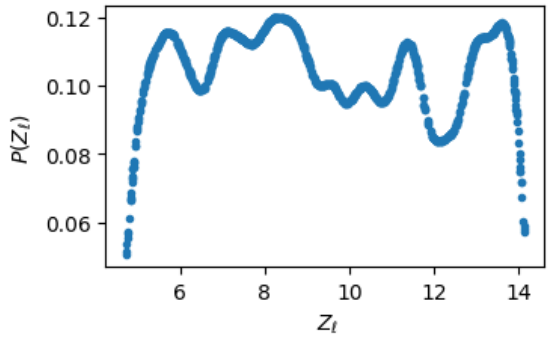
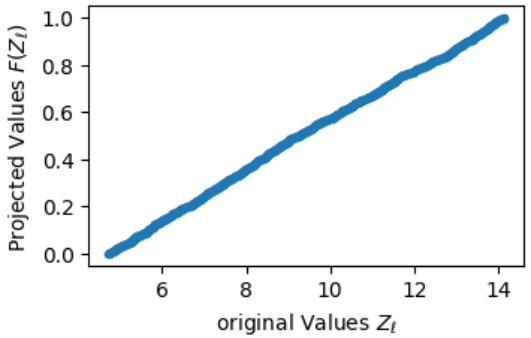
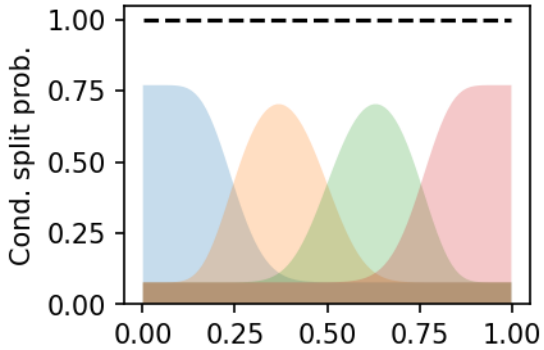
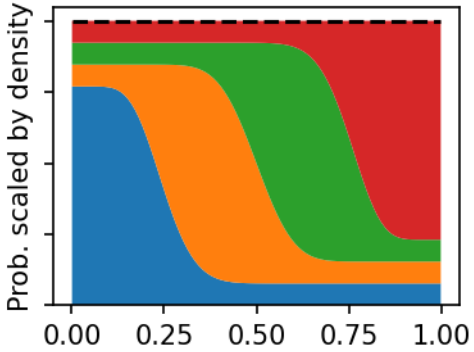
Compute $P(Z_t)$ and $F(Z_t)$

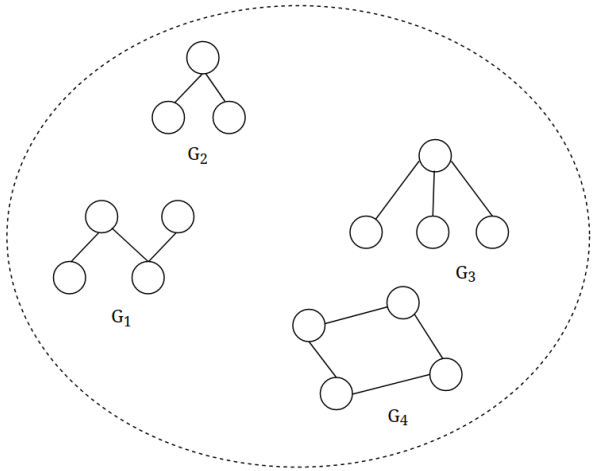






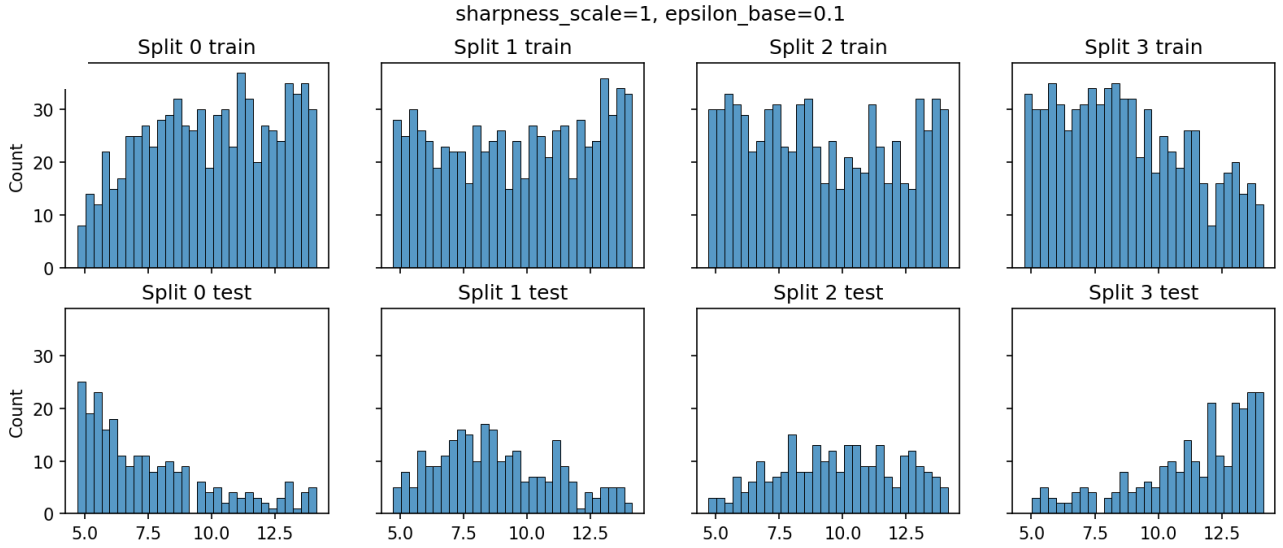
Compute conditionals

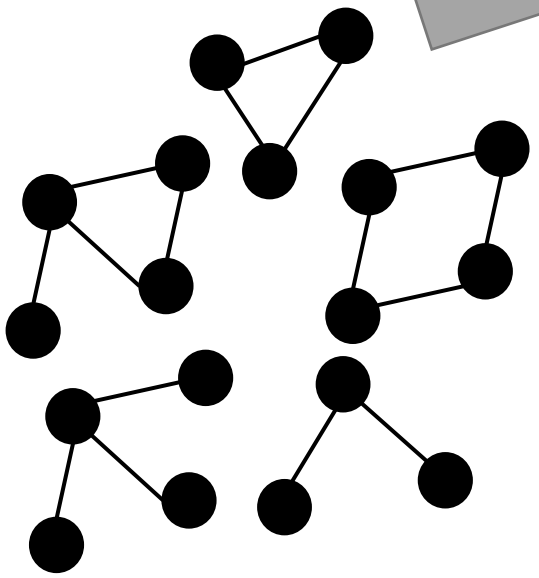




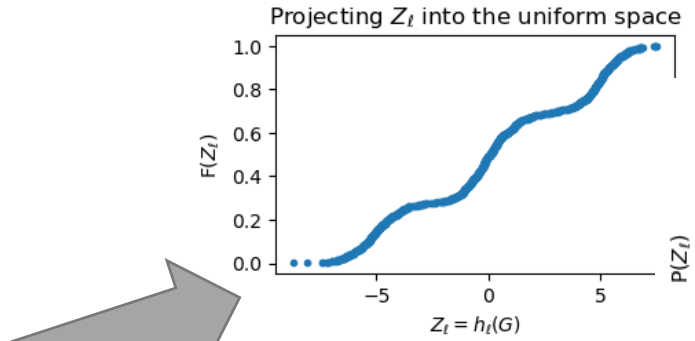
Dataset of Graphs

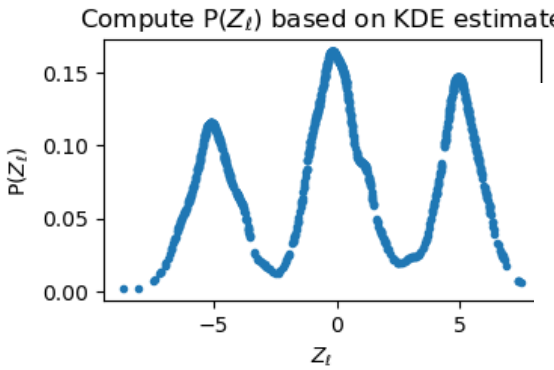
Split the data based on conditionals

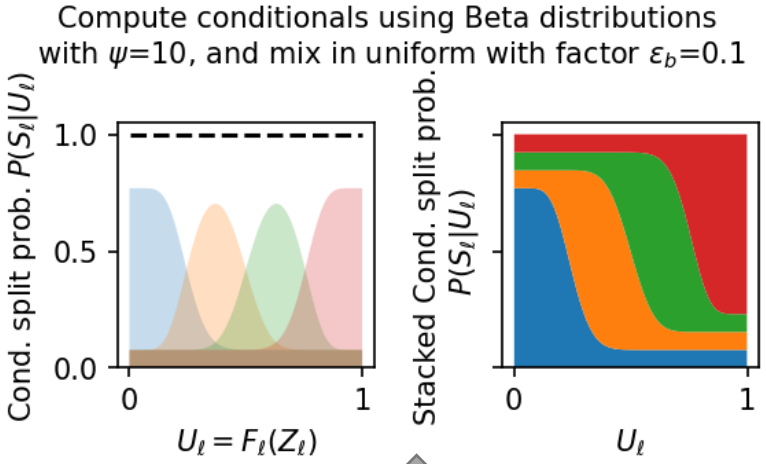




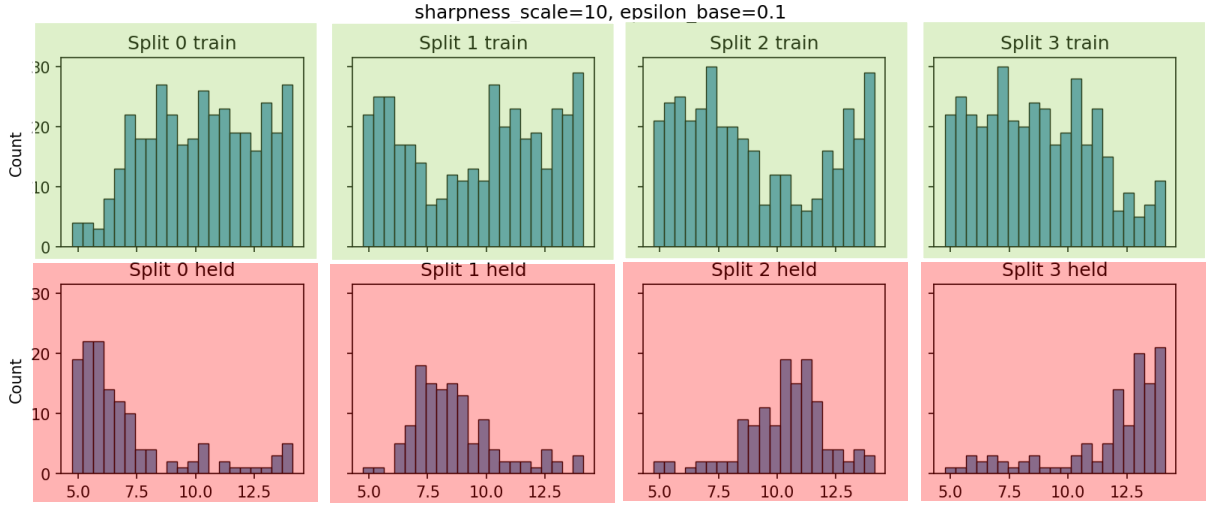
Dataset of Graphs

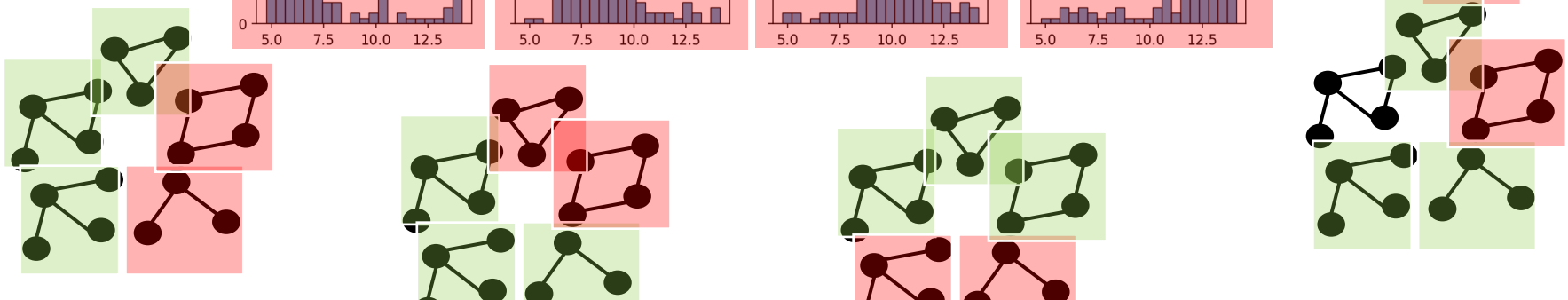


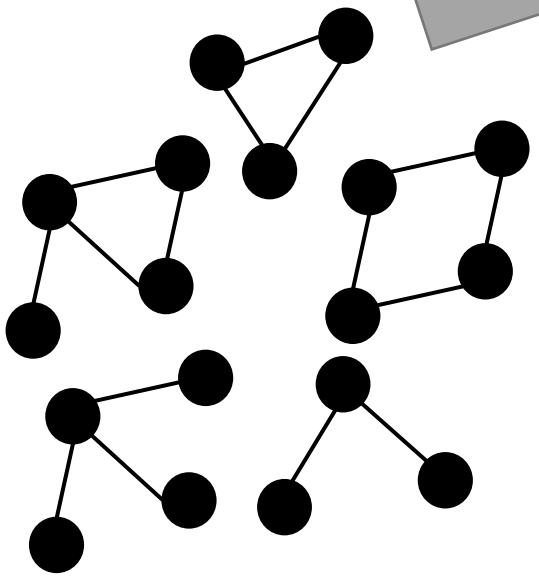




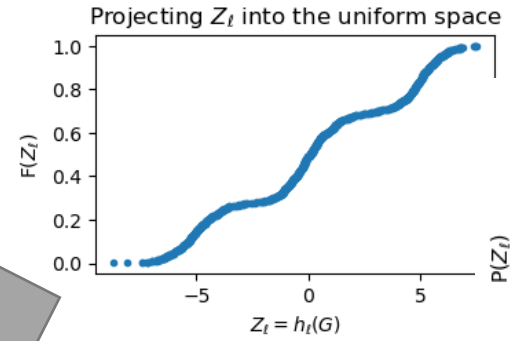
Split the data based on conditionals

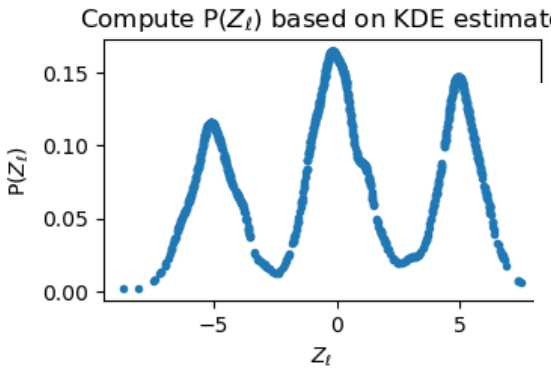


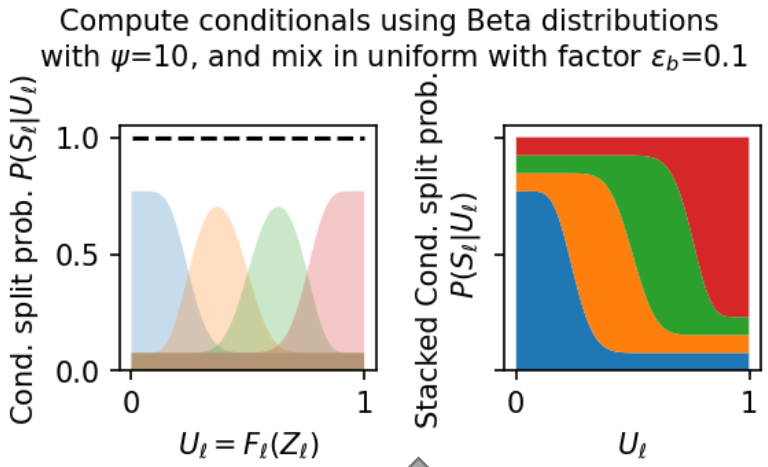




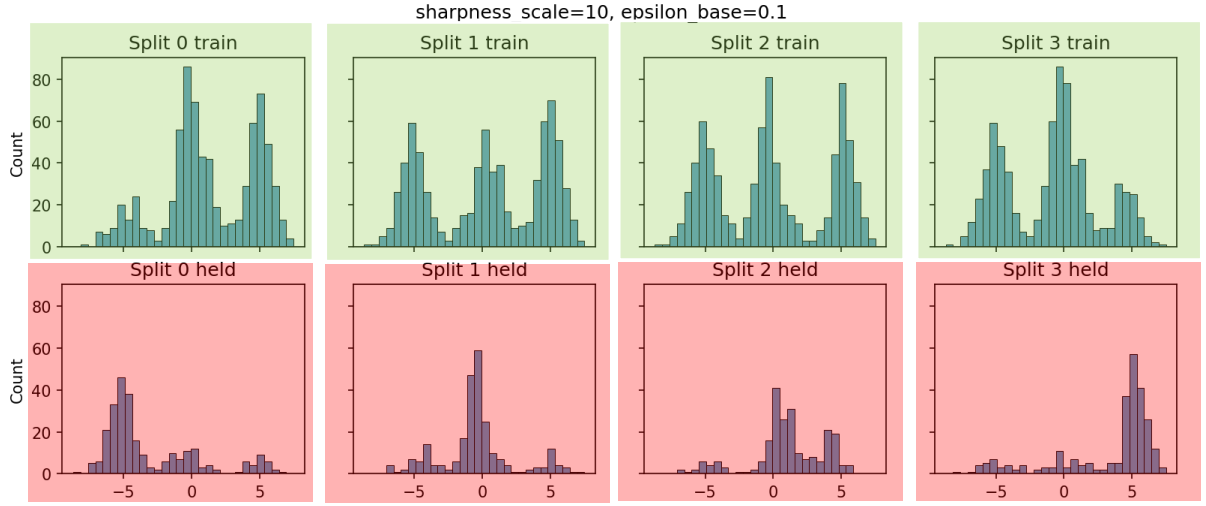
Dataset of Graphs

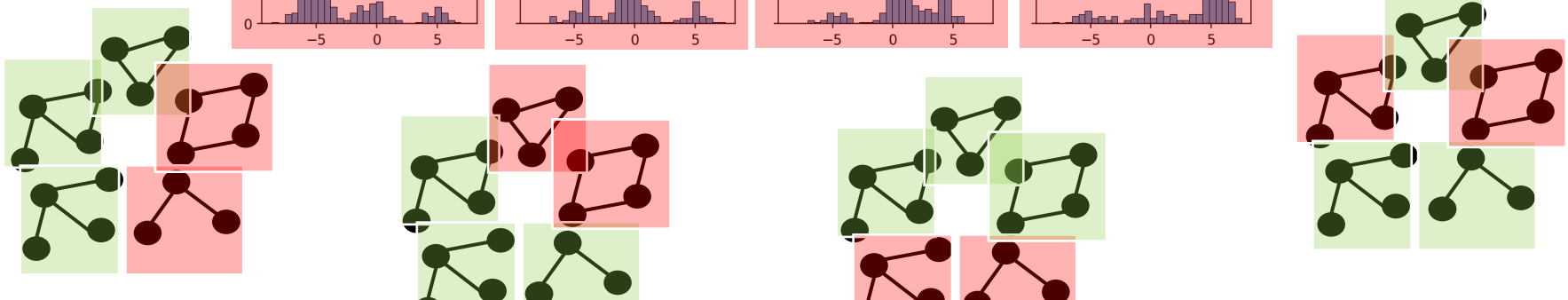




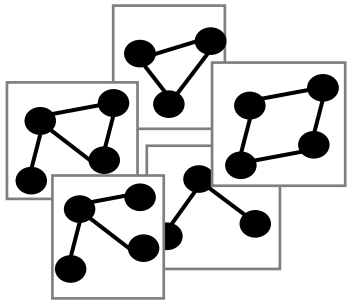


Split the data based on conditionals



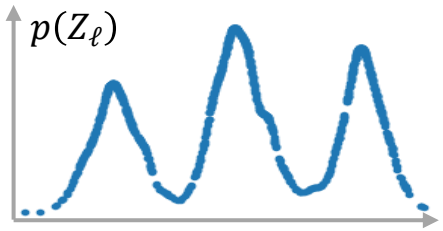


Dataset of Graphs



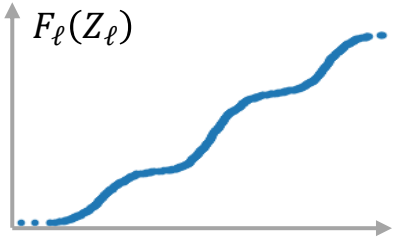
$\mathcal{G} = \{G_i\}_{i=1}^{|\mathcal{G}|}$

1. Compute Graph Properties



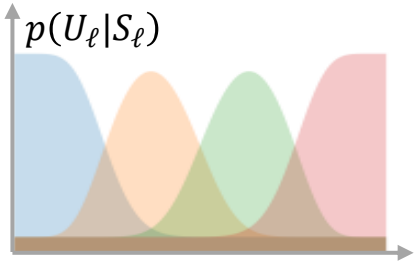
$Z_\ell = h_\ell(G)$

2. Project to unit interval via CDF



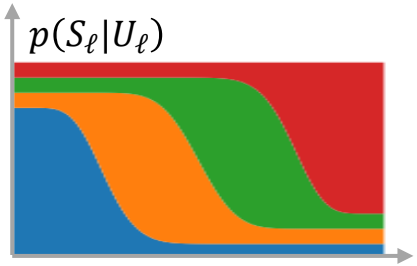
$Z_\ell = h_\ell(G)$

3. Define split distributions on [0,1]

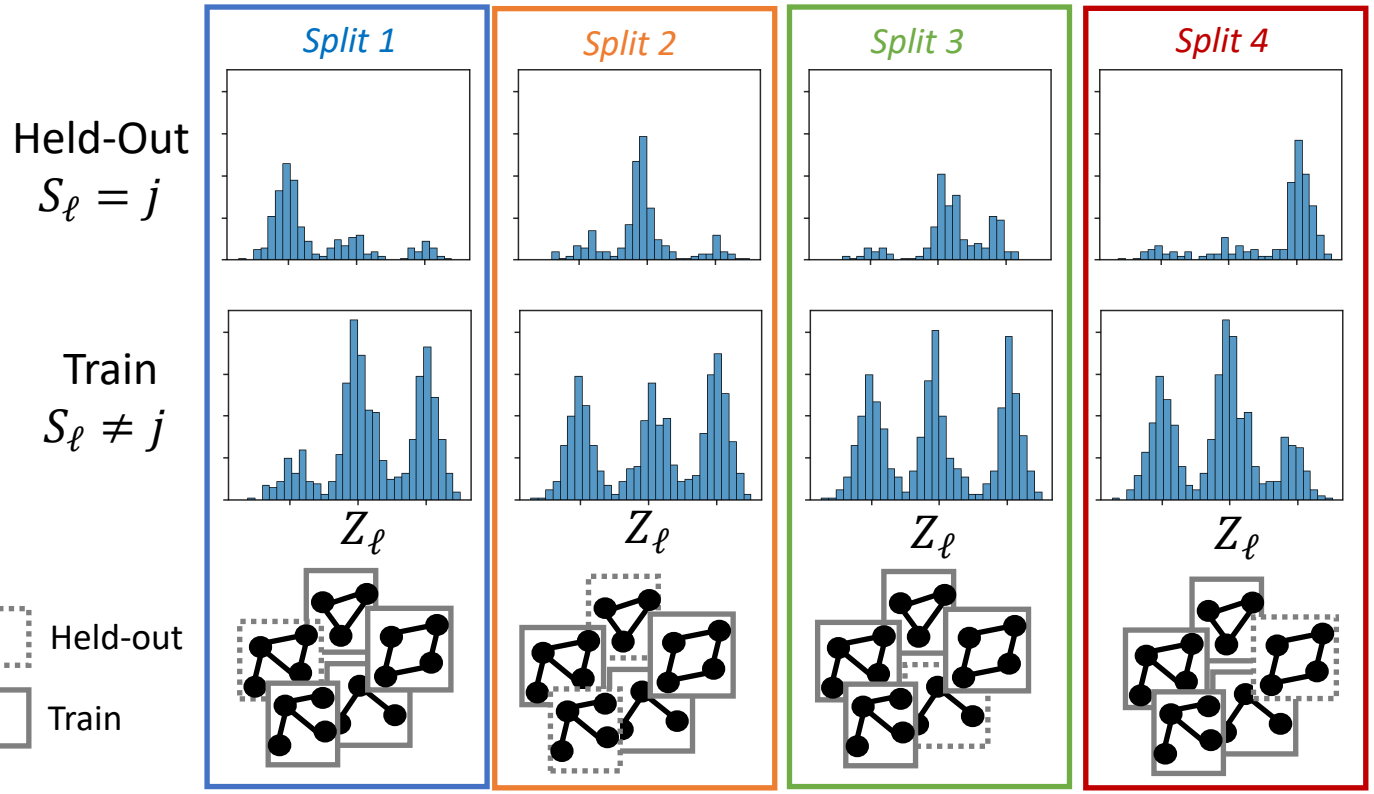


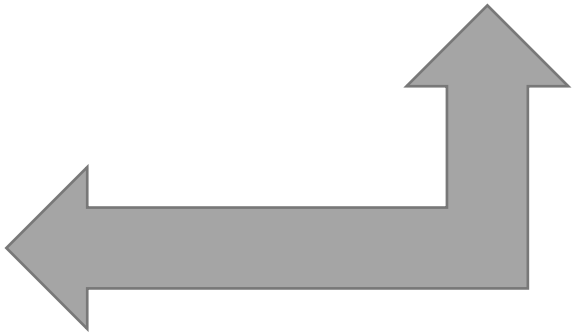
$U_\ell = F_\ell(Z_\ell)$

4. Compute split prob *given* U_ℓ



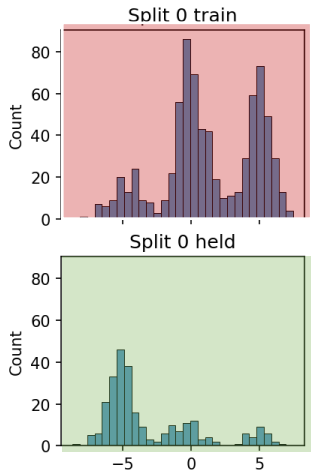
$U_\ell = F(Z_\ell)$

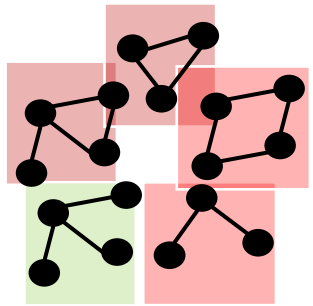




5. Sample $S_\ell \sim p(S_\ell|U_\ell)$ for each graph

Focusing on one of the splits





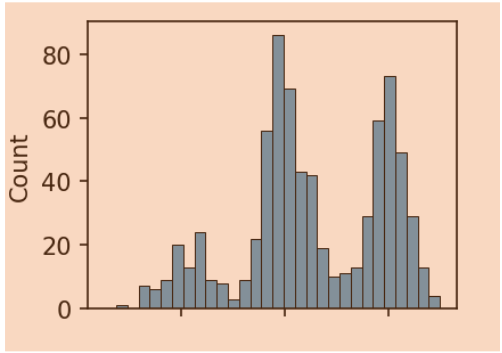
$\mathcal{G}_{\text{train}}^{(\ell,j)}$

Use to train model

Model

Model generates

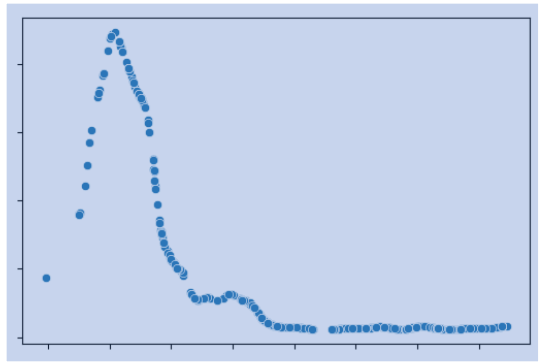
$\mathcal{G}_{\text{gen}}^{(\ell,j)}$



Calculated weights

$\mathcal{G}_{\text{held}}^{(\ell,j)}$

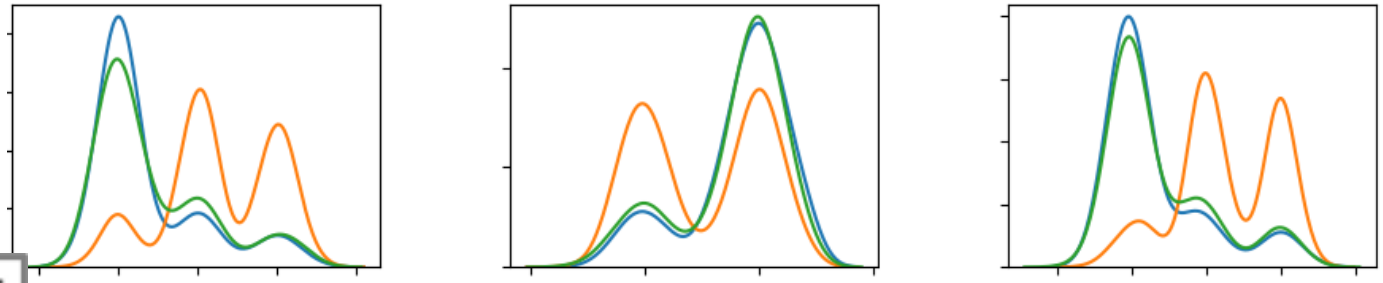
Use to Calculate Weights to reweight $\mathcal{G}_{\text{gen}}$



Property 1

Property 2

Property 3



$\mathcal{G}_{gen}$ Reweighted

$\mathcal{G}_{gen}$ Unweighted

$\mathcal{G}_{held}$

$\mathcal{G}_{\text{gen}}^{(\ell,j)}$ Reweighted

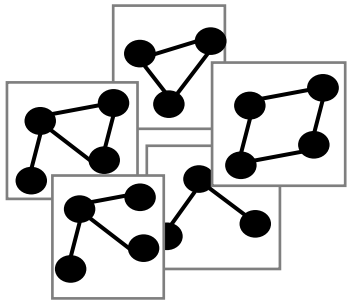
$\mathcal{G}_{\text{gen}}^{(\ell,j)}$ Unweighted

$\mathcal{G}_{\text{held}}^{(\ell,j)}$

Use weights to reweight samples of $\mathcal{G}_{\text{gen}}^{(\ell,j)}$ for all the dataset properties

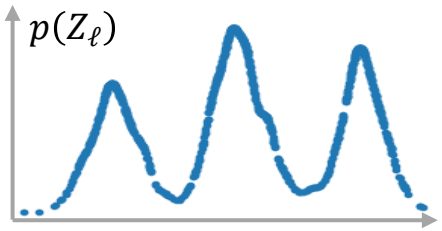
Finally calculate KS statistic between reweighted $\mathcal{G}_{\text{gen}}^{(\ell,j)}$ and $\mathcal{G}_{\text{held}}^{(\ell,j)}$

Dataset of Graphs



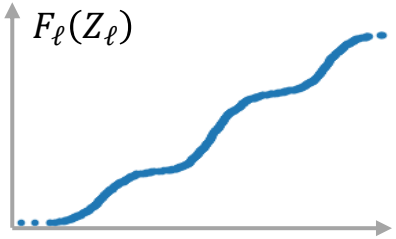
$\mathcal{G} = \{G_i\}_{i=1}^{|\mathcal{G}|}$

1. Compute Graph Properties



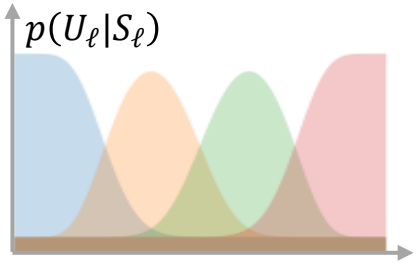
$Z_\ell = h_\ell(G)$

2. Project to unit interval via CDF



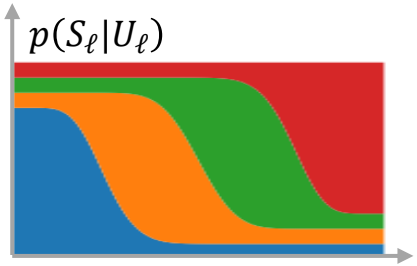
$Z_\ell = h_\ell(G)$

3. Define split distributions on [0,1]

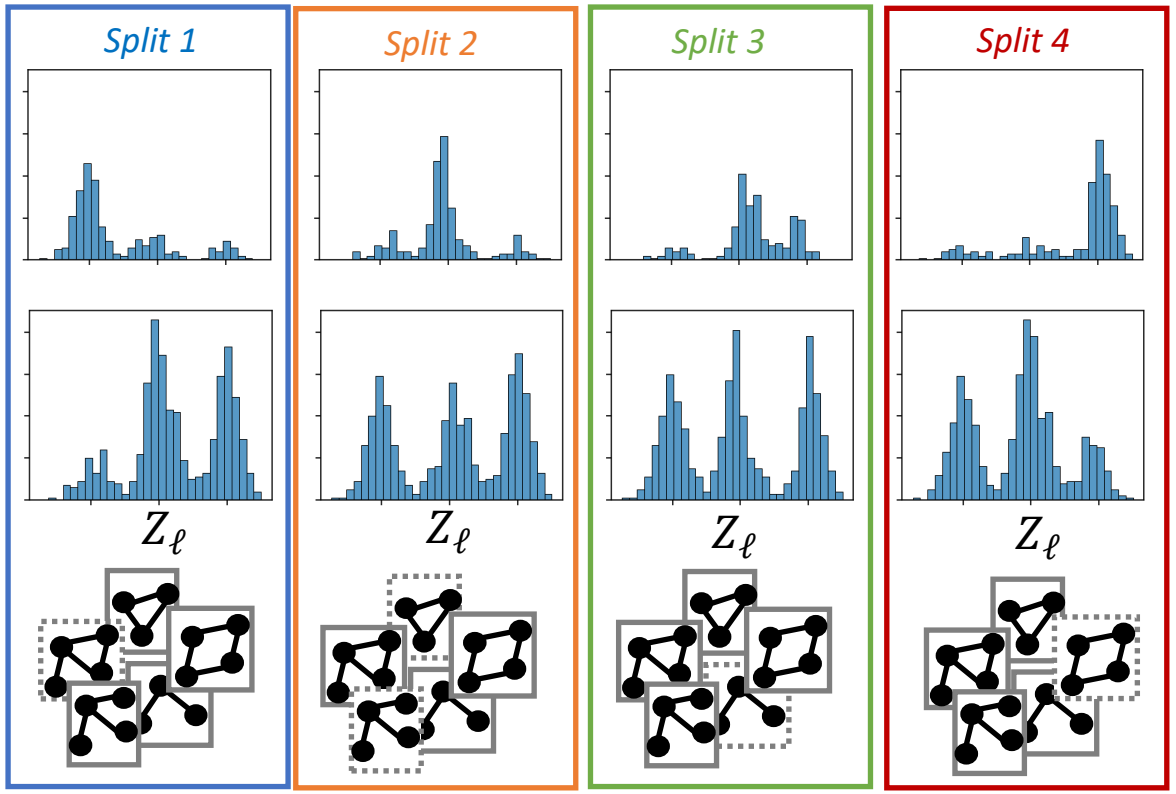


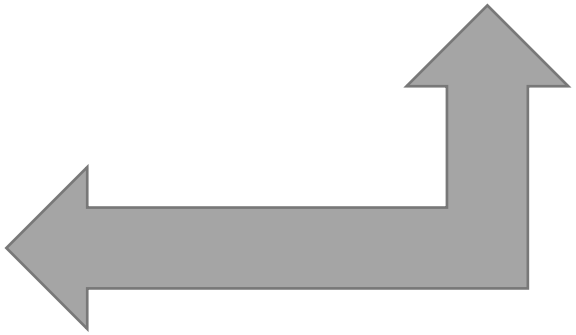
$U_\ell = F_\ell(Z_\ell)$

4. Compute split prob *given* U_ℓ



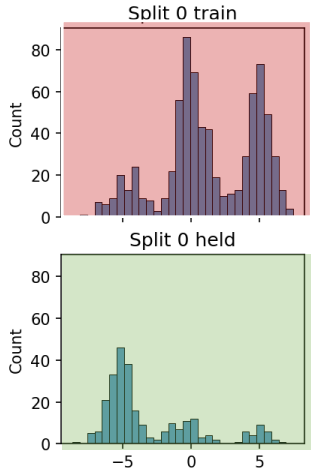
$U_\ell = F(Z_\ell)$

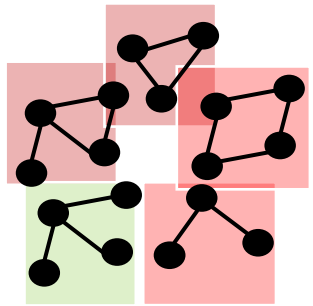




5. Sample $S_{i,\ell} \sim p(S_{i,\ell}|U_{i,\ell})$ for each graph

Focusing on one of the splits





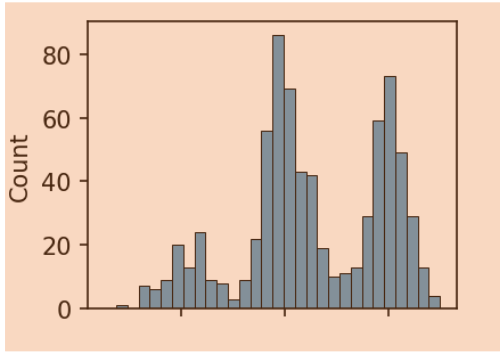
$\mathcal{G}_{\text{train}}^{(\ell,j)}$

Use to train model

Model

Model generates

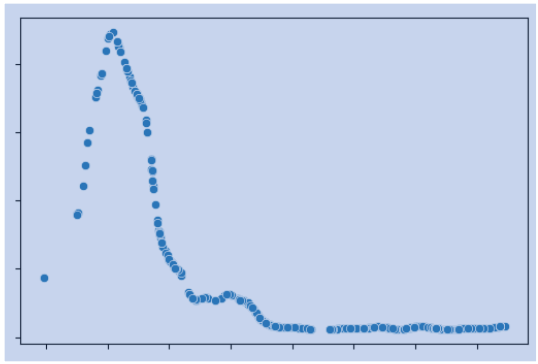
$\mathcal{G}_{\text{gen}}^{(\ell,j)}$

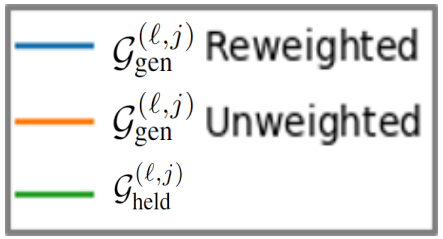


Calculated weights

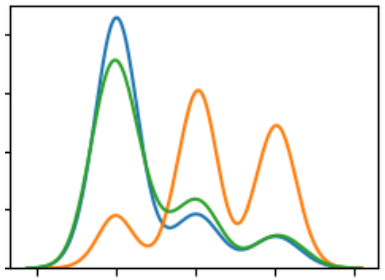
$\mathcal{G}_{\text{held}}^{(\ell,j)}$

Use to Calculate Weights to reweight $\mathcal{G}_{\text{gen}}$

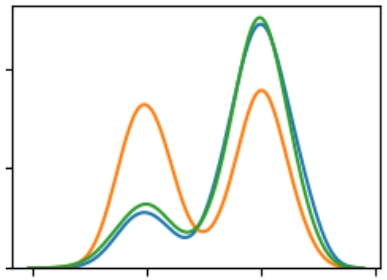




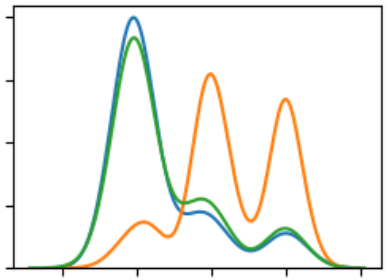
Property 1



Property 2

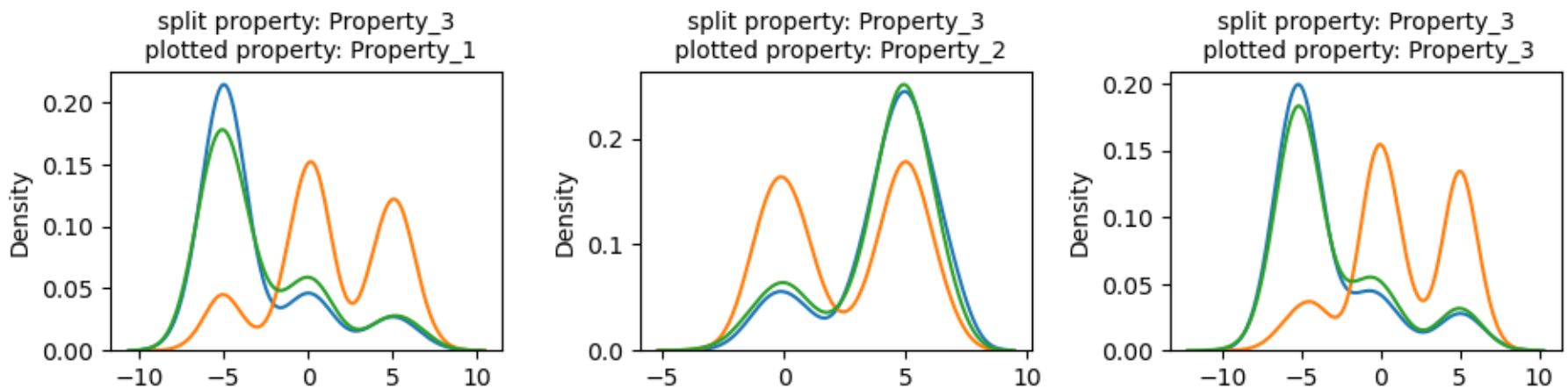


Property 3

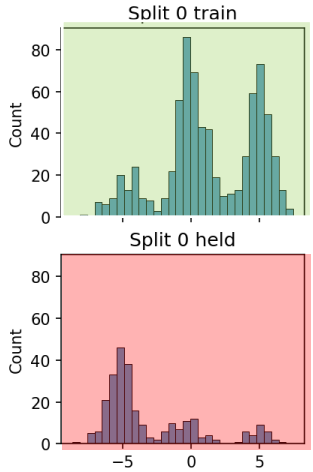


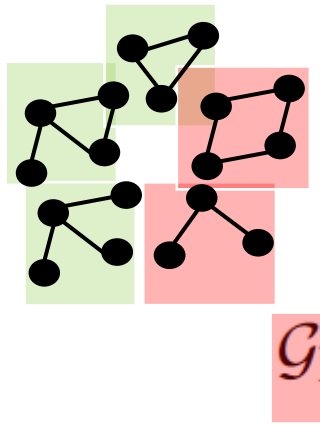
Use weights to reweight samples of $\mathcal{G}_{\text{gen}}^{(\ell,j)}$ for all the dataset properties

Finally calculate KS statistic between reweighted $\mathcal{G}_{\text{gen}}^{(\ell,j)}$ and $\mathcal{G}_{\text{held}}^{(\ell,j)}$


split property: Property_3
plotted property: Property_1
Density
0.20
0.15
0.10
0.05
0.00
−10
−5
0
5
10
split property: Property_3
plotted property: Property_2
Density
0.2
0.1
0.0
−5
0
5
10
split property: Property_3
plotted property: Property_3
Density
0.20
0.15
0.10
0.05
0.00
−10
−5
0
5
10

Focusing on one of the splits





$\mathcal{G}_{\text{train}}$

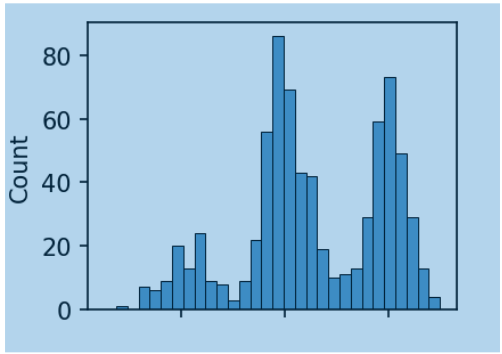
$\mathcal{G}_{\text{held}}$

Use to train model

Model

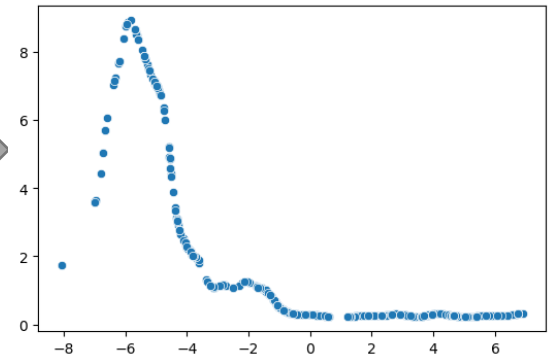
Model generates

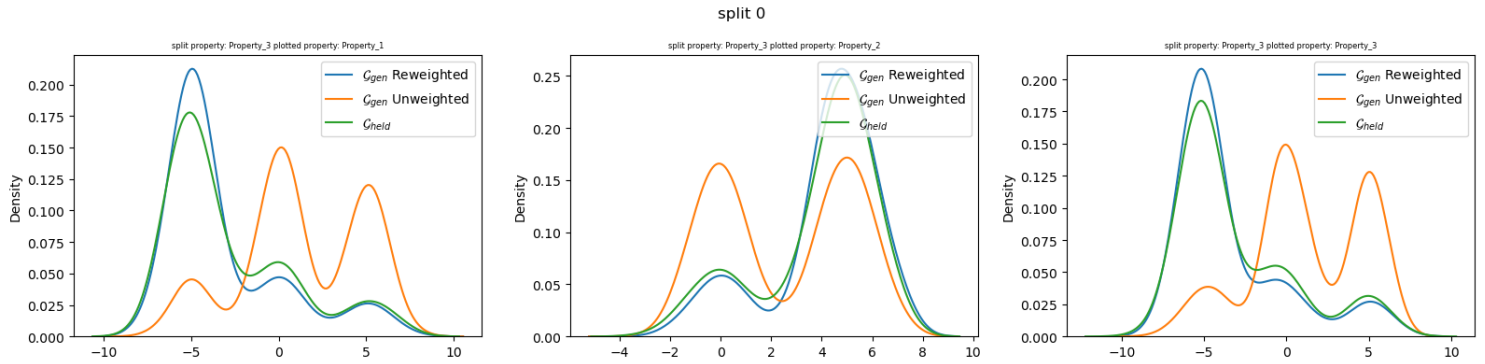
$\mathcal{G}_{\text{gen}}$



Use to Calculate Weights to reweight $\mathcal{G}_{\text{gen}}$

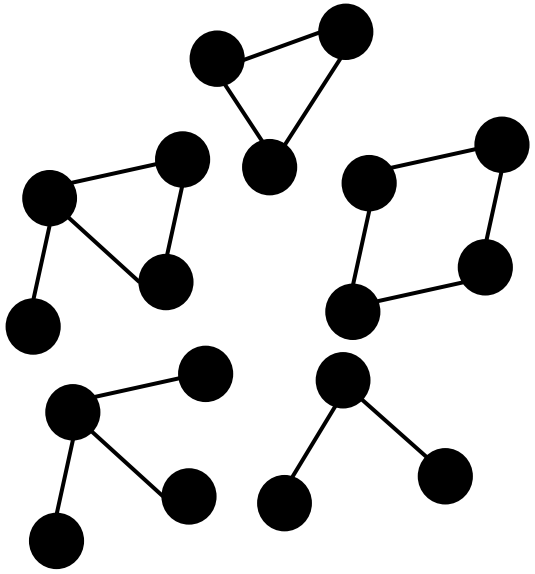
Calculated weights





Use weights to reweight samples of $\mathcal{G}_{\text{gen}}$ for all the dataset properties

Finally calculate KS statistic between reweighted $\mathcal{G}_{\text{gen}}$ and $\mathcal{G}_{\text{held}}$



Dataset of Graphs