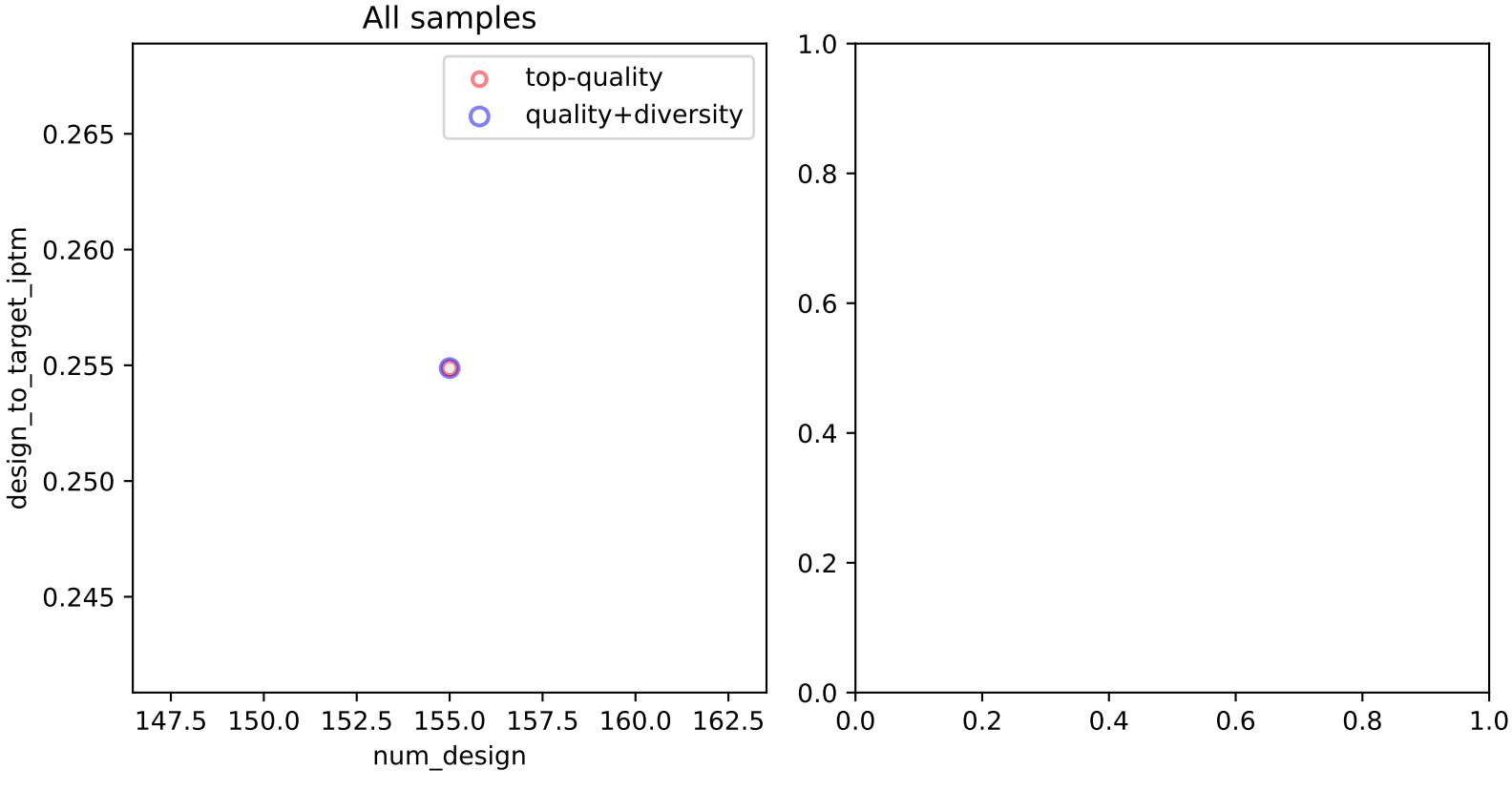

All samples
top-quality
quality+diversity
plip_saltbridge_refolded
num_design
0.04
0.02
0.00
−0.02
−0.04
147.5
150.0
152.5
155.0
157.5
160.0
162.5
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

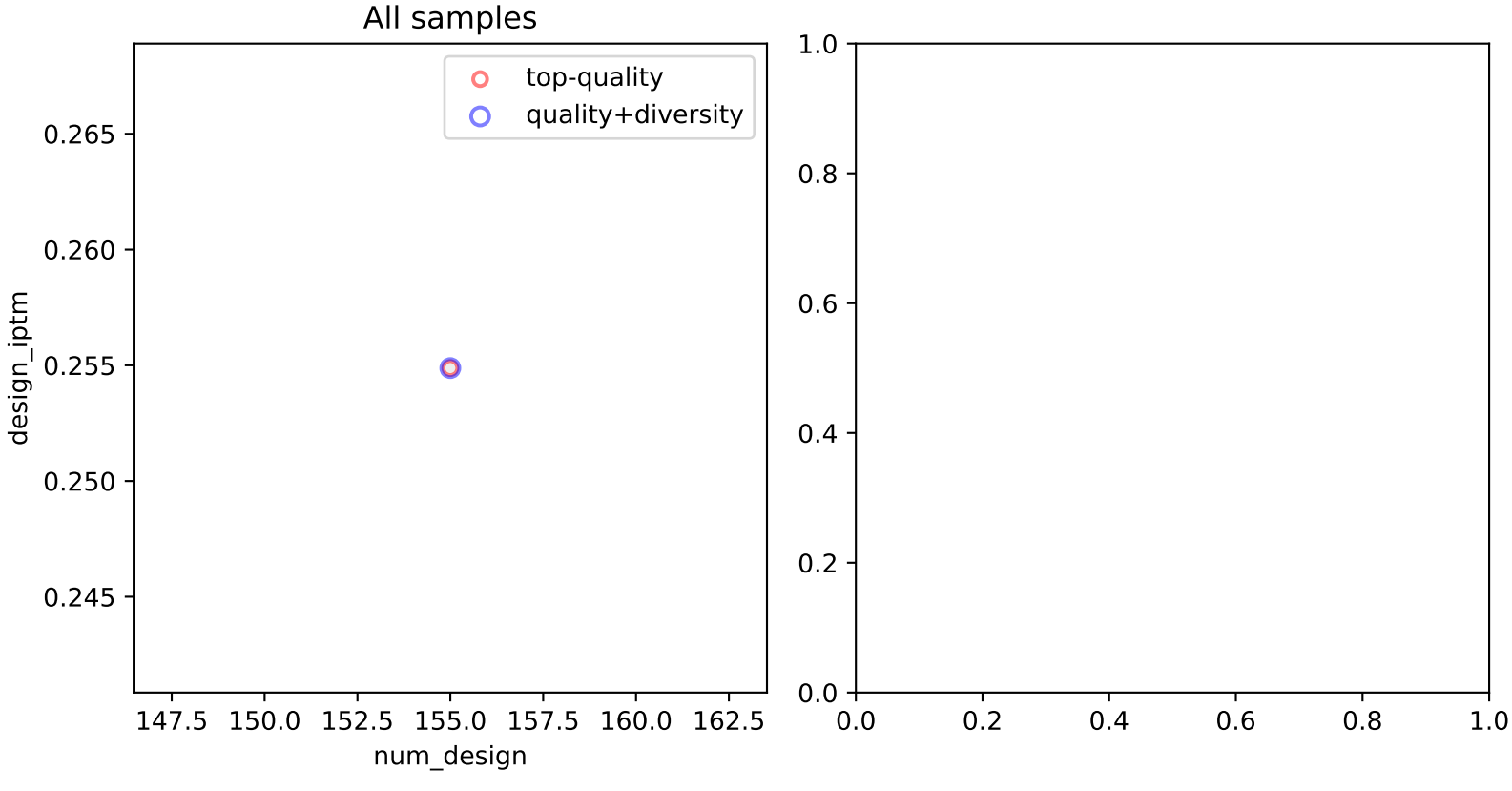

All samples
top-quality
quality+diversity
plip_hbonds_refolded
0.04
0.02
0.00
−0.02
−0.04
147.5
150.0
152.5
155.0
157.5
160.0
162.5
num_design
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

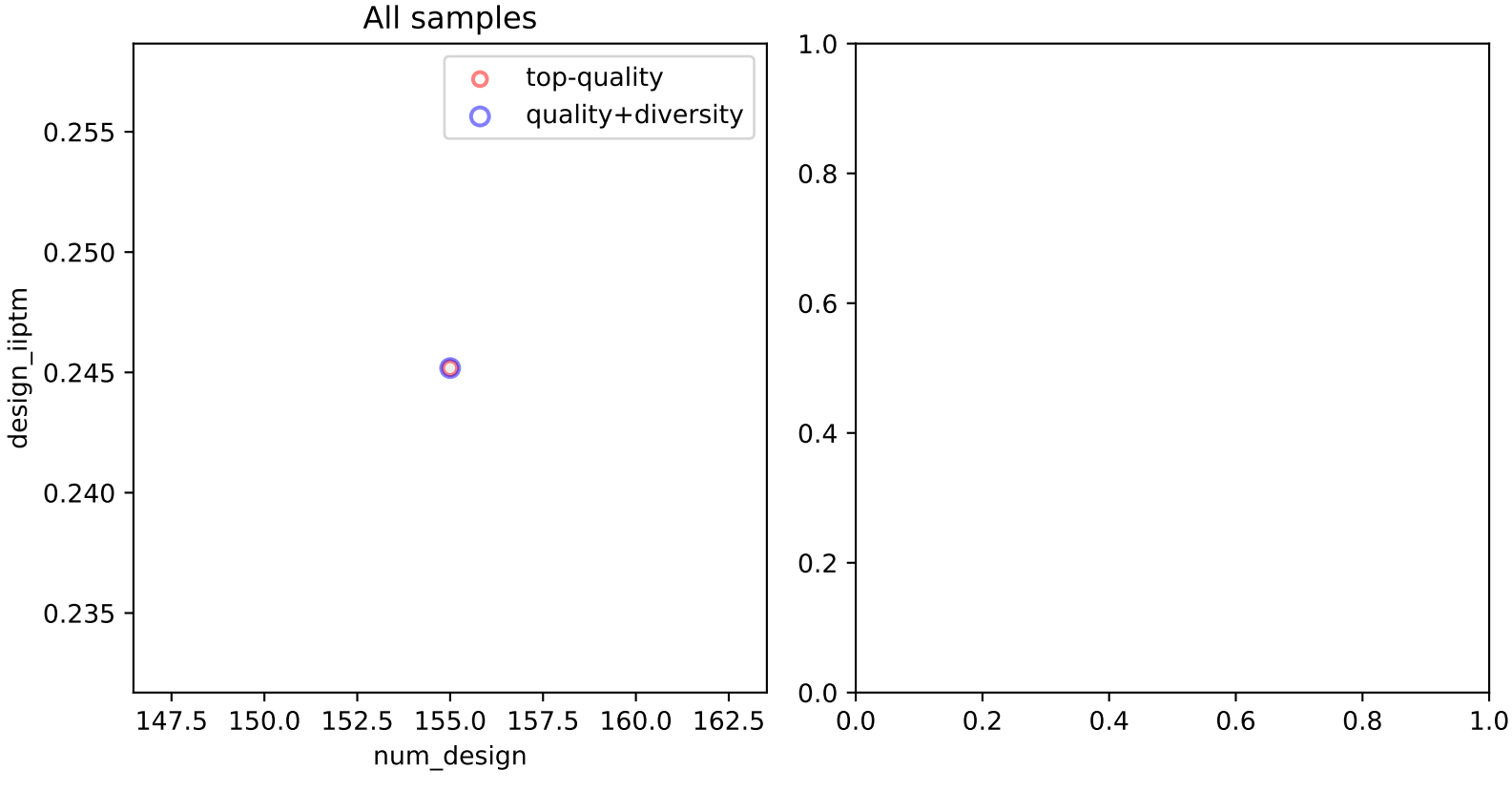

All samples
top-quality
quality+diversity
delta_sasa_refolded
num_design
395
390
385
380
375
370
365
360
355
147.5
150.0
152.5
155.0
157.5
160.0
162.5
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

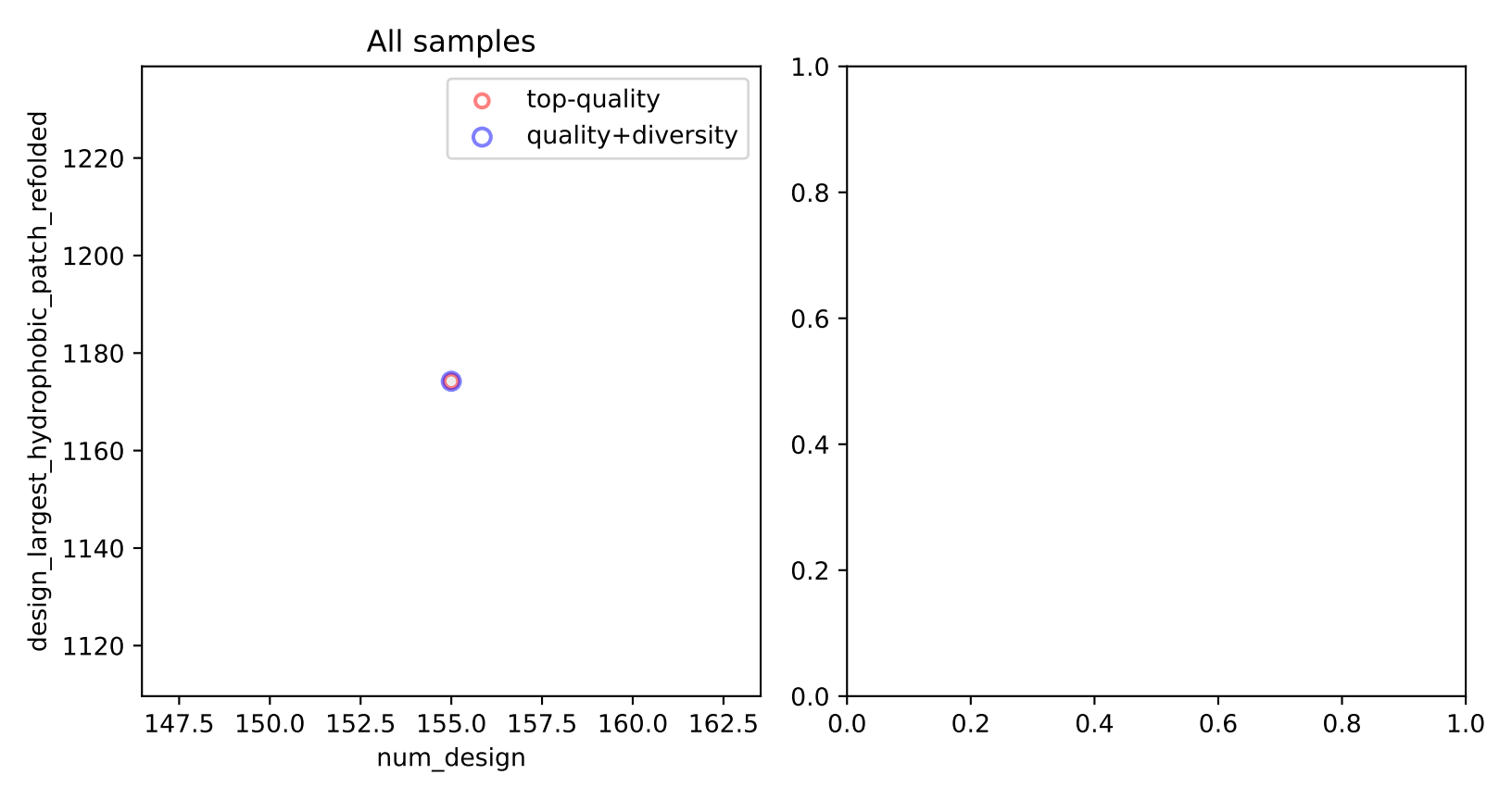

All samples
top-quality
quality+diversity
design_ptm
num_design
0.300
0.295
0.290
0.285
0.280
0.275
147.5
150.0
152.5
155.0
157.5
160.0
162.5
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

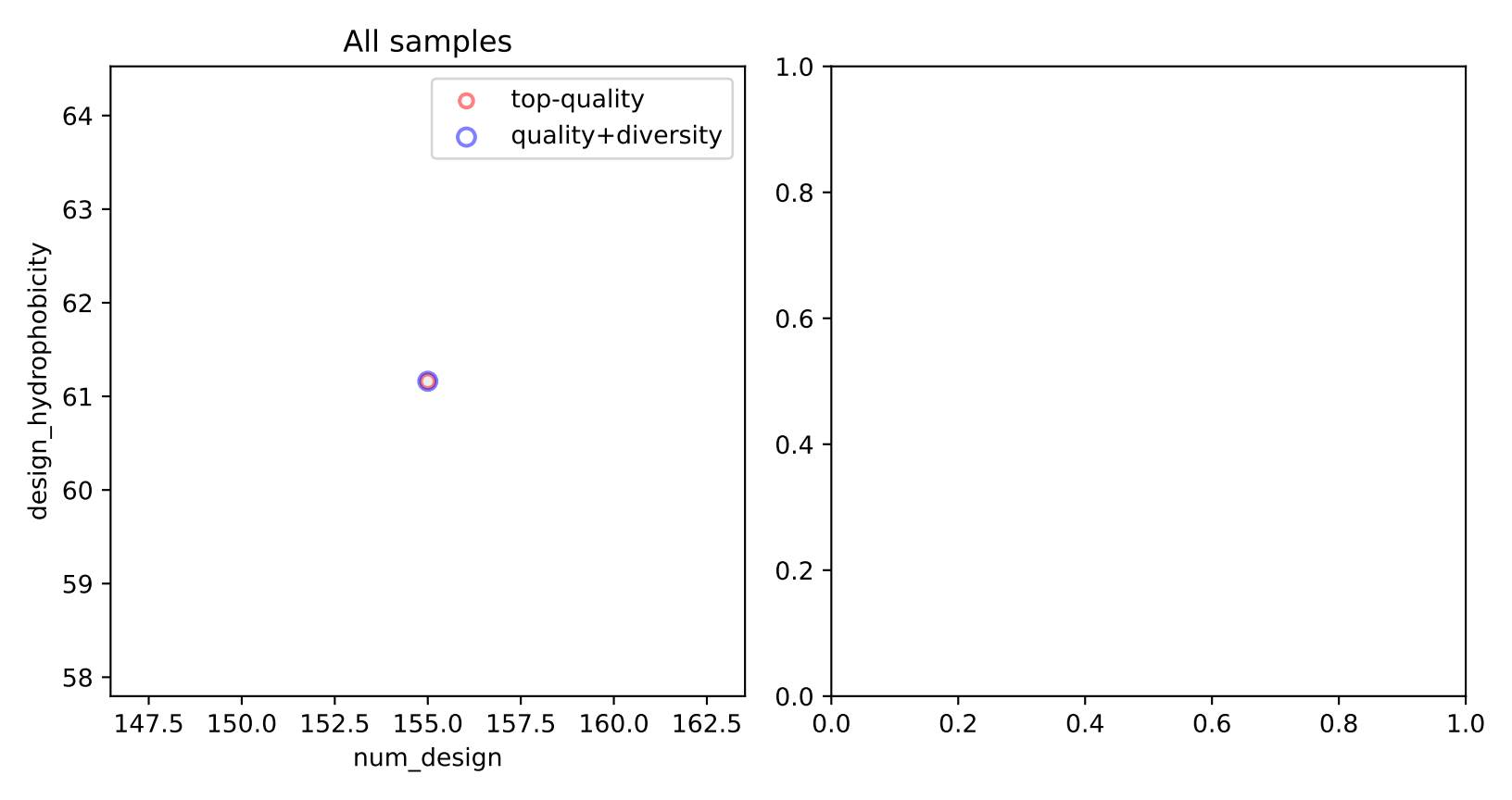

All samples
top-quality
quality+diversity
min_design_to_target_pae
num_design
12.4
12.2
12.0
11.8
11.6
11.4
147.5
150.0
152.5
155.0
157.5
160.0
162.5
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


All samples
top-quality
quality+diversity
design_to_target_iptm
0.265
0.260
0.255
0.250
0.245
147.5
150.0
152.5
155.0
157.5
160.0
162.5
num_design
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


All samples
top-quality
quality+diversity
design_iptm
num_design
0.265
0.260
0.255
0.250
0.245
147.5
150.0
152.5
155.0
157.5
160.0
162.5
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


All samples
top-quality
quality+diversity
design_iiptm
num_design
0.255
0.250
0.245
0.240
0.235
147.5
150.0
152.5
155.0
157.5
160.0
162.5
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

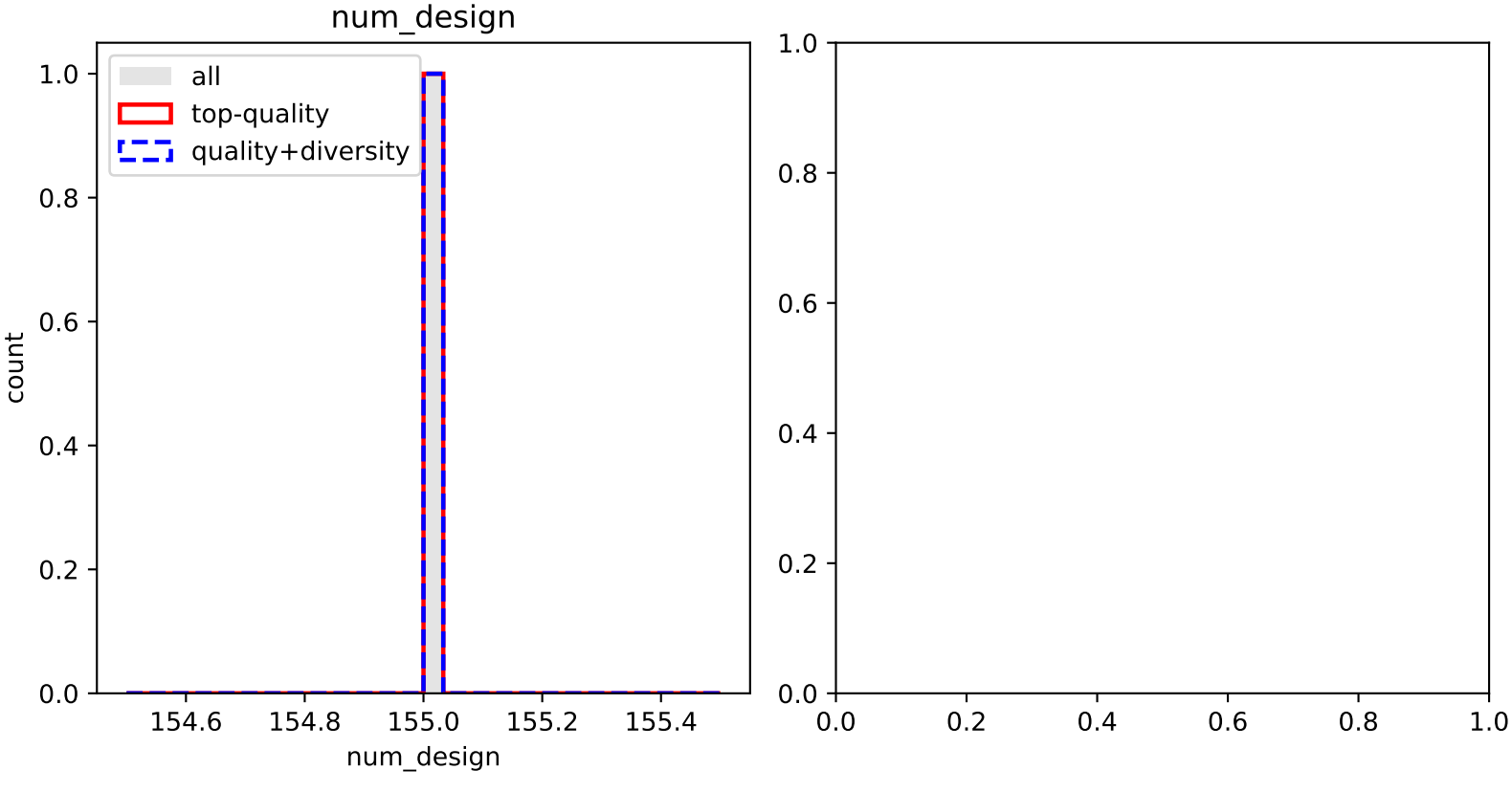

All samples
top-quality
quality+diversity
design_largest_hydrophobic_patch_refolded
1220
1200
1180
1160
1140
1120
147.5
150.0
152.5
155.0
157.5
160.0
162.5
num_design
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

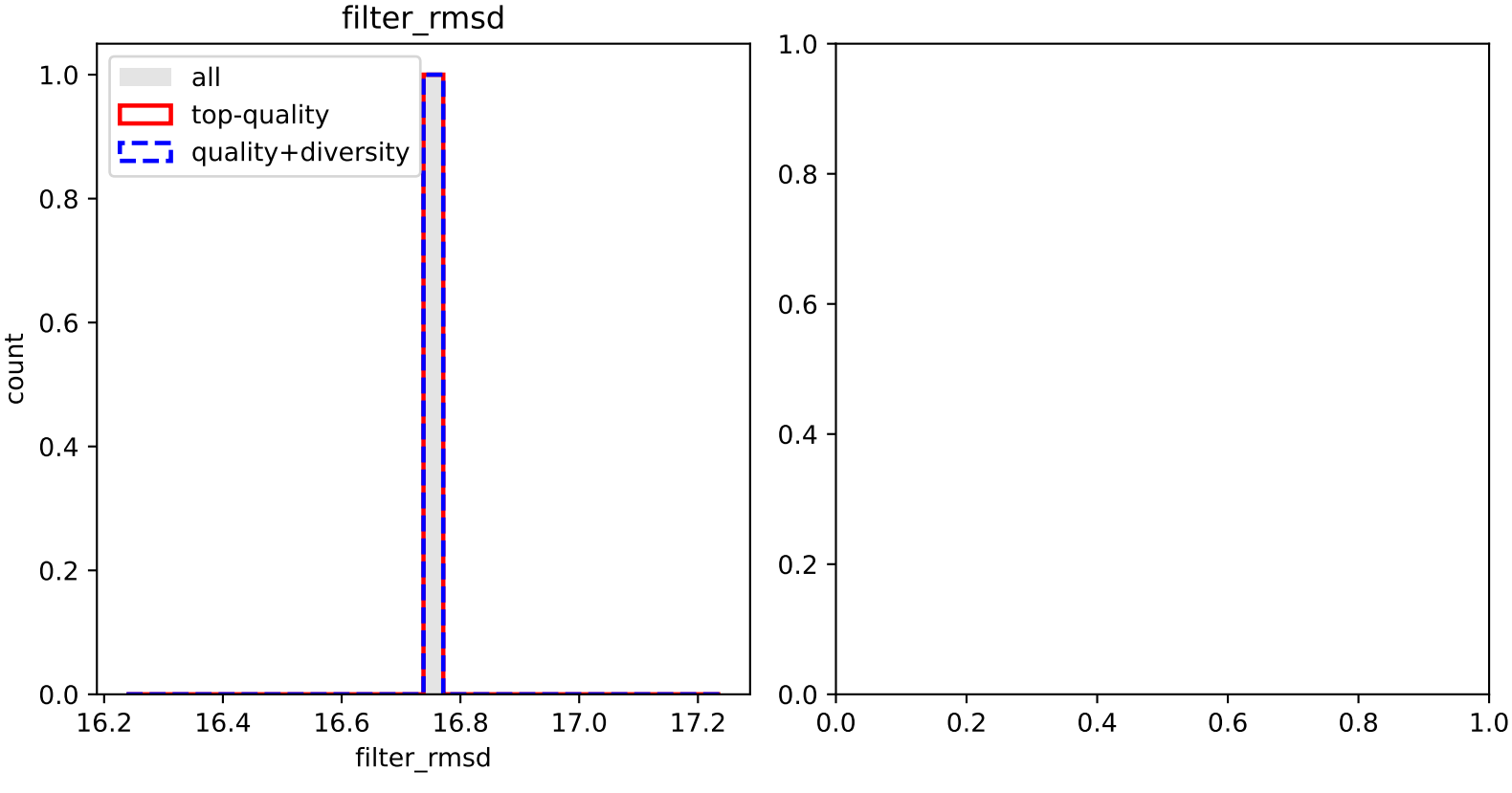

All samples
top-quality
quality+diversity
design_hydrophobicity
num_design
64
63
62
61
60
59
58
147.5
150.0
152.5
155.0
157.5
160.0
162.5
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

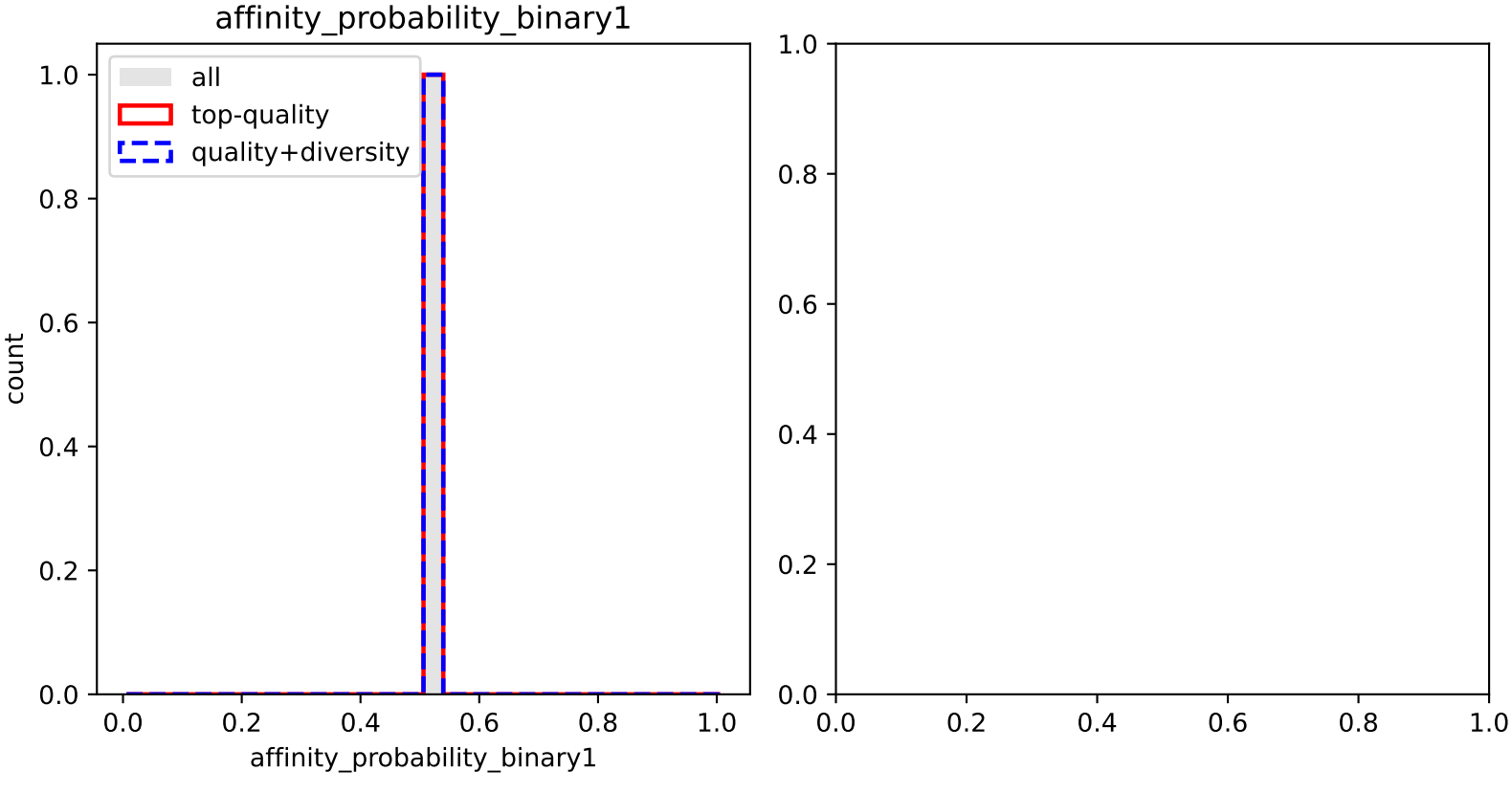
Metric Distributions

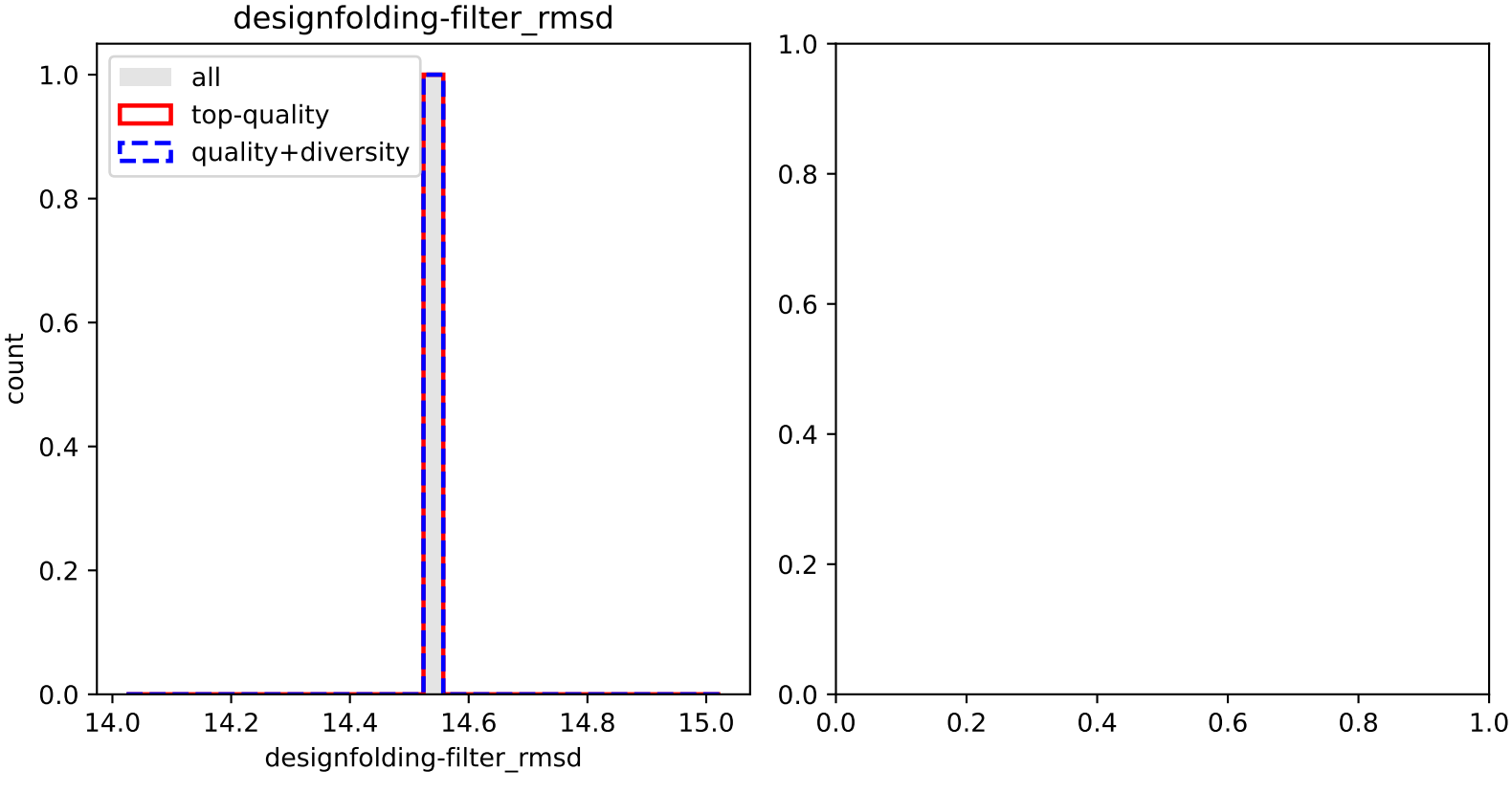
Histograms

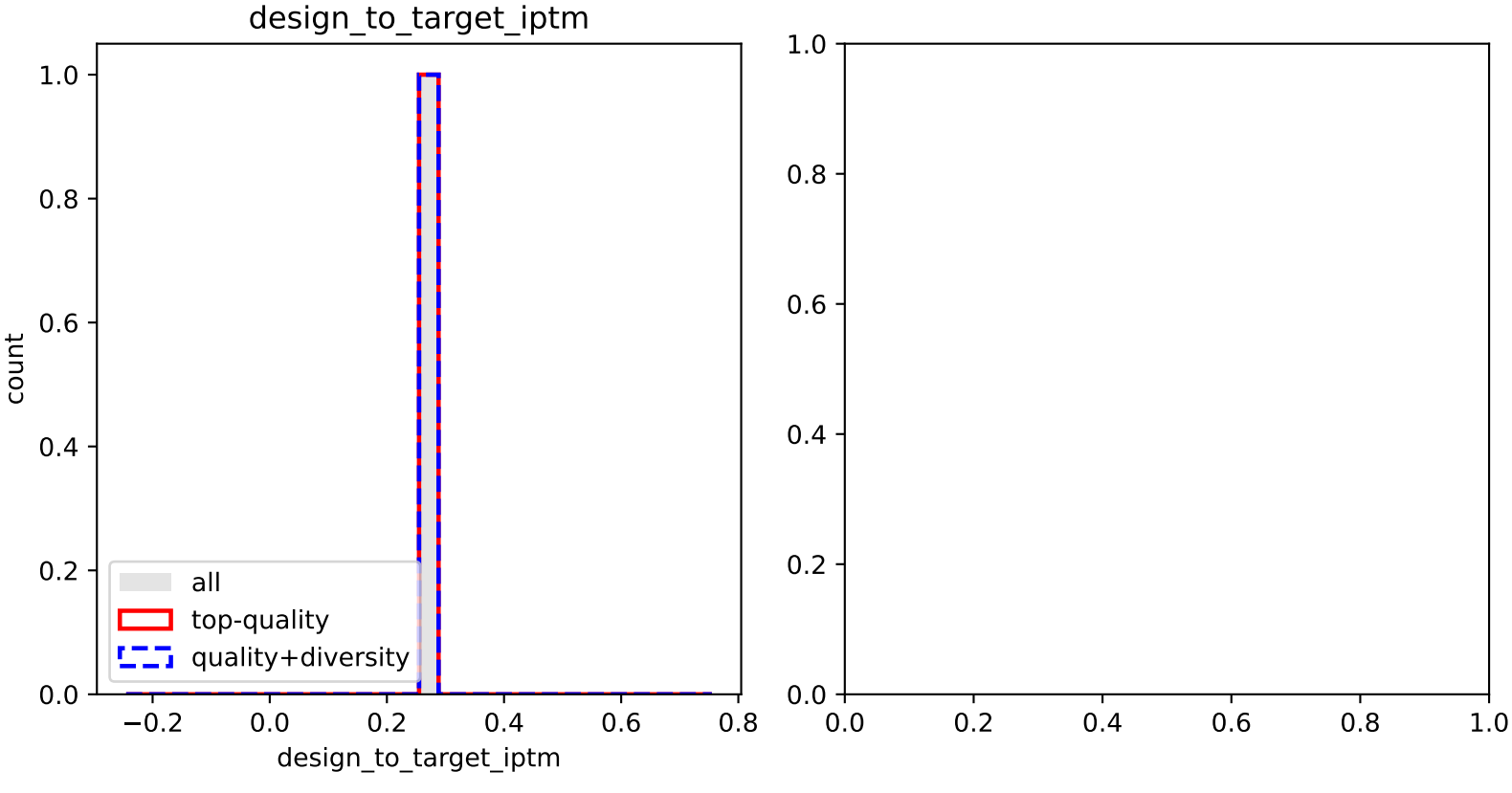
Distribution of each metric across all designs (grey) with overlays for Top-quality (red outline) and Diversity-optimised (blue dashed) subsets. The right panel repeats the histogram but only for designs that pass the RMSD threshold.

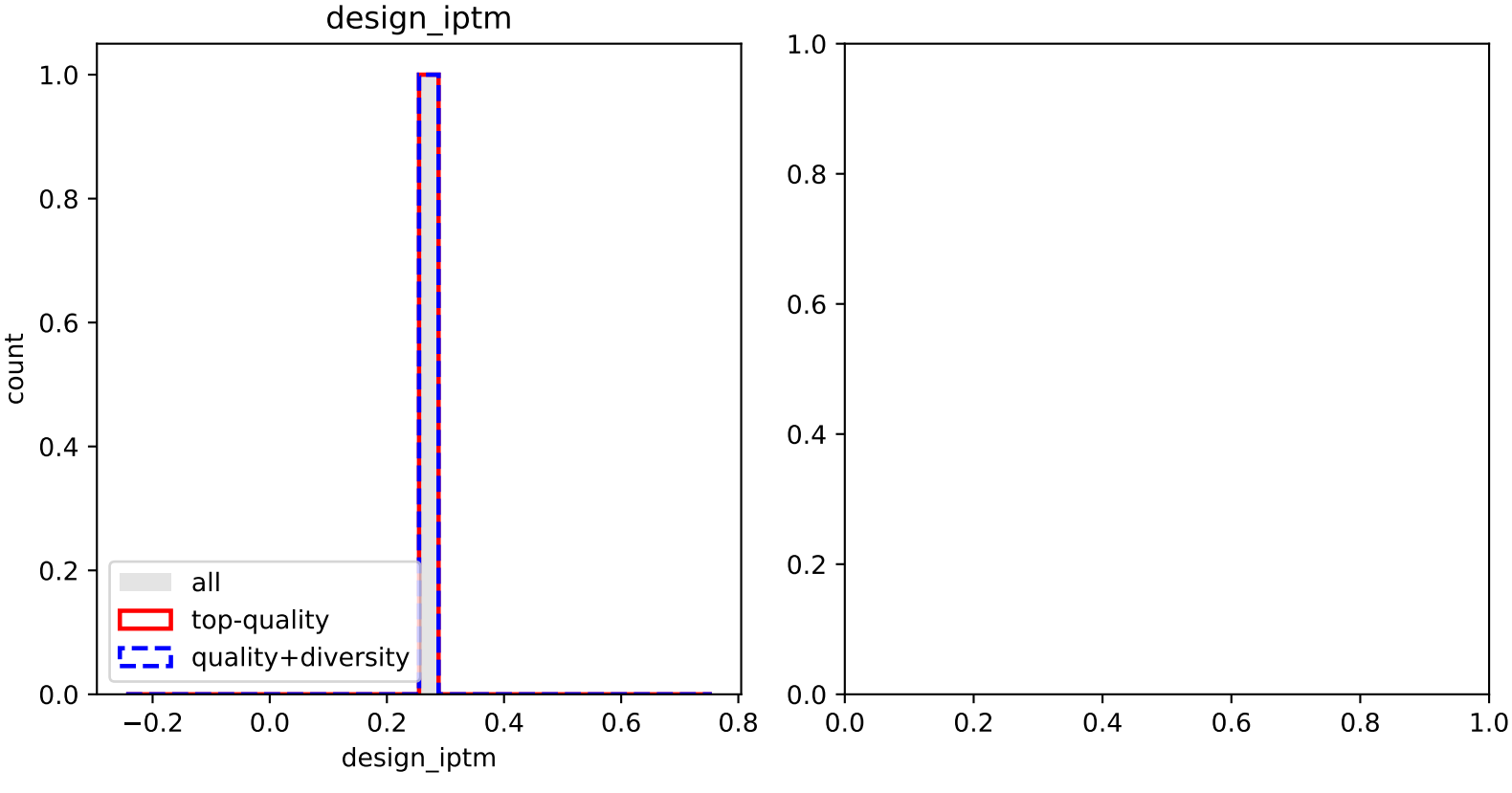

num_design
all
top-quality
quality+diversity
count
num_design
1.0
0.8
0.6
0.4
0.2
0.0
154.6
154.8
155.0
155.2
155.4
0.0
0.2
0.4
0.6
0.8
1.0

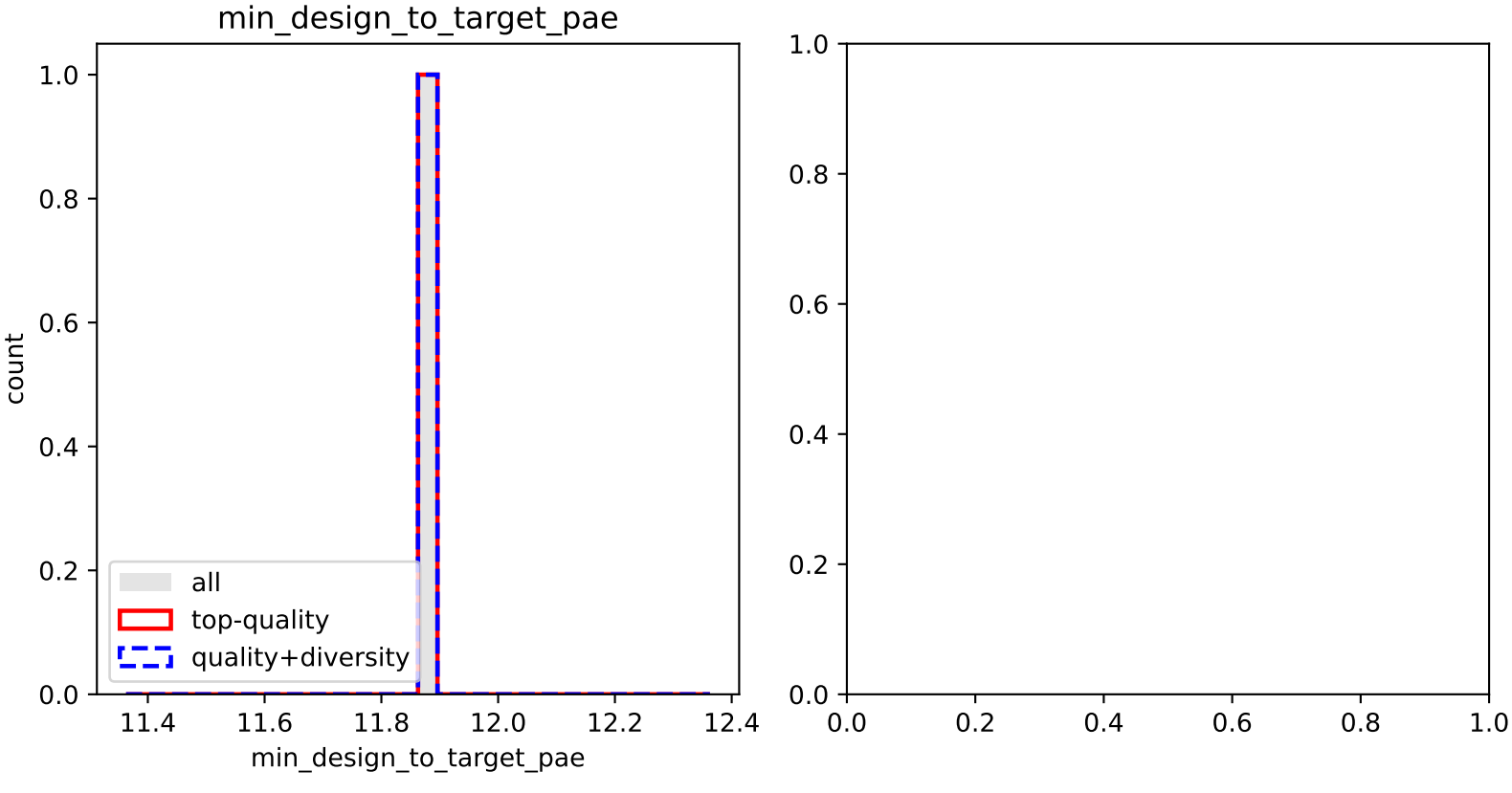

filter_rmsd
all
top-quality
quality+diversity
count
filter_rmsd
1.0
0.8
0.6
0.4
0.2
0.0
16.2
16.4
16.6
16.8
17.0
17.2
0.0
0.2
0.4
0.6
0.8
1.0

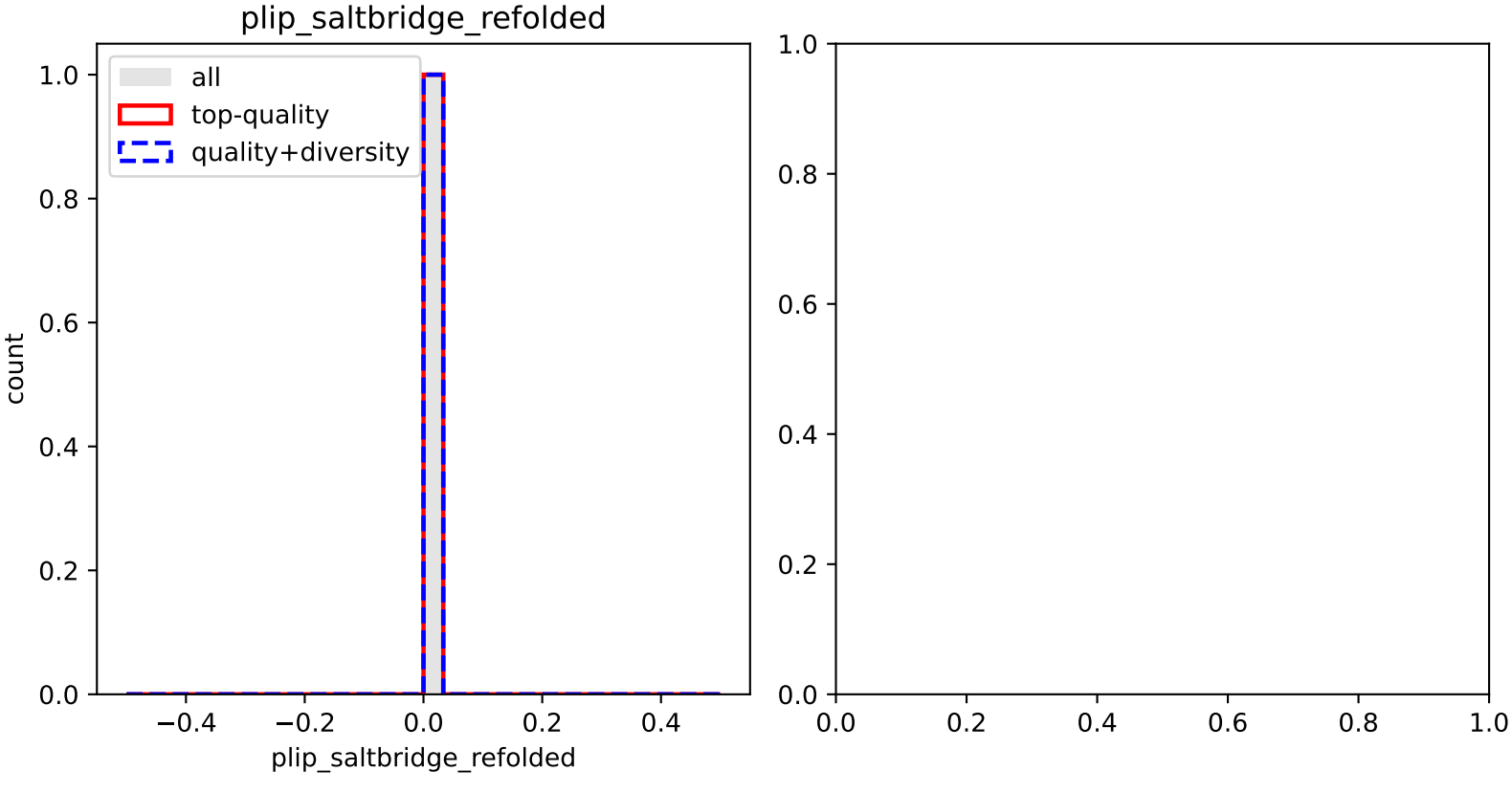

affinity_probability_binary1
all
top-quality
quality+diversity
count
affinity_probability_binary1
0.0
0.2
0.4
0.6
0.8
1.0

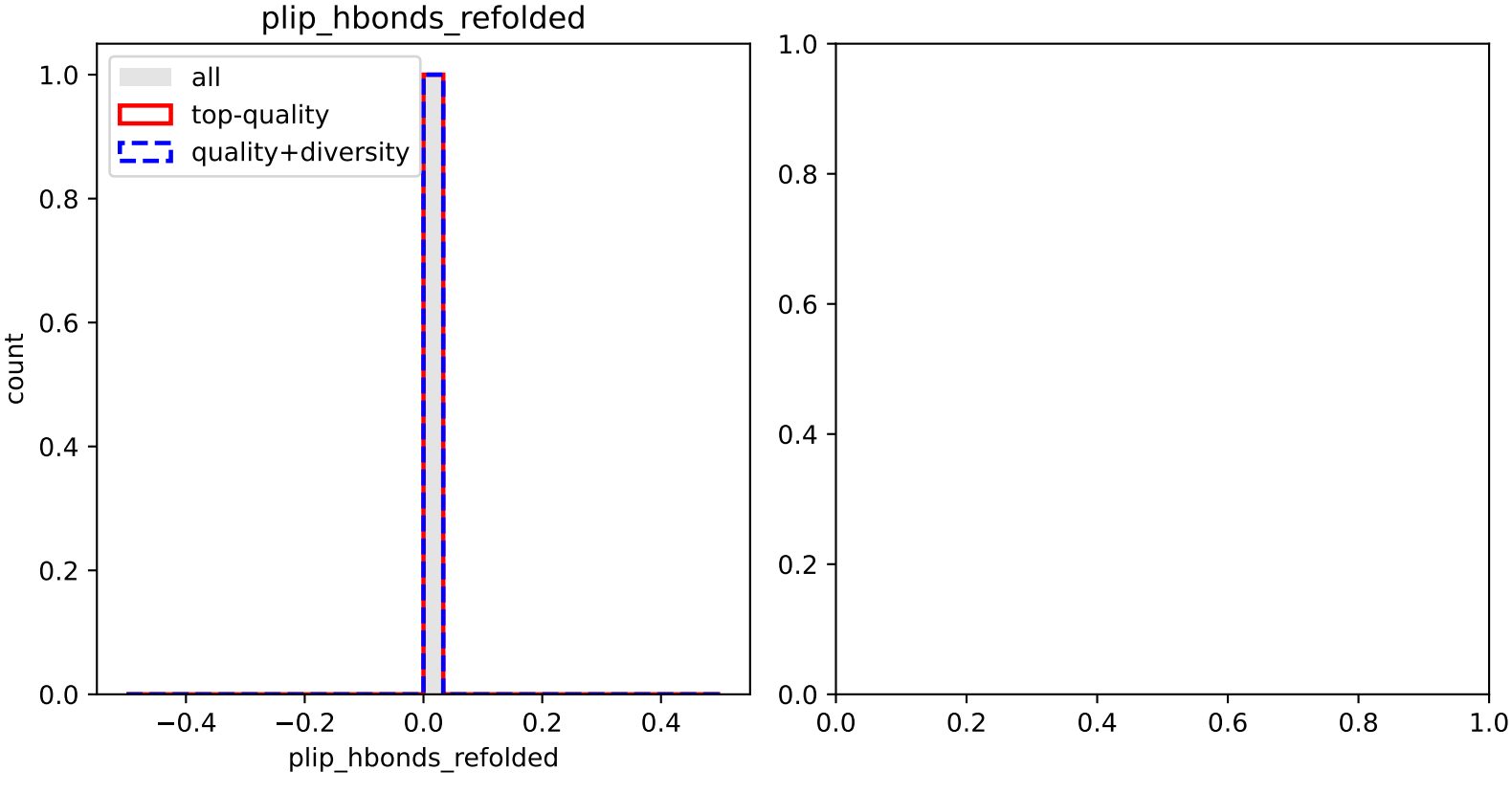

designfolding-filter_rmsd
all
top-quality
quality+diversity
count
designfolding-filter_rmsd
1.0
0.8
0.6
0.4
0.2
0.0
14.0
14.2
14.4
14.6
14.8
15.0
0.0
0.2
0.4
0.6
0.8
1.0

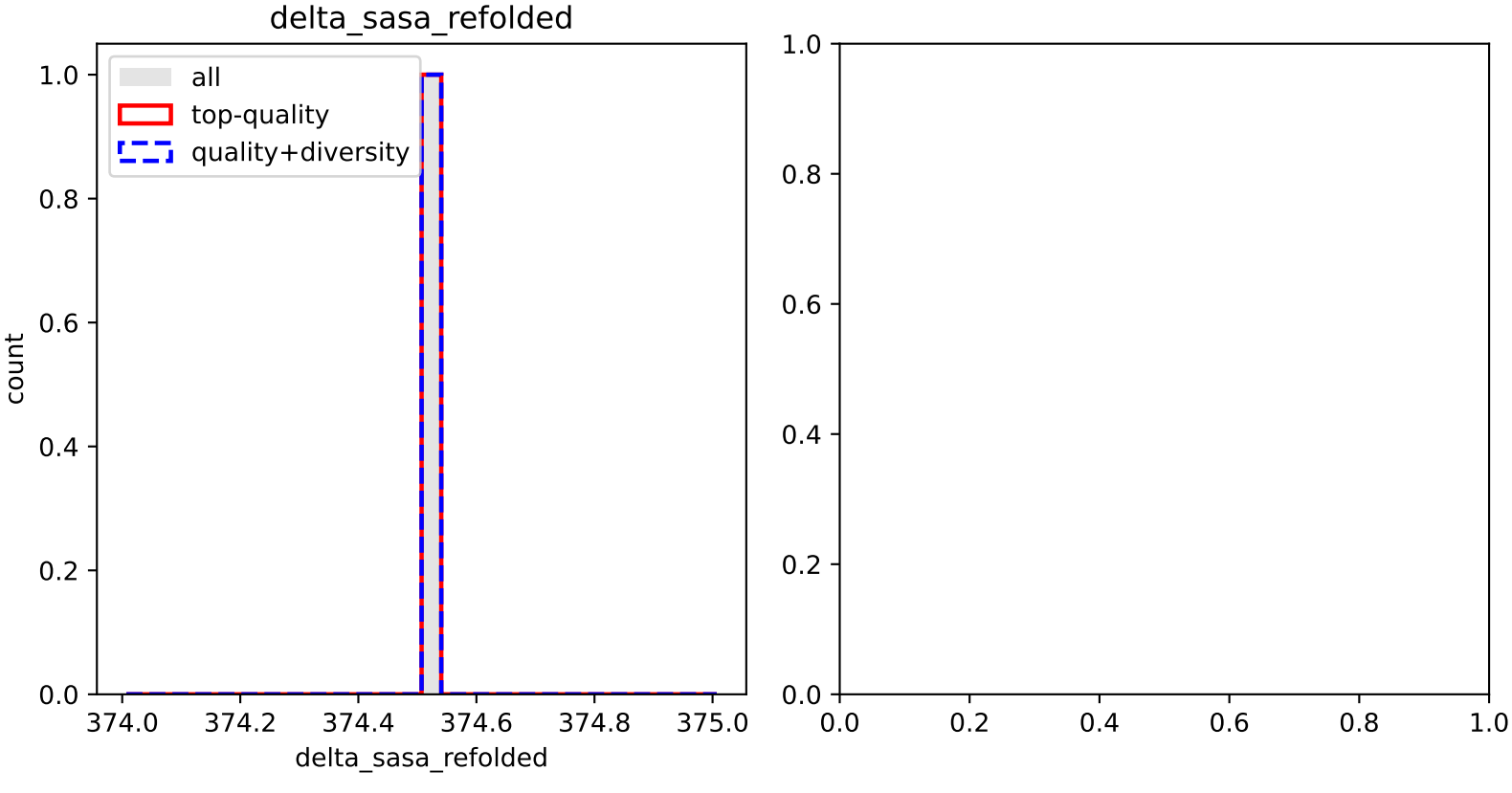

design_to_target_iptm
count
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0.0
0.2
0.4
0.6
0.8
all
top-quality
quality+diversity
design_to_target_iptm
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0

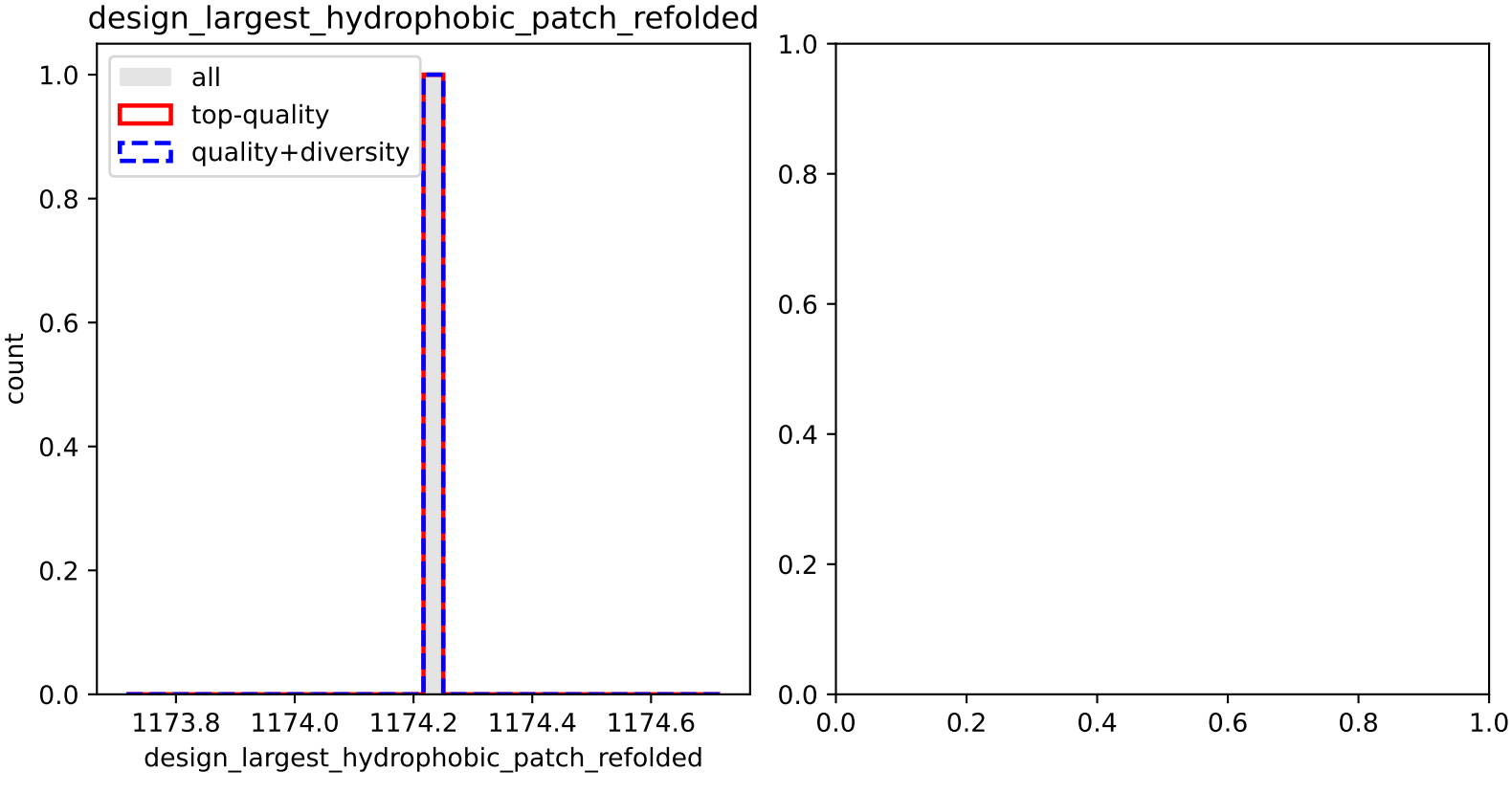

design_iptm
count
design_iptm
all
top-quality
quality+diversity
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0.0
0.2
0.4
0.6
0.8

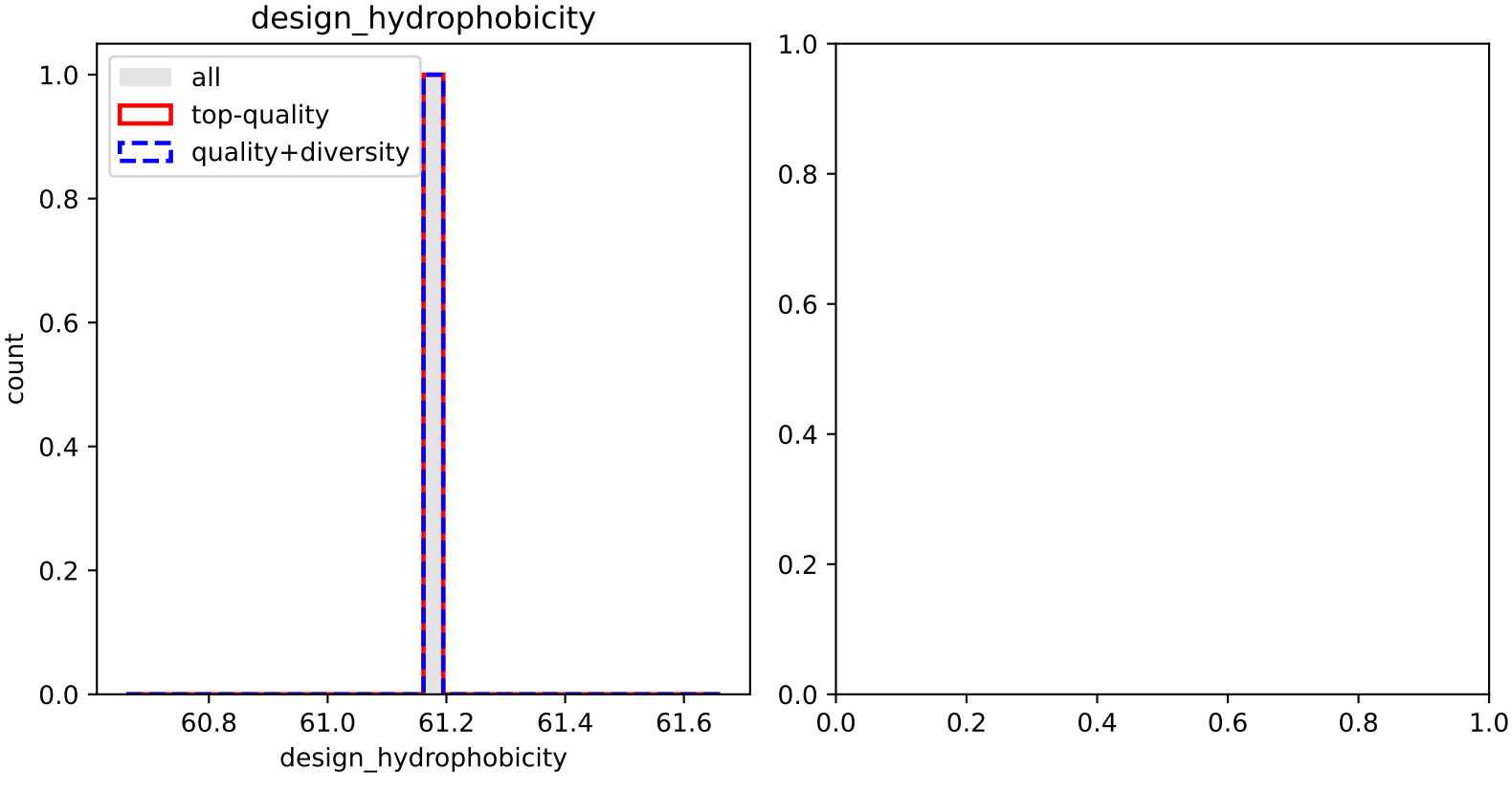

min_design_to_target_pae
count
1.0
0.8
0.6
0.4
0.2
0.0
11.4
11.6
11.8
12.0
12.2
12.4
min_design_to_target_pae
all
top-quality
quality+diversity
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


plip_saltbridge_refolded
all
top-quality
quality+diversity
count
plip_saltbridge_refolded
1.0
0.8
0.6
0.4
0.2
0.0
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
1.0


plip_hbonds_refolded
all
top-quality
quality+diversity
count
plip_hbonds_refolded
1.0
0.8
0.6
0.4
0.2
0.0
−0.4
−0.2
0.0
0.2
0.4
0.0
0.2
0.4
0.6
0.8
1.0


delta_sasa_refolded
all
top-quality
quality+diversity
count
delta_sasa_refolded
0.0
0.2
0.4
0.6
0.8
1.0
374.0
374.2
374.4
374.6
374.8
375.0


design_largest_hydrophobic_patch_refolded
all
top-quality
quality+diversity
count
1.0
0.8
0.6
0.4
0.2
0.0
1173.8
1174.0
1174.2
1174.4
1174.6
design_largest_hydrophobic_patch_refolded
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0


design_hydrophobicity
all
top-quality
quality+diversity
count
1.0
0.8
0.6
0.4
0.2
0.0
60.8
61.0
61.2
61.4
61.6
design_hydrophobicity
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0