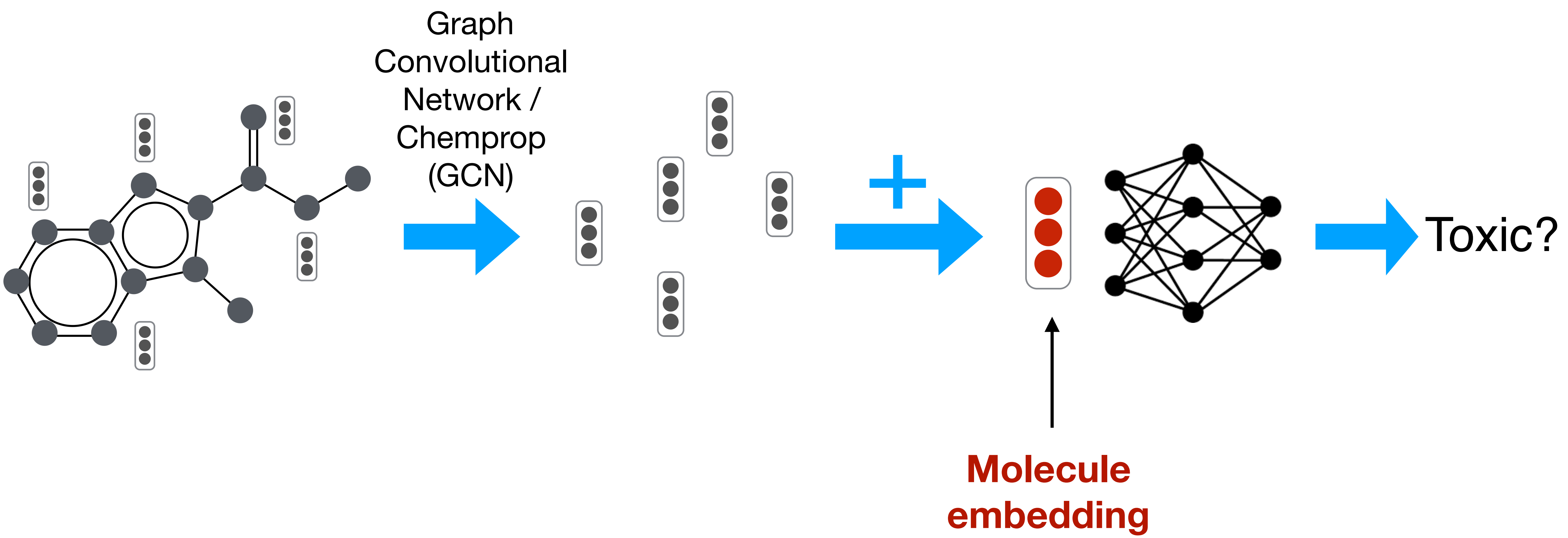
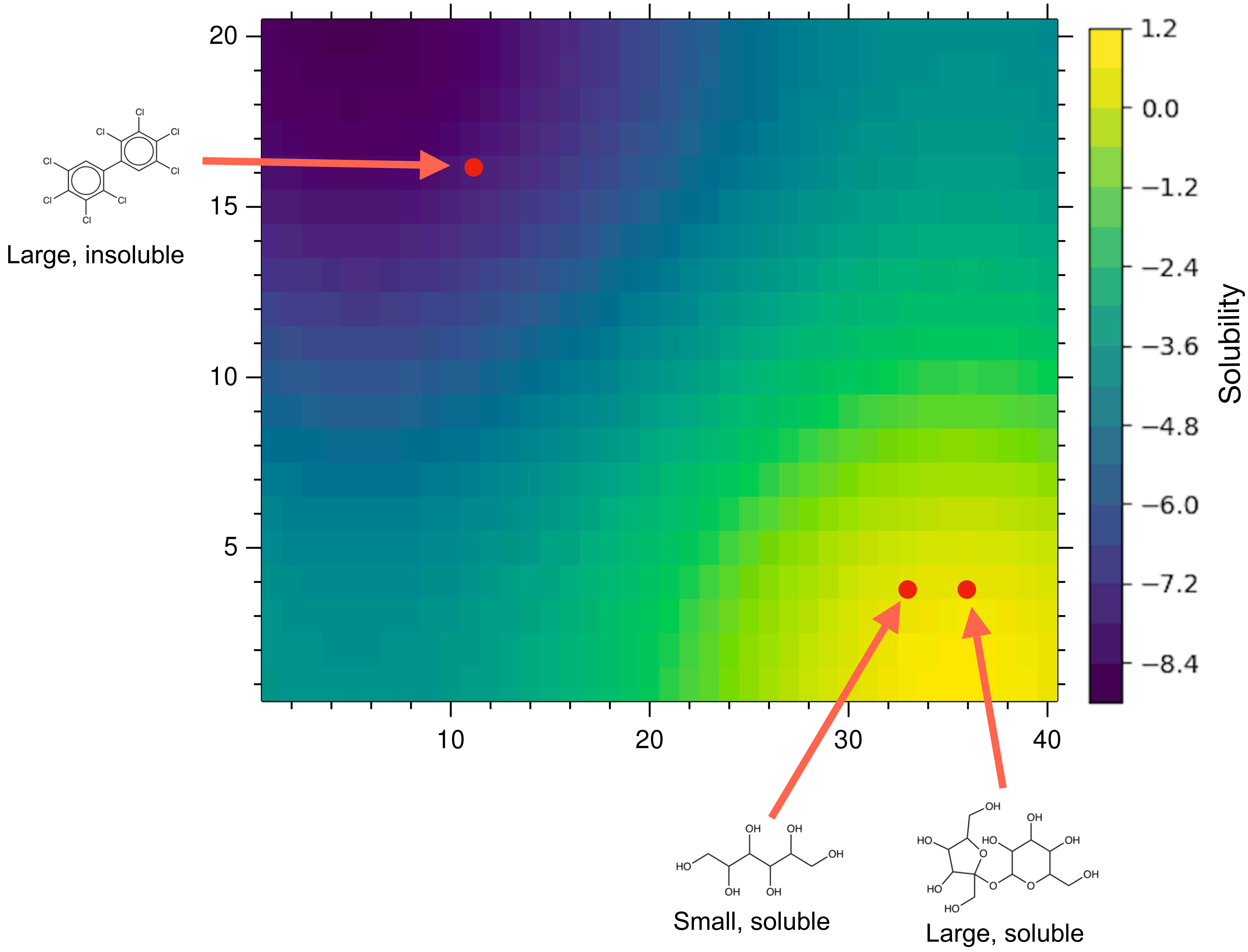
Optimal Transport Graph Neural Networks

Gary Bécigneul* **Octavian-Eugen Ganea*** **Benson Chen***
Regina Barzilay **Tommi Jaakkola**
Computer Science and Artificial Intelligence Lab, MIT

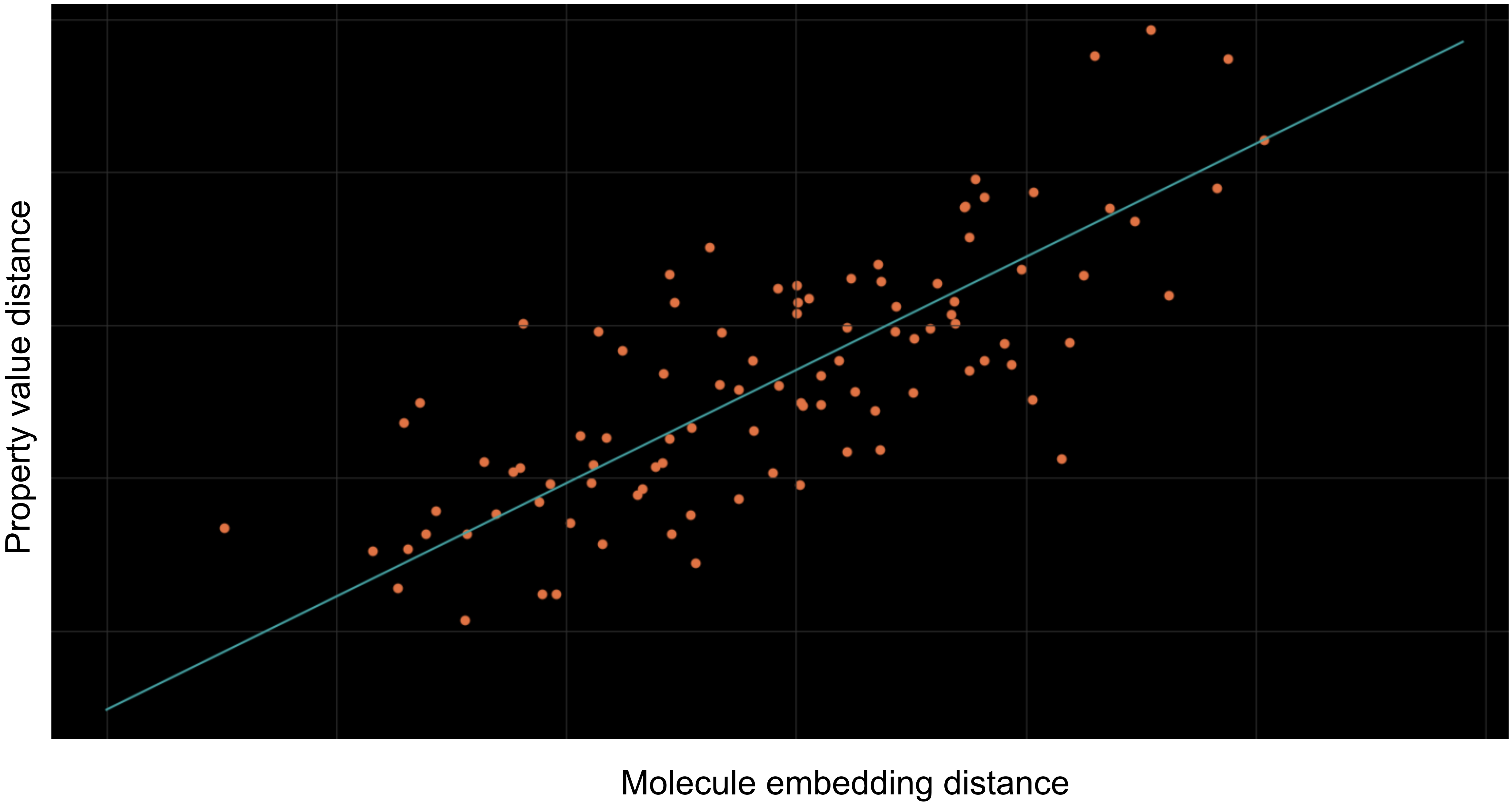
Graph Convolutions and Chemprop Model



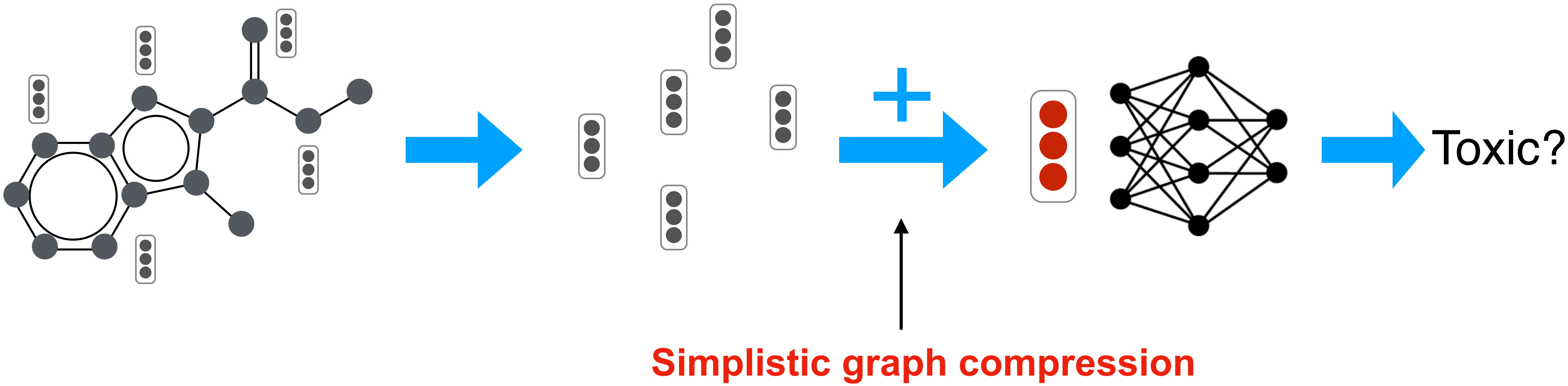
Property smoothness & robustness of representations



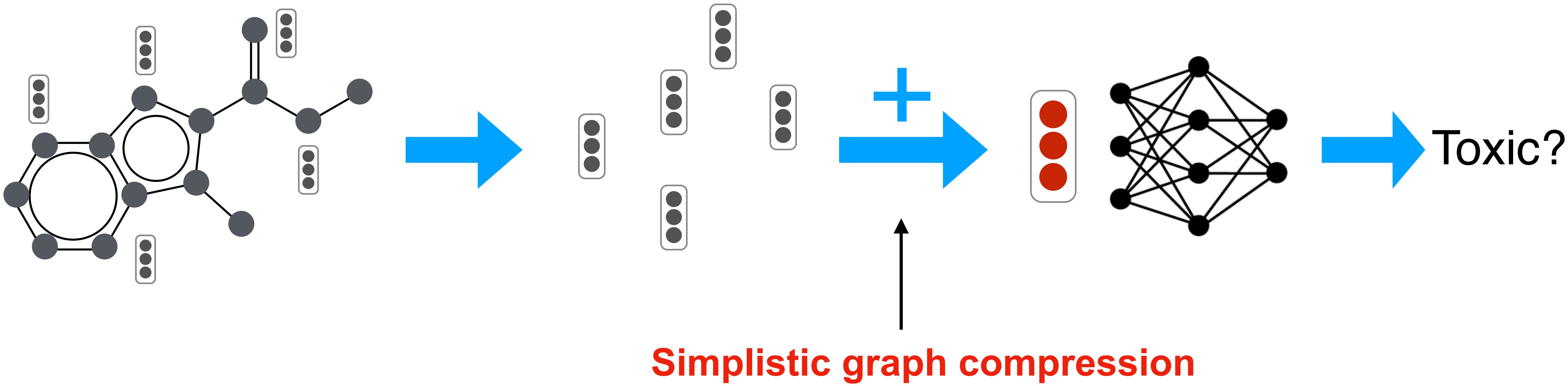
Property smoothness & robustness of representations



From atom embeddings to molecule embedding

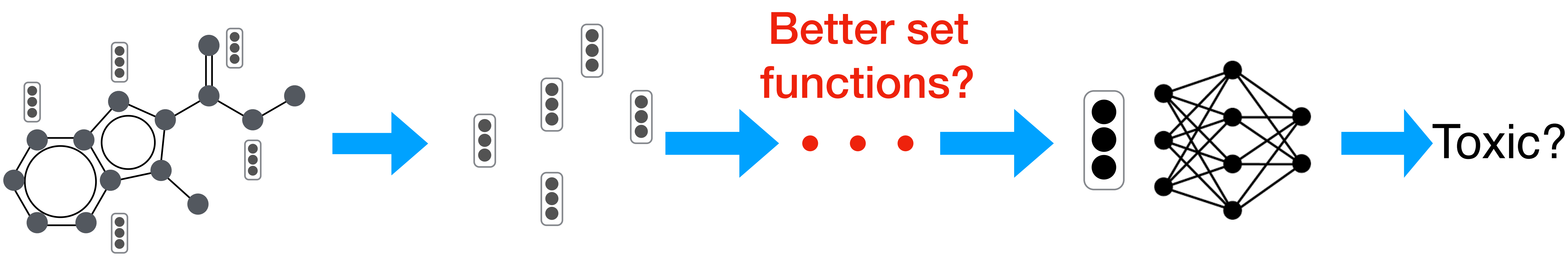


From atom embeddings to molecule embedding



Green, red and orange sets have the same sum and same average.

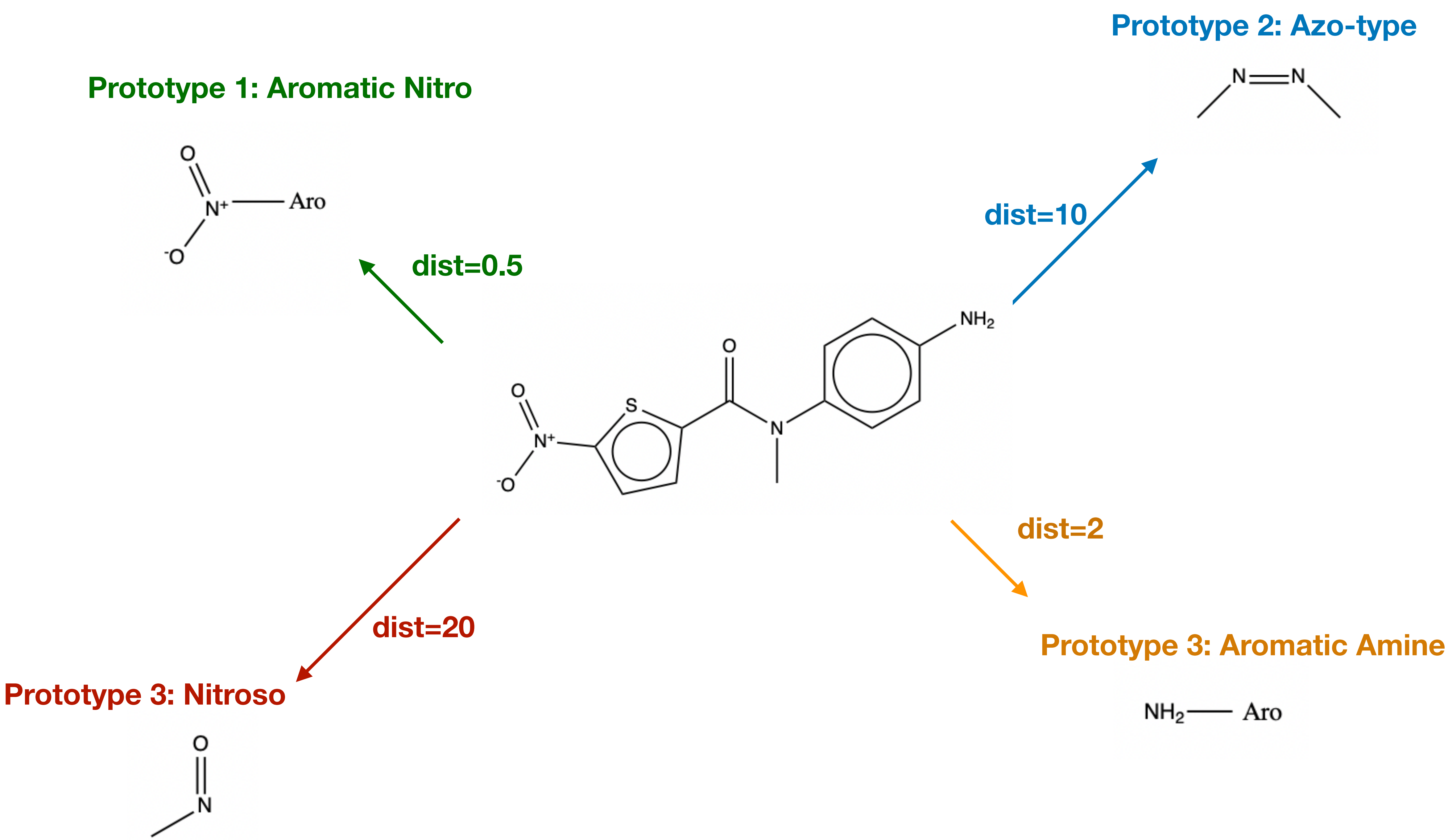
From atom embeddings to molecule embedding



Prototypes Inspired From Functional Groups

Our idea:

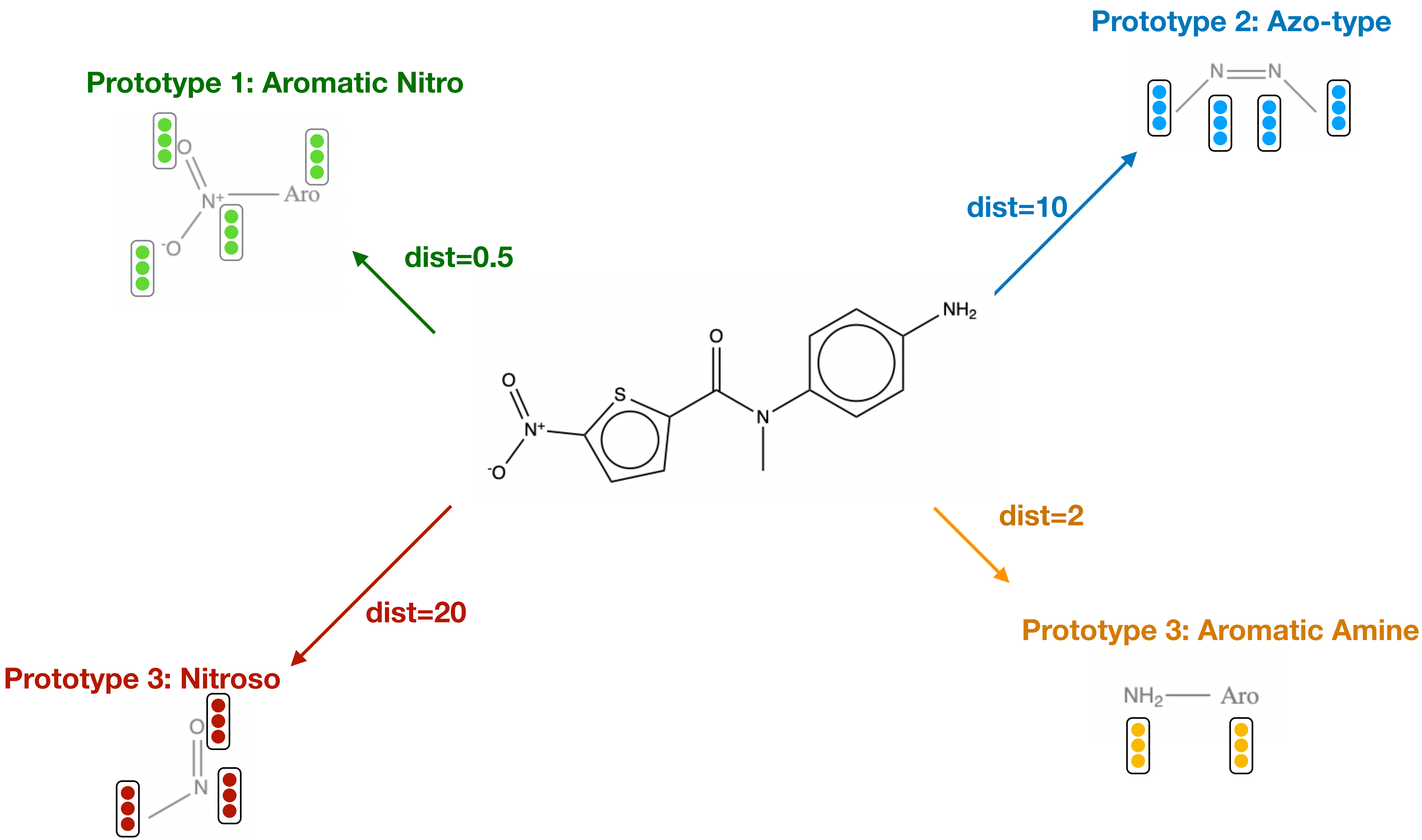
- *learn a dictionary of structural basis functions that serve to highlight key facets of compounds*
- *express molecules by relating them to abstract molecular prototypes*
- *prototypes highlight property values associated with different structural features (e.g. solubility)*



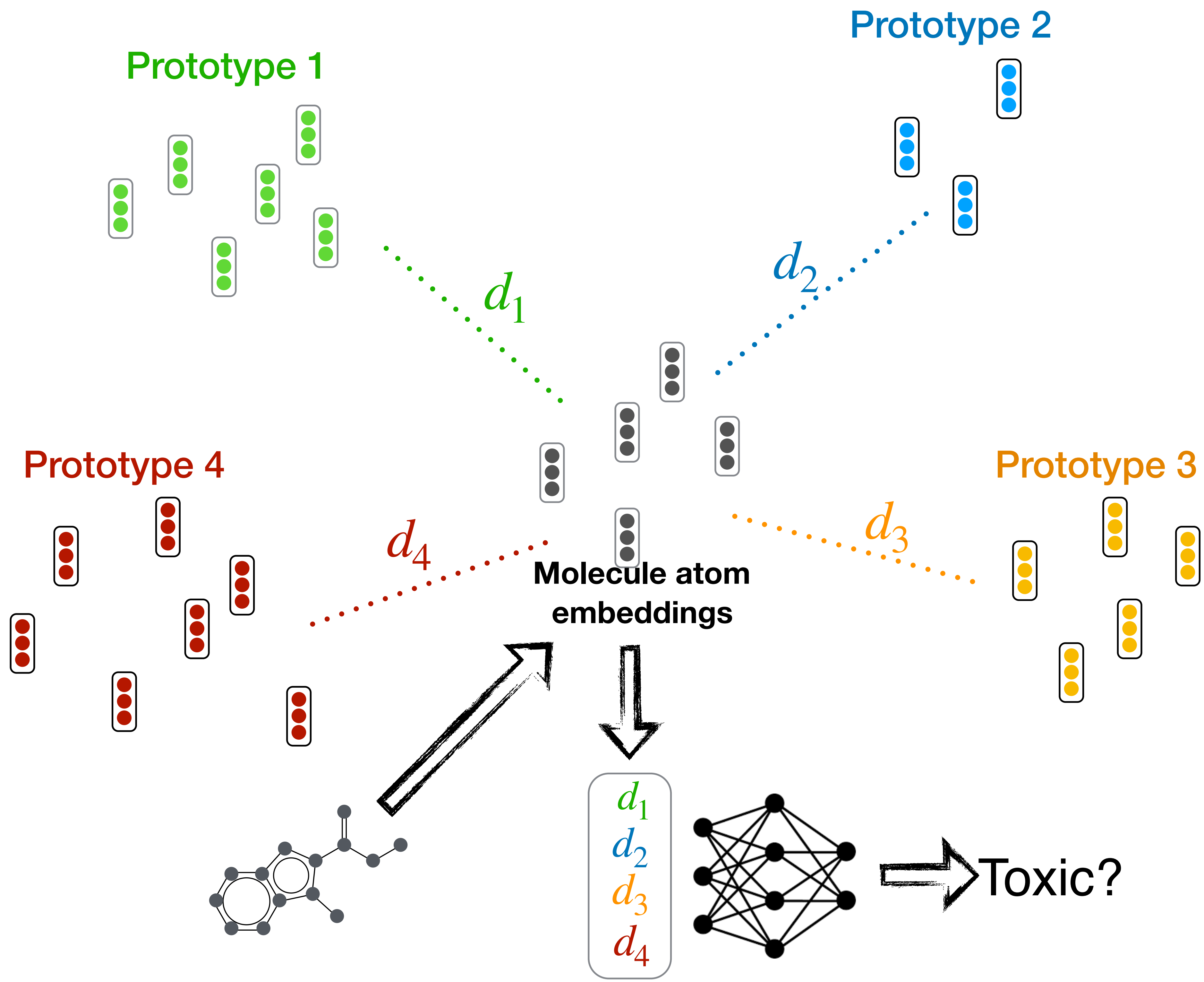
Prototypes as GCN Embedding Point Clouds

Our idea:

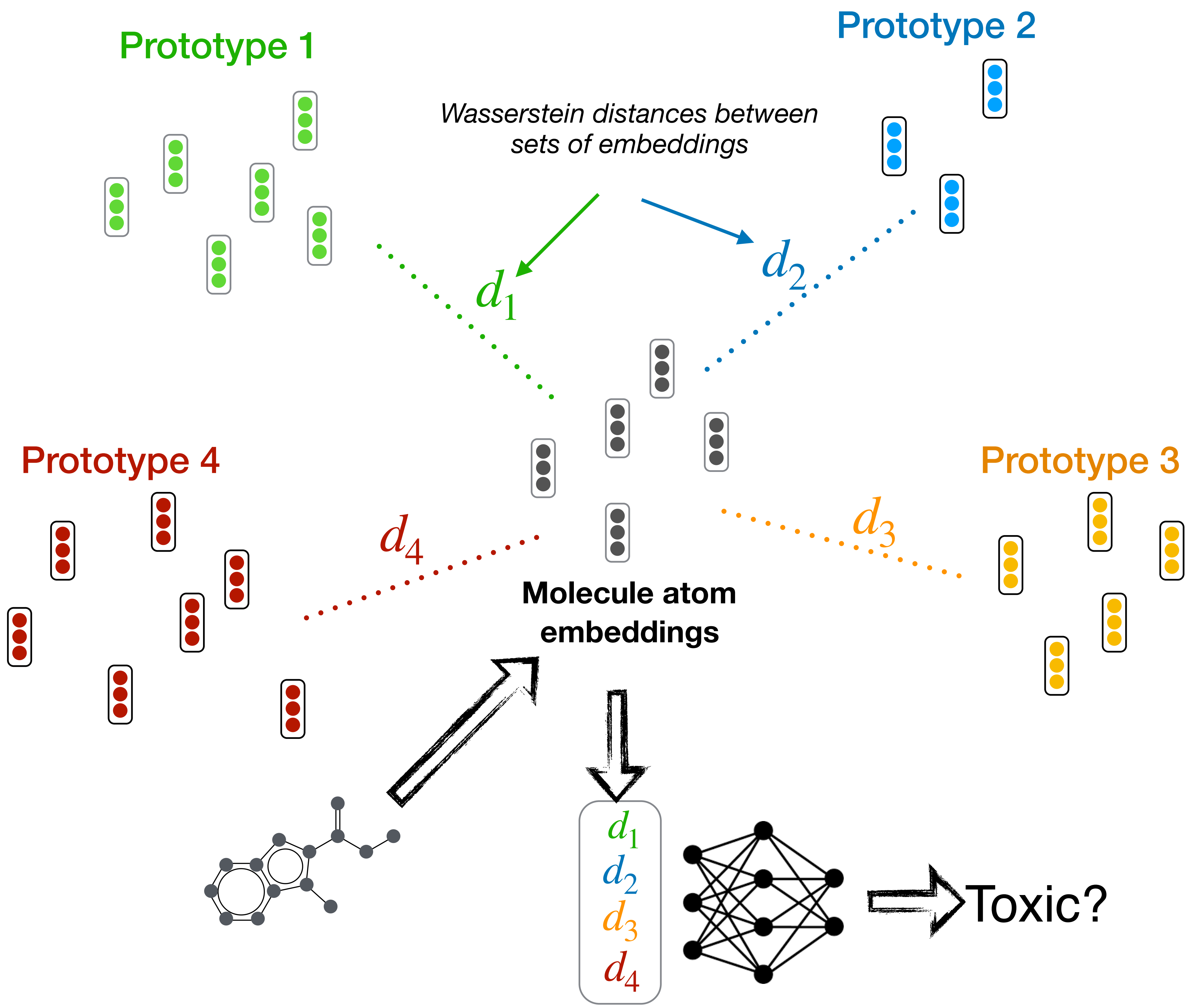
- *learn a dictionary of structural basis functions that serve to highlight key facets of compounds*
- *express molecules by relating them to abstract molecular prototypes*
- *prototypes highlight property values associated with different structural features (e.g. solubility)*



Wasserstein Prototypes



Wasserstein Prototypes



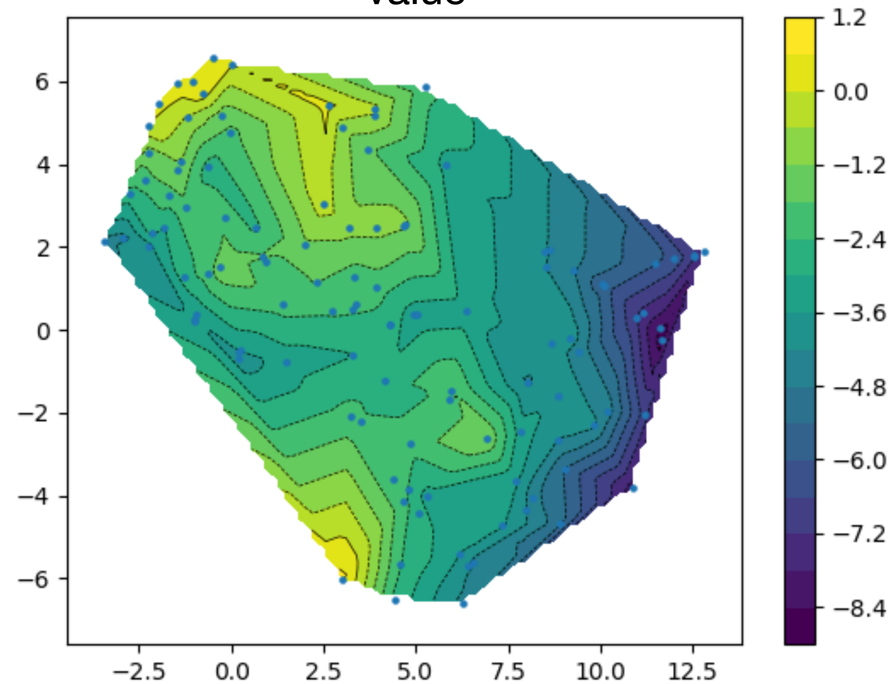
Property smoothness: Latent Space for Real Models

Graph Convolution w/ sum aggregation

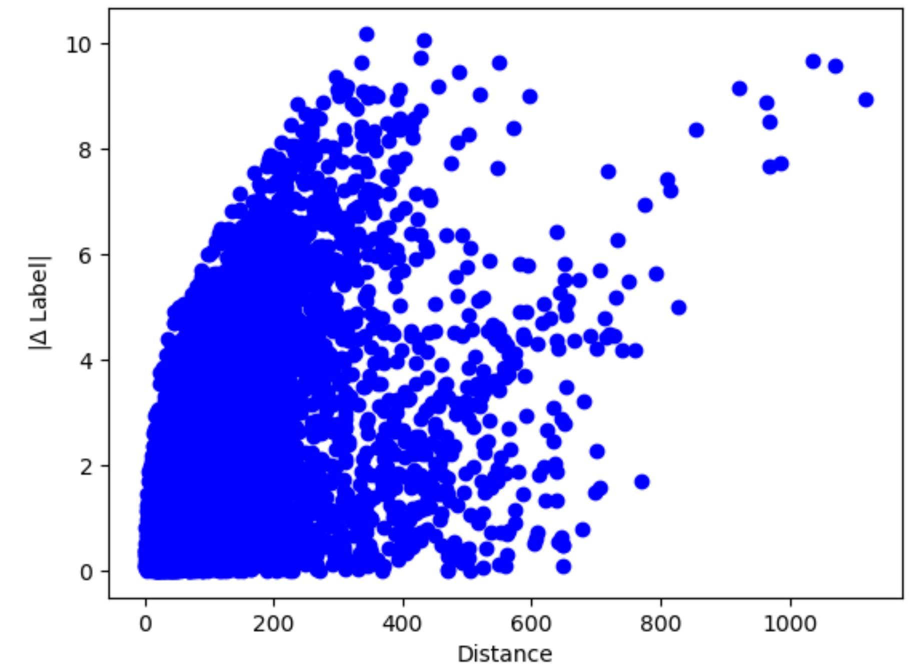
Spearman ρ = .424±.029

Pearson r = .393±.049

Molecule embedding in 2D vs property value



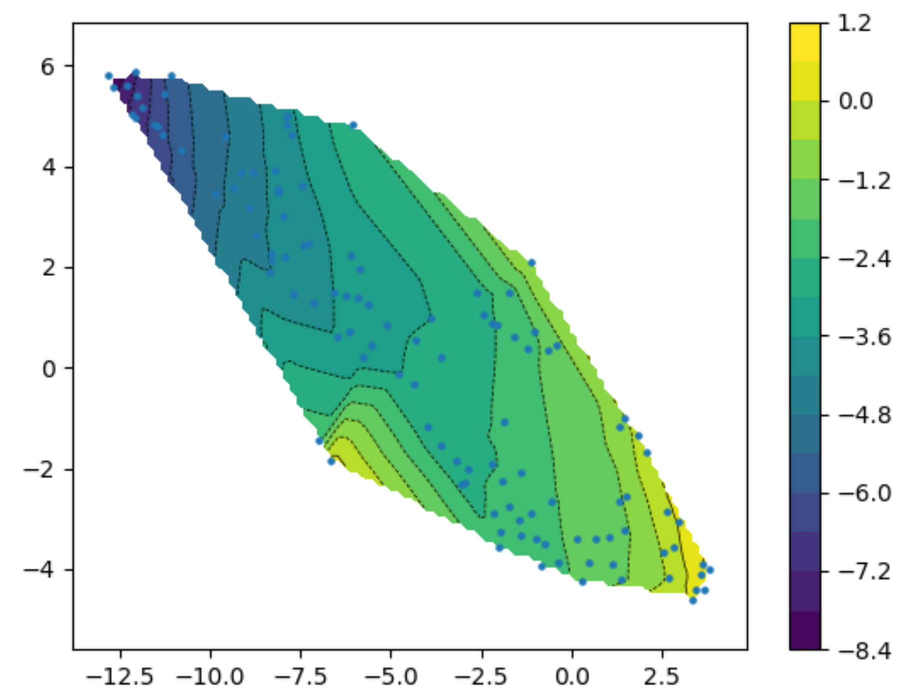
Molecule embedding distance vs property value distance

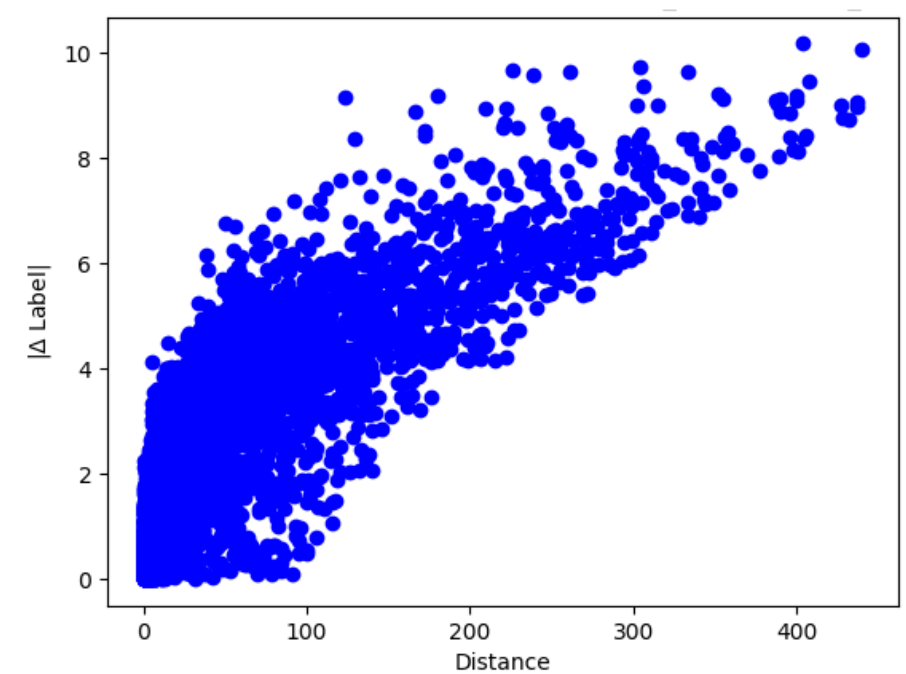


Wasserstein Prototypes

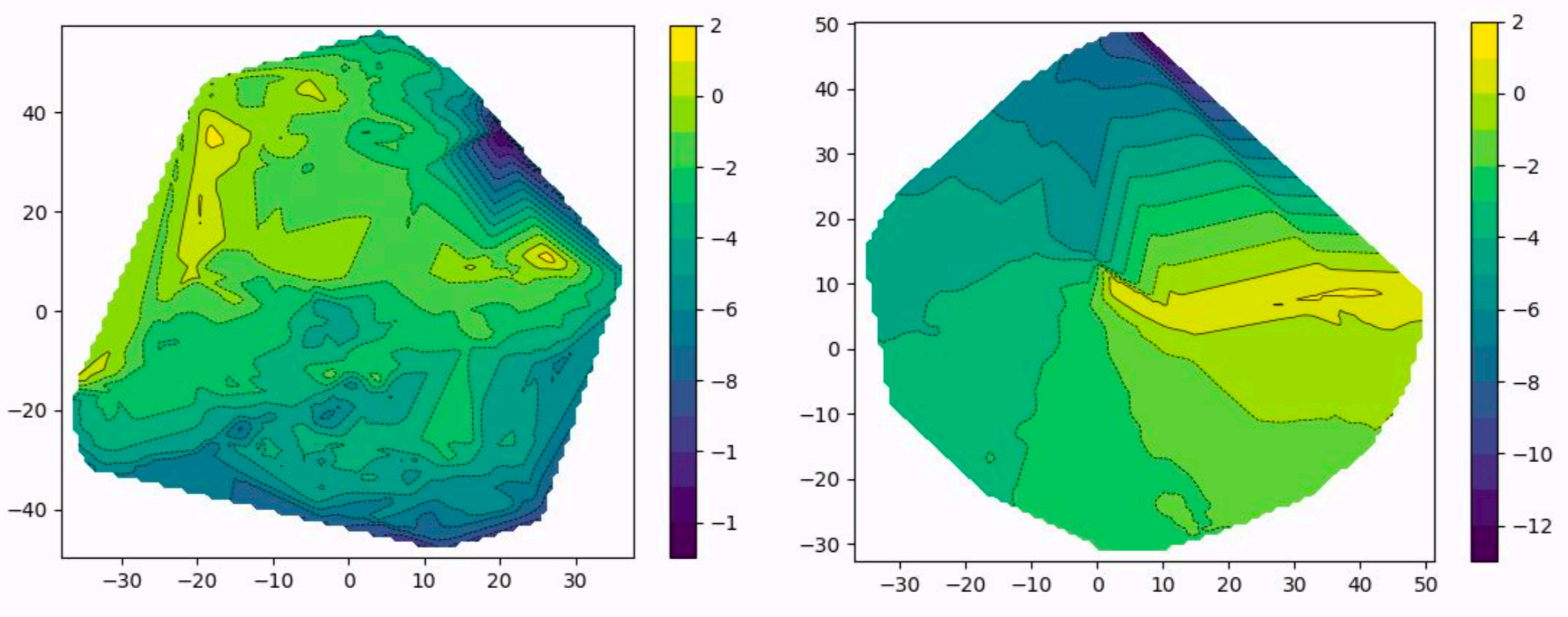
Spearman ρ = **.815±.026**

Pearson r = **.828±.020**



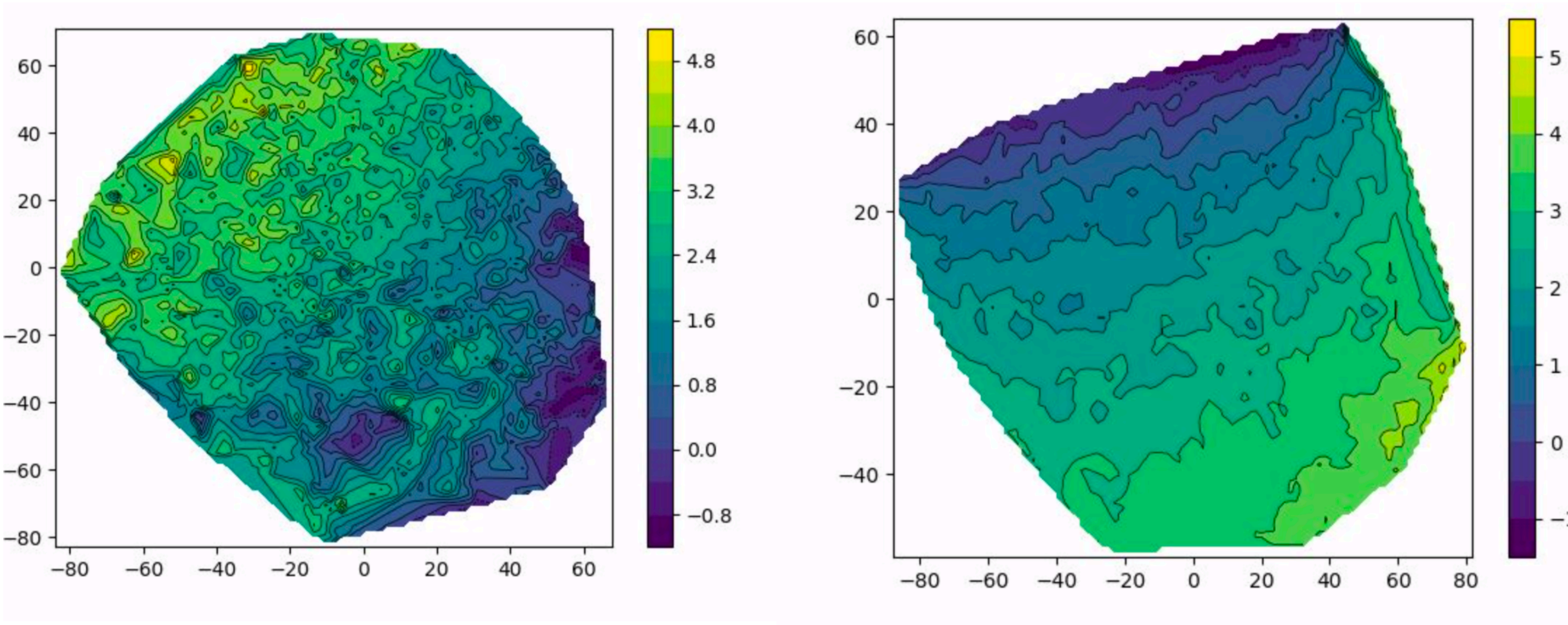


Property smoothness: Latent Space for Real Models



Solubility values in the latent space
Left: using sum aggregation; **Right:** using Wasserstein with prototypes

Property smoothness: Latent Space for Real Models



Lipophilicity values in the latent space
Left: using sum aggregation; **Right:** using Wasserstein with prototypes

Results on Molecular Property Prediction

	Solubility	Lipophilicity	Inhibitors of human β-secretase 1 (BACE-1)	Blood-brain barrier penetration (permeability)
# graphs / molecules	ESOL (RMSE) $\downarrow$ $n = 1128$	Lipo (RMSE) $\downarrow$ $n = 4199$	BACE (AUC) $\uparrow$ $n = 1512$	BBBP (AUC) $\uparrow$ $n = 2039$
GNN/Chemprop	$.635 \pm .027$	$.646 \pm .041$	$.865 \pm .013$	$.915 \pm .010$
ProtoS-L2	$.611 \pm .034$	$\mathbf{.580 \pm .016}$	$.865 \pm .010$	$.918 \pm .009$
ProtoW-Dot *(no reg.)*	$.608 \pm .029$	$.637 \pm .018$	$.867 \pm .014$	$.919 \pm .009$
ProtoW-Dot	$\mathbf{.594 \pm .031}$	$.629 \pm .015$	$.871 \pm .014$	$.919 \pm .009$
ProtoW-L2 *(no reg.)*	$.616 \pm .028$	$.615 \pm .025$	$\underline{.870 \pm .012}$	$\underline{.920 \pm .010}$
ProtoW-L2	$\underline{.605 \pm .029}$	$\underline{.604 \pm .014}$	$\mathbf{.873 \pm .015}$	$\mathbf{.920 \pm .010}$