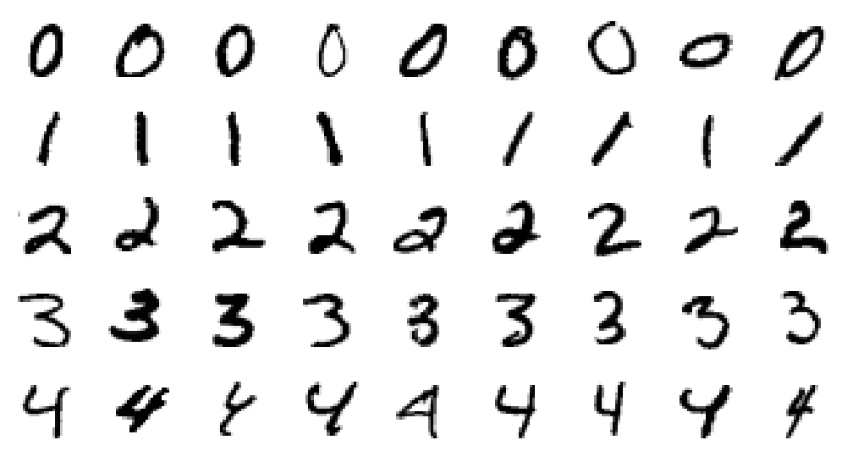
Perturbation Theory for the Information Bottleneck

Wave Ngampruetikorn & David J Schwab

Initiative for the Theoretical Sciences, The Graduate Center, CUNY

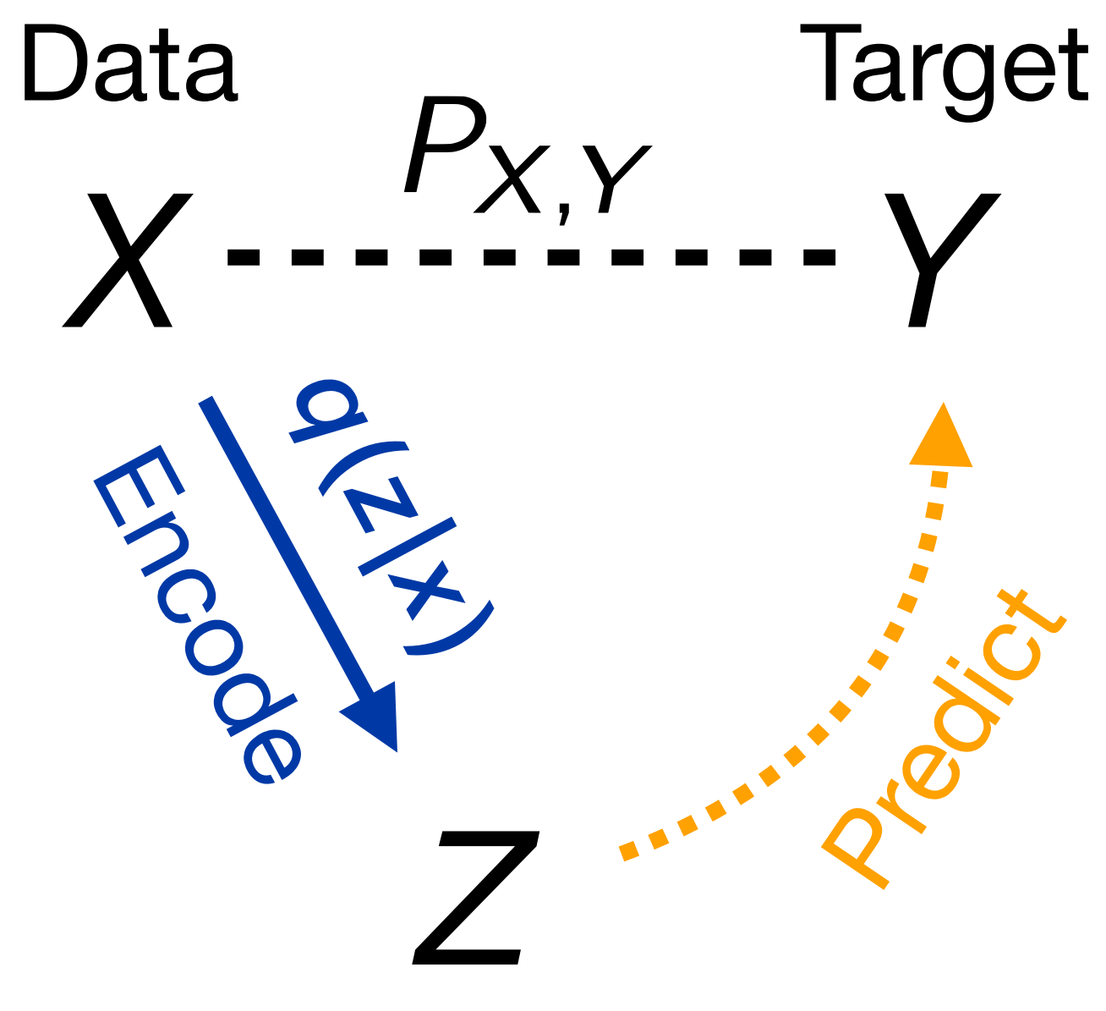
Extracting 'relevant' information from data underpins all forms of learning



Classifying handwritten digits requires extracting the *right* features from the space of pixels

The *right* features depend on what we want to know, *eg*, the digit, identity of writer, or brand of pen/pencil

Relevant information is the bits that can predict the 'target'



Representation of relevant information in X

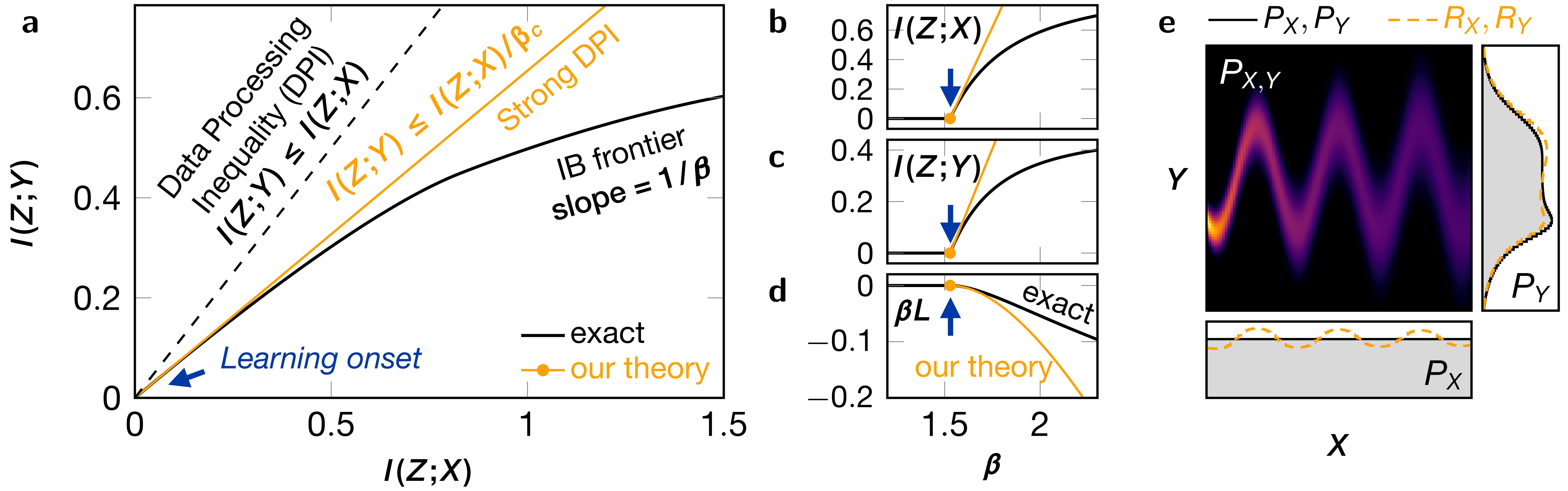
Encoder $q(z|x)$ defines a mapping from X to Z that compresses X by discarding irrelevant information

In information bottleneck (IB), relevant information extracting is optimization

$$\min_{q(z|x)} L \quad \text{with} \quad L = \underbrace{I(Z;X)}_{\text{favors compression}} - \beta \underbrace{I(Z;Y)}_{\text{favors prediction}}$$

β controls compression-prediction trade-off

While precise and appealing, the IB problem is analytically intractable in general



Our theory predicts the maximum 'relevant' information ratio ($1/\beta_c$) and connects the learning onset in the information bottleneck problem to the strong data processing inequality

Relevant information per extracted bit must be bounded from above

data processing inequality

$$\text{Relevant information ratio} = \frac{I(Z;Y)}{I(Z;X)} \le \beta_c^{-1} \le 1$$

strong data processing inequality

- **$1/\beta_c$ = maximum relevant information ratio**
- **No informative representation for $\beta \le \beta_c$**
- **$\beta = \beta_c$ marks the IB learning onset**

Our analytical results for learning onset have potential implications in fundamental research and in practice

Maximum relevant information ratio (**$1/\beta_c$**)

- Related to contraction coefficient in information theory [Anantharam etal arXiv:1304.6133]
- Tight bound of the thermodynamic efficiency in predictive systems [Still etal PRL 2012]
- Useful measure of correlations [Kim etal NeurIPS 2017]
- Might help tune hyperparameters in deep learning techniques such as VIB [Wu etal Entropy 2019]

vngampruetikorn@gc.cuny.edu