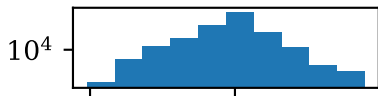


number of model predictions

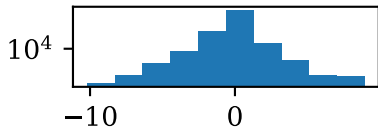
proportional  
change



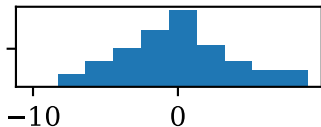
proportional  
change  
baseline



conditional  
weakening



conditional  
weakening  
baseline



entropy(clean) - entropy(ablated)