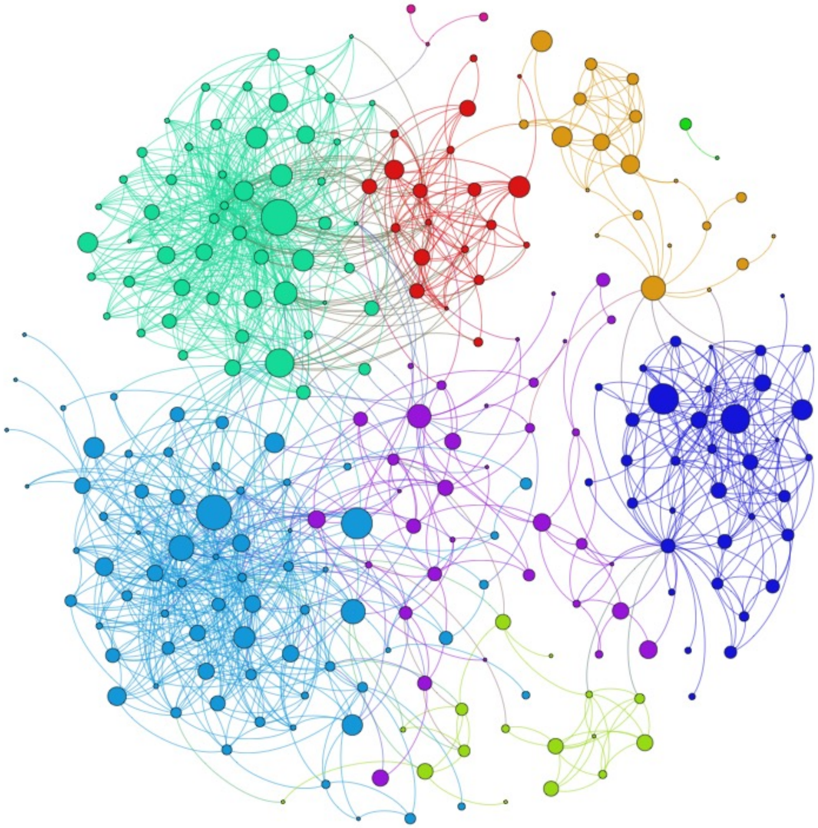
FISHNETS

T. Lucas Makinen [Imperial College London]

ML x Cosmo Conference, IAP, November 2023

with Justin Alsing + Ben Wandelt

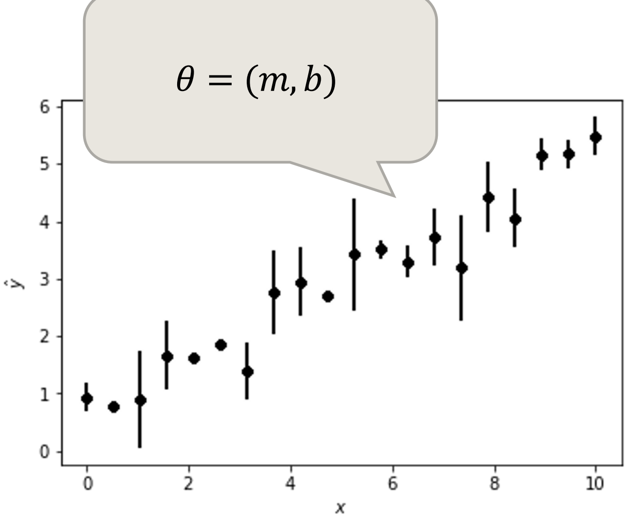
arxiv:2310.03812



FISHNETS

Given some data: $\{\mathbf{d}\} = \{\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}\}$

$$\ln p(\{\boldsymbol{d}\}|\boldsymbol{\theta}) = \sum_{i=1}^{n_{\text{data}}} \ln p(d_i|\boldsymbol{\theta})$$



FISHNETS

Given some data: $\{\mathbf{d}\} = \{\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}\}$

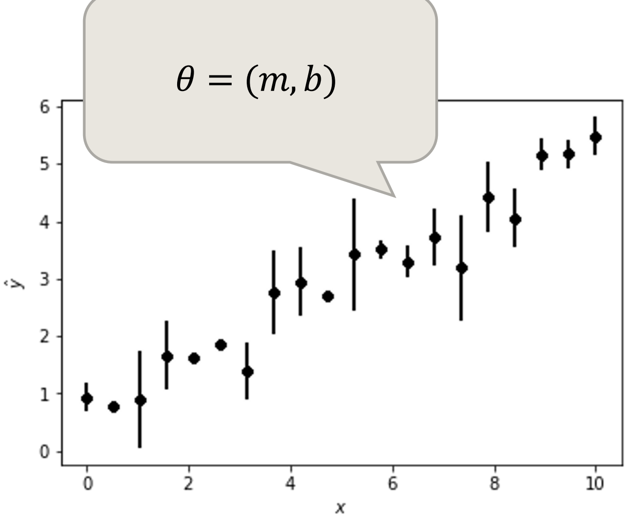
$$\ln p(\{\boldsymbol{d}\}|\boldsymbol{\theta}) = \sum_{i=1}^{n_{\text{data}}} \ln p(d_i|\boldsymbol{\theta})$$

Dataset score is additive !

$$\boldsymbol{t} = \nabla_\theta \ln p(\{\boldsymbol{d}\}|\boldsymbol{\theta})$$

$$= \sum_{i=1}^{n_{\text{data}}} \nabla_\theta \ln p(d_i|\boldsymbol{\theta})$$

$$= \sum_{i=1}^{n_{\text{data}}} \boldsymbol{t}_i(d_i)$$

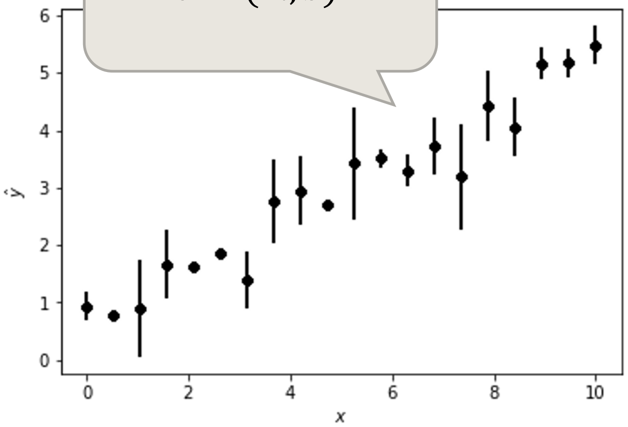


FISHNETS

Given some data: $\{\mathbf{d}\} = \{\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}\}$

$$\ln p(\{\boldsymbol{d}\}|\boldsymbol{\theta}) = \sum_{i=1}^{n_{\text{data}}} \ln p(d_i|\boldsymbol{\theta})$$

$\theta = (m, b)$



Dataset score is additive !

$$\boldsymbol{t} = \nabla_\theta \ln p(\{\boldsymbol{d}\}|\boldsymbol{\theta})$$

$$= \sum_{i=1}^{n_{\text{data}}} \nabla_\theta \ln p(d_i|\boldsymbol{\theta})$$

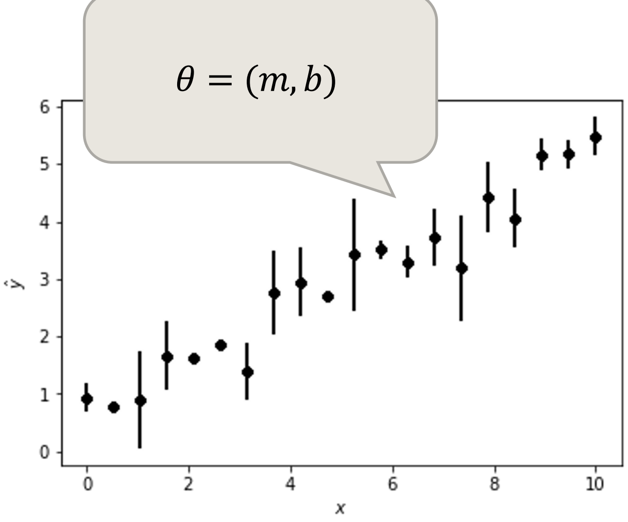
$$= \sum_{i=1}^{n_{\text{data}}} \boldsymbol{t}_i(d_i)$$

Fisher is also additive !

$$\mathbf{F}(\{\mathbf{d}\}) = \sum_{\boldsymbol{i=1}}^{n_{\text{data}}} \mathbf{F}_{\text{i}}(x_i, \sigma_i)$$

FISHNETS

Given some data: $\{\mathbf{d}\} = \{\mathbf{x}, \mathbf{y}, \boldsymbol{\sigma}\}$



$$\ln p(\{\boldsymbol{d}\}|\boldsymbol{\theta}) = \sum_{i=1}^{n_{\rm data}} \ln p(d_i|\boldsymbol{\theta})$$

Dataset score is additive !

$$\boldsymbol{t} = \nabla_\theta \ln p(\{\boldsymbol{d}\}|\boldsymbol{\theta})$$

$$= \sum_{i=1}^{n_{\rm data}} \nabla_\theta \ln p(d_i|\boldsymbol{\theta})$$

$$= \sum_{i=1}^{n_{\rm data}} \boldsymbol{t}_i(d_i)$$

Fisher is also additive !

$$\mathbf{F}(\{\mathbf{d}\}) = \sum_{\boldsymbol{i}=1}^{n_{\rm data}} \mathbf{F}_{\rm i}(x_i, \sigma_i)$$

learn with neural networks !

FISHNETS

create pseudo-MLE estimate

$$\widehat{\boldsymbol{\theta}}^{\mathrm{MLE}} = \boldsymbol{F}_{NN}^{-1}\boldsymbol{t}_{NN} + \boldsymbol{\theta}_{fid}$$

Train the networks using a "KL"-type loss:

$$\mathcal{L} = \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{\mathrm{MLE}})^T \mathbf{F}_{NN}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{\mathrm{MLE}}) - \frac{1}{2} \ln \det \mathbf{F}_{NN}$$

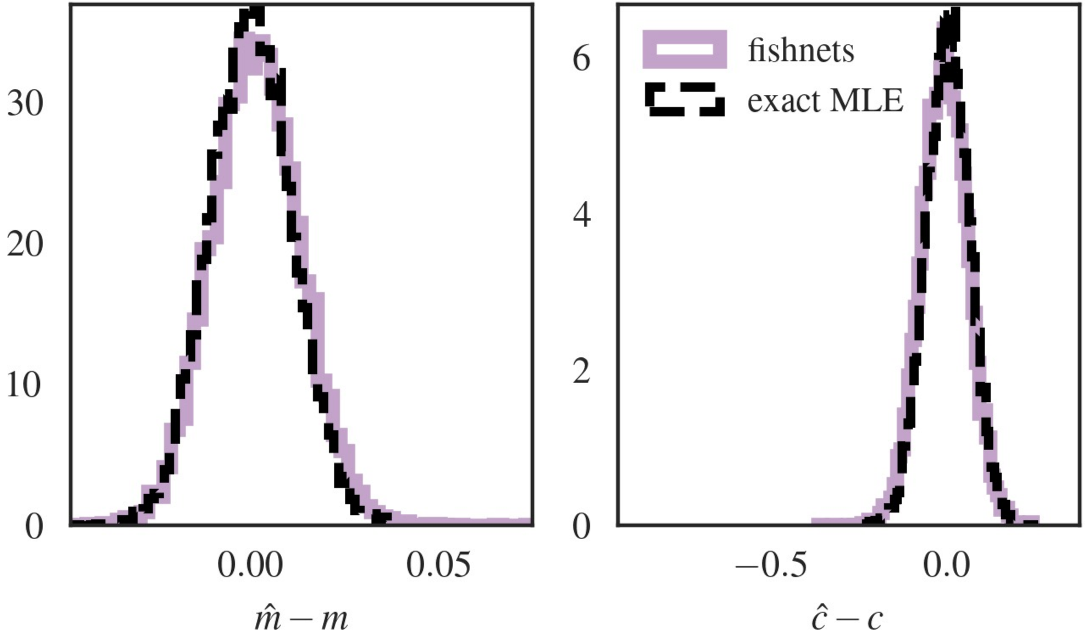
Train the networks using a "KL"-type loss:

fisher network

$$\mathcal{L} = \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{\mathrm{MLE}})^T \mathbf{F}_{NN}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{\mathrm{MLE}}) - \frac{1}{2} \ln \det \mathbf{F}_{NN}$$

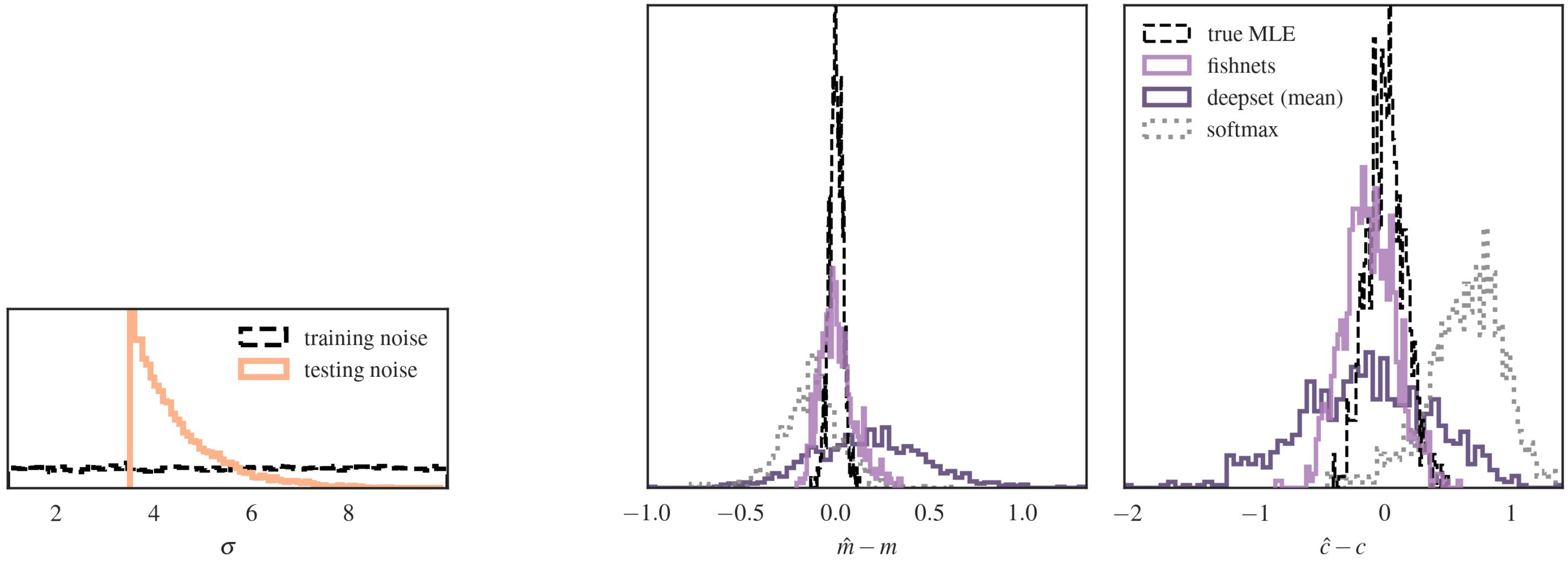
score network

optimal for linear regression ...



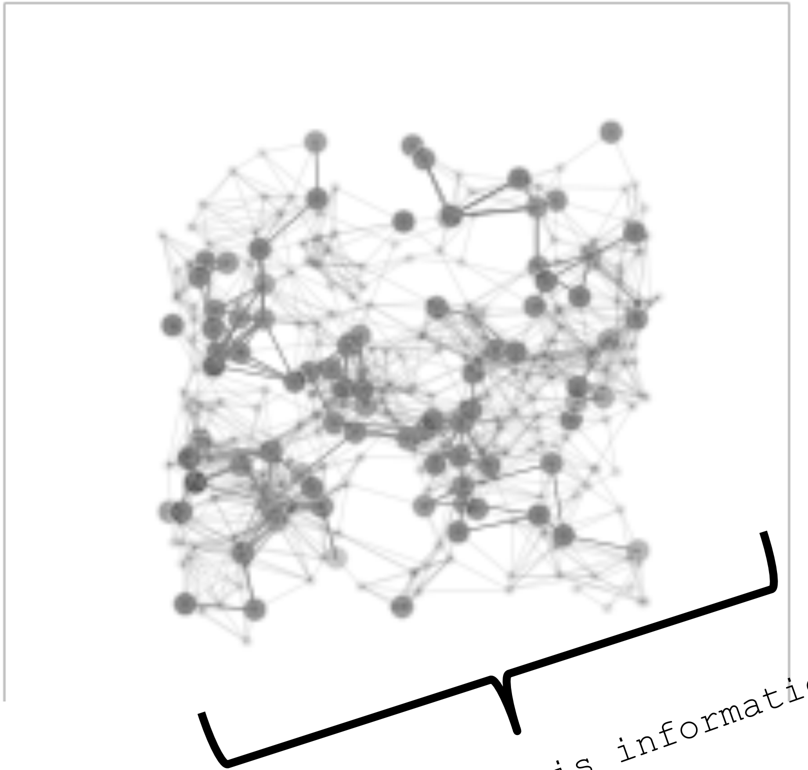
arxiv:2310.03812

... even when the global noise distribution changes !



arxiv:2310.03812

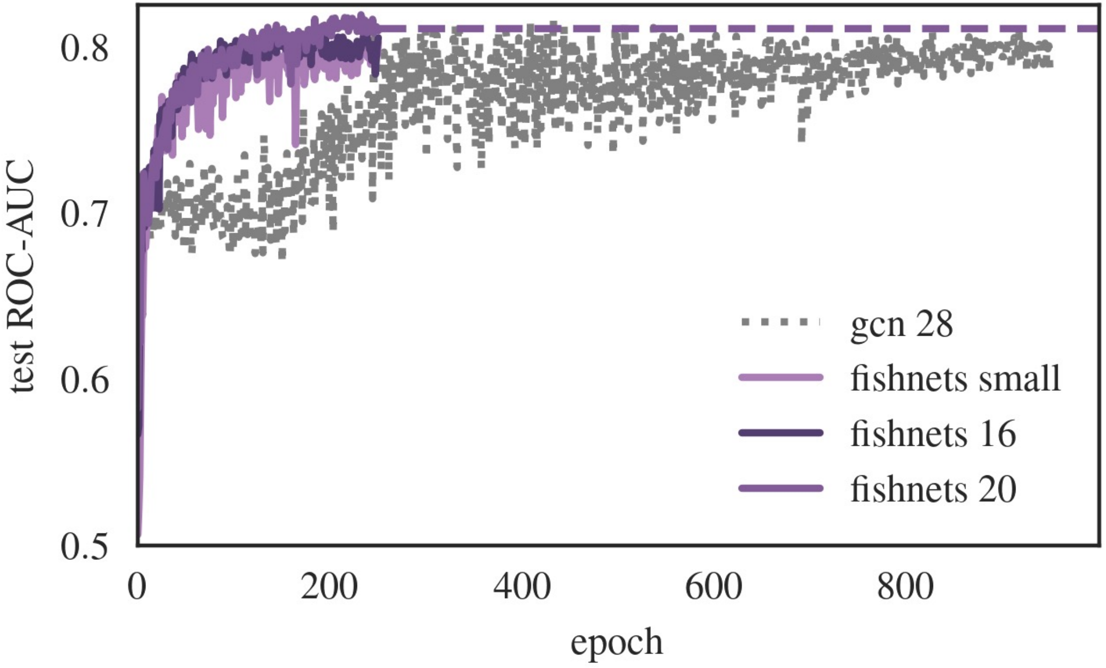
what about graphs ?



aggregate this information !

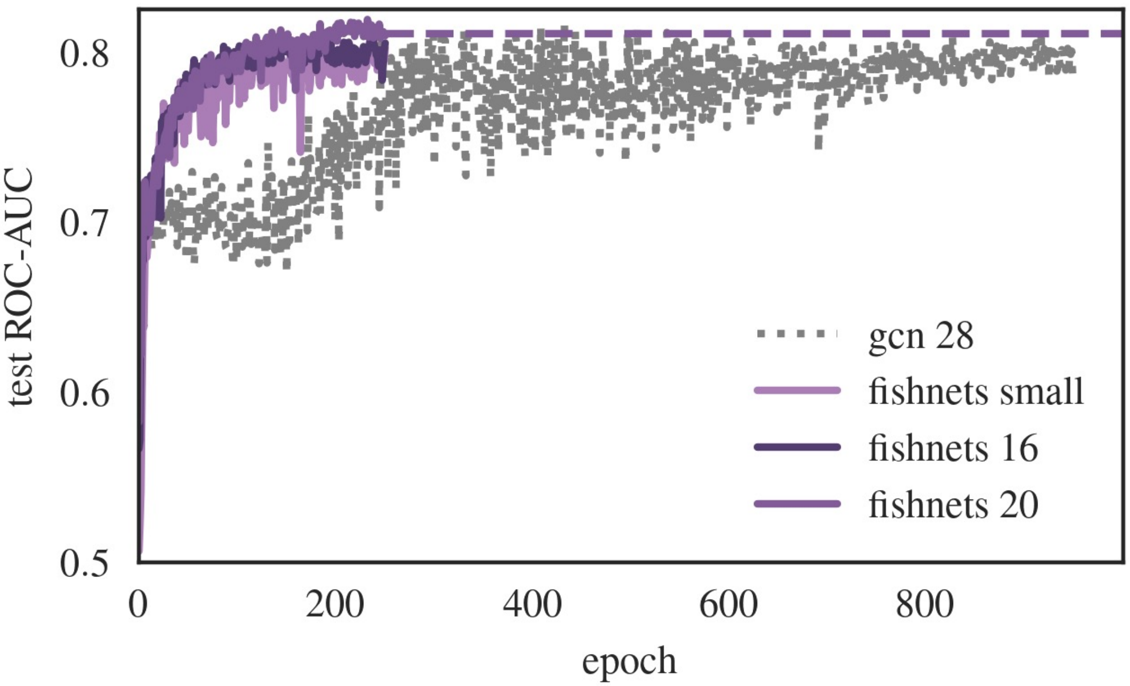
arxiv:2310.03812

Use as a drop-in replacement for graph network aggregation



arxiv:2310.03812

improve information embedding + extraction for common benchmarks



arxiv:2310.03812