Manifold Learning for Complex Visual Analytics:

Benefits from and to Neural Architectures

Stephane Marchand-Maillet
Viper group
University of Geneva
Switzerland

Edgar Roman-Rangel, Ke Sun (Viper)
A. Agocs, D. Dardanis, R. Forster, J.-M. Le Goff, X. Ouvrard (CERN)

Outline

- Visual Analytics and Manifold Learning
- Deriving manifold
  - Learning strategies
  - Spacetime
  - Information geometry
- Make Manifold Learning inductive with Neural Architectures
- Application potential: Visualising Neuroscience data

Manifold learning

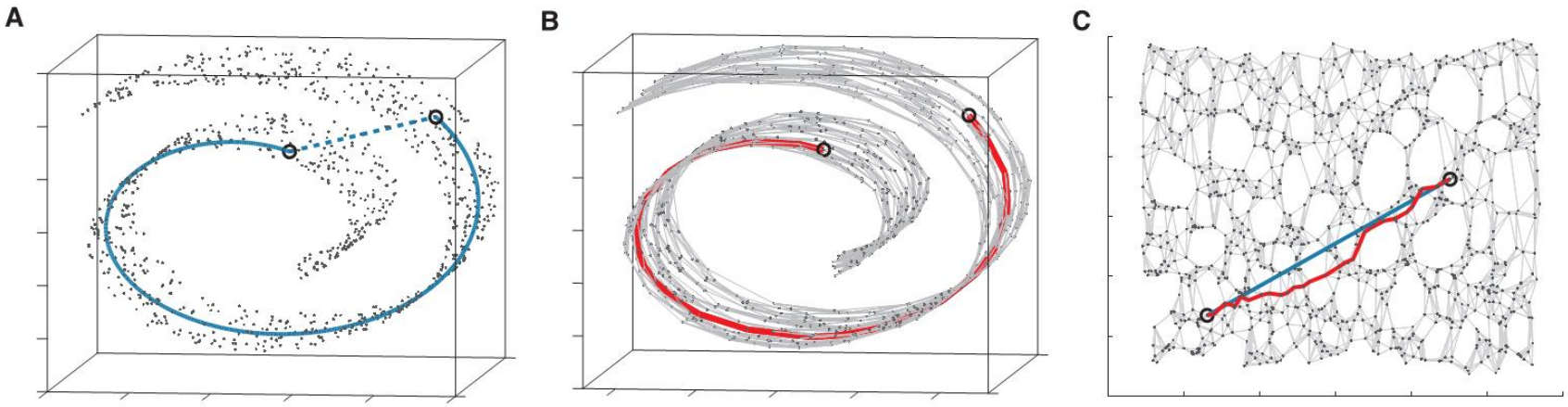


Fig 3. from J. B. Tenenbaum, V. de Silva, J. C. Langford, A Global Geometric Framework for Nonlinear Dimensionality Reduction, Science 290, (2000), 2319–2323

- Choice of features
- Preservation of local information
  - MDS : preserve exact neighborhood
  - t-SNE : preserve neighborhood distribution
- At the heart of visualisation (and Visual Analytics)

Preserving local information

$$\text{Given} \quad \{x_i\}_{i=1\ldots N} \in \Re^D \quad \text{find} \quad \{y_i\}_{i=1\ldots N} \in \Re^d \quad (d < D)$$

- Distance-based

$$d_{ij} = \left\| x_i - x_j \right\| \quad \delta_{ij} = \left\| y_i - y_j \right\| \Rightarrow \min_y \frac{\sum_i \sum_j w_{ij} (d_{ij} - \delta_{ij})^2}{scale}$$

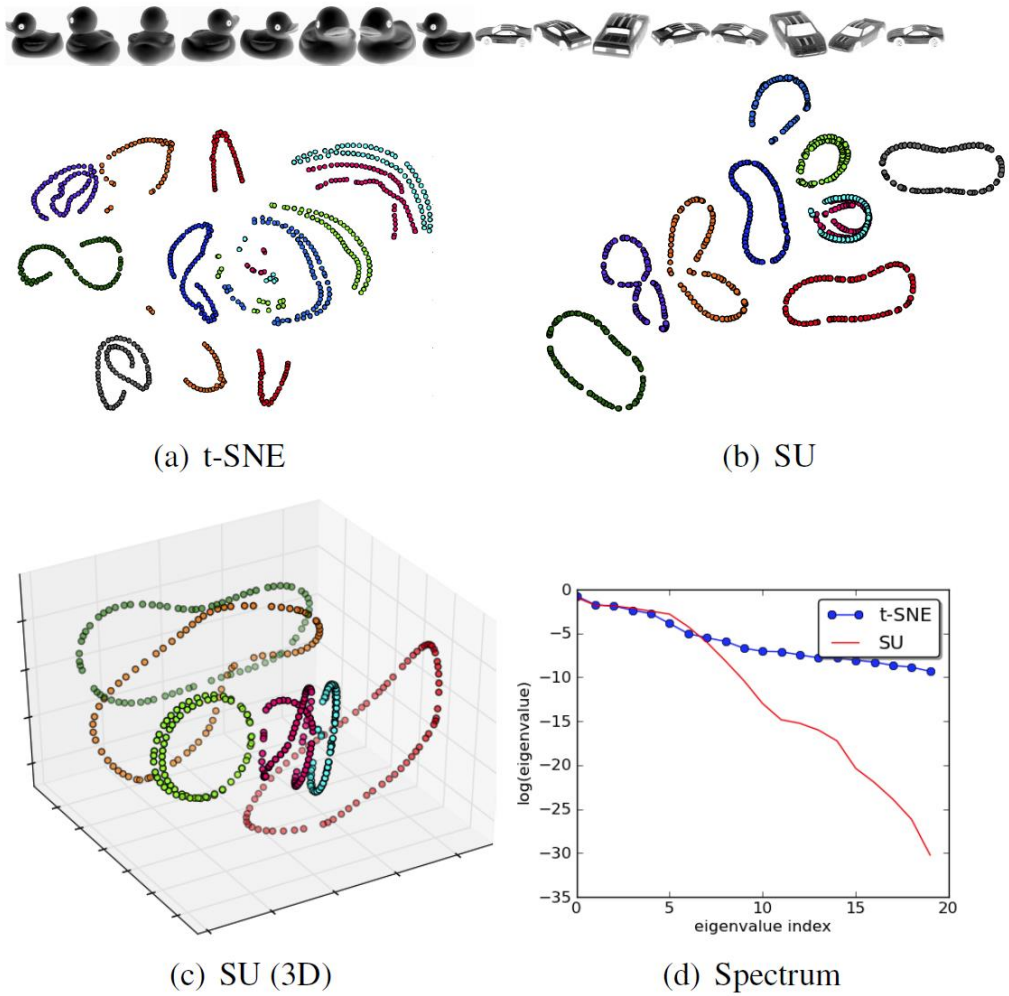
- Stochastic neighbourhood

$$p_{j|i} = \frac{h\left(\left\| x_i - x_j \right\|^2\right)}{\sum_{j \neq i} h\left(\left\| x_i - x_j \right\|^2\right)} \quad q_{j|i} = \frac{h\left(\left\| y_i - y_j \right\|^2\right)}{\sum_{j \neq i} h\left(\left\| y_i - y_j \right\|^2\right)}$$

$$E(\mathbf{y}) = -\sum_{i=1}^{N} \sum_{j:j \neq i} q_{j|i}(\mathbf{y}) \log p_{j|i}$$

Sun, K., Bruno, E., & Marchand-Maillet, S. (2012). *Stochastic Unfolding*. In IEEE Machine Learning for Signal Processing Workshop (MLSP'2012), Santander, Spain.

Stochastic Unfolding (SU)



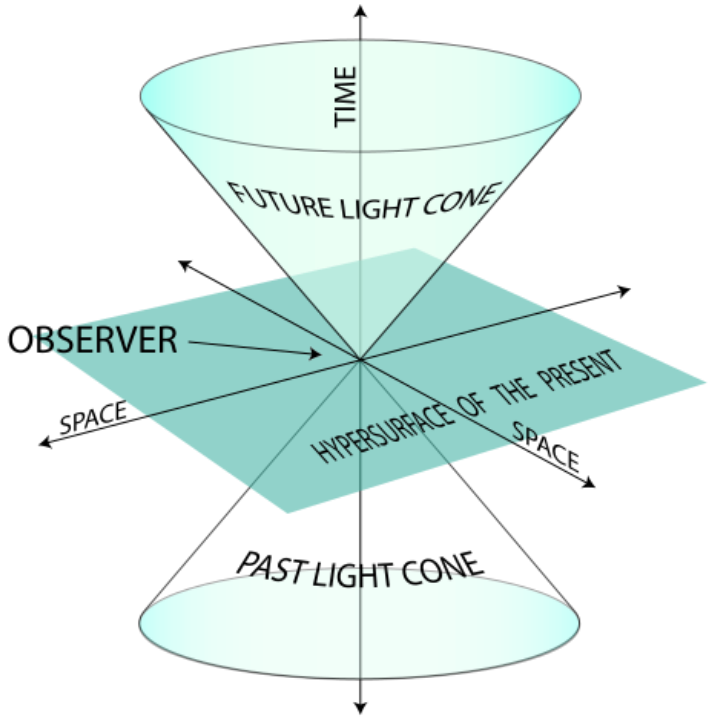
(a) t-SNE (b) SU

(c) SU (3D) (d) Spectrum

Sun, K., Bruno, E., & Marchand-Maillet, S. (2012). *Stochastic Unfolding*. In IEEE Machine Learning for Signal Processing Workshop (MLSP'2012), Santander, Spain.

Extension to spacetime

- Use relativistic pseudo-metric tensor for including a “time” (negative) dimension



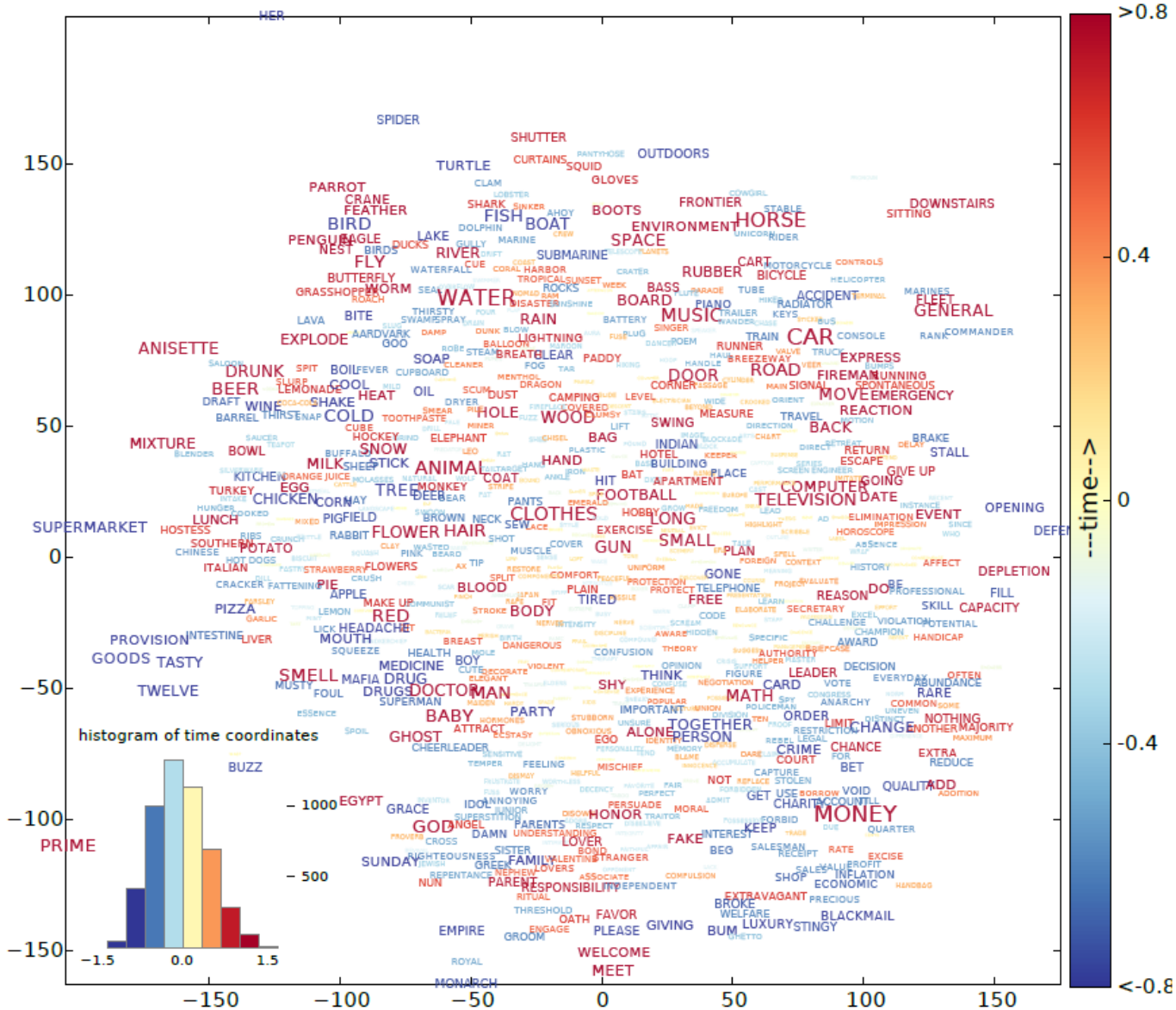
Similar stochastic embedding formulation using

$$c(x,y)^2 = \sum_{space}(x_i - y_i)^2 - \sum_{time}(x_i - y_i)^2$$

Provides more power for representation

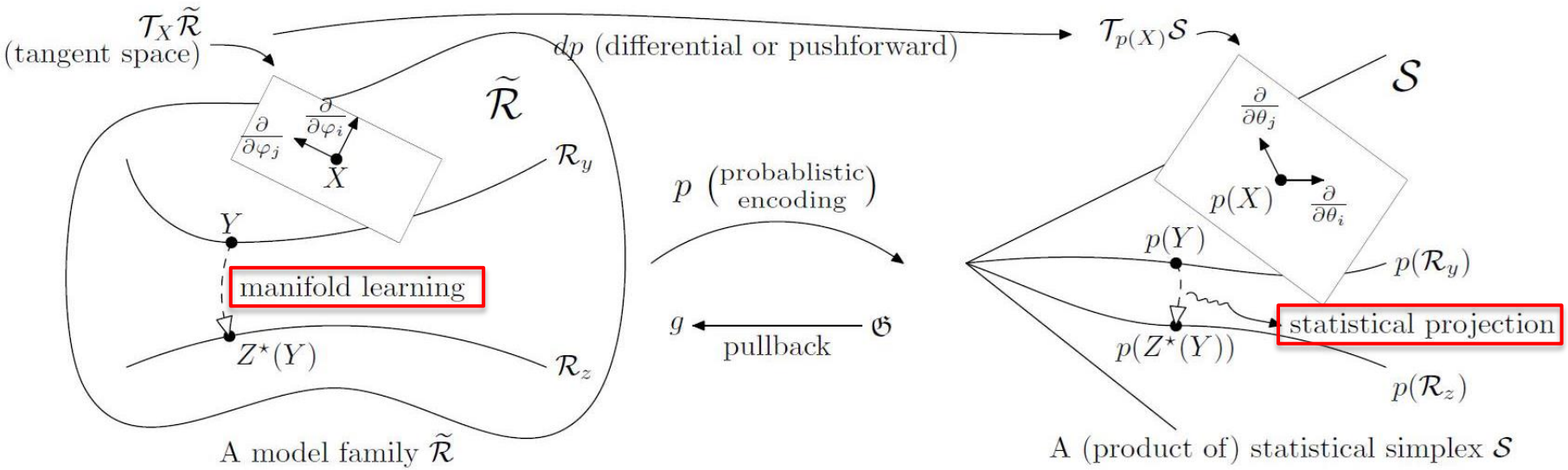
Sun, K., Wang, J., Kalousis, A., & Marchand-Maillet, S. (2015). *Space-Time Local Embeddings*. In Proceedings of Advances in Neural Information Processing Systems 28 (NIPS 2015), Montreal, Canada, December 2015.

Visualising Spacetime



A geometric view of Machine Learning

Information Geometry allows use to consider statistical machine learning as geometric operations (eg projections) over statistical manifolds



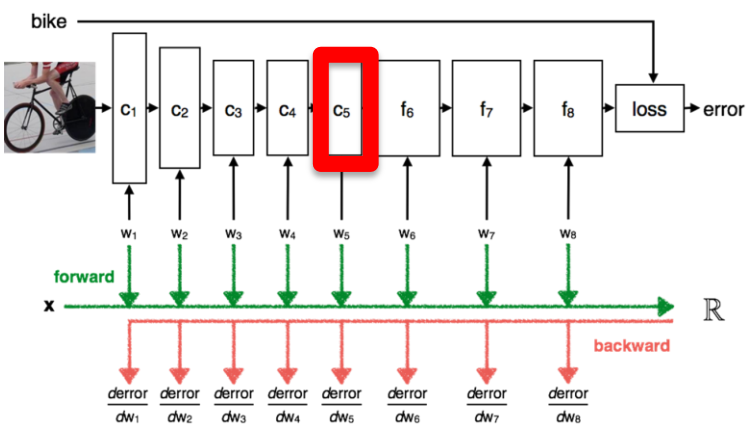
Sun, K., & Marchand-Maillet, S. (2014). An Information Geometry of Statistical Manifold Learning. In Proceedings of the International Conference on Machine Learning (ICML 2014), Beijing, China.

Embarking Neural Architectures

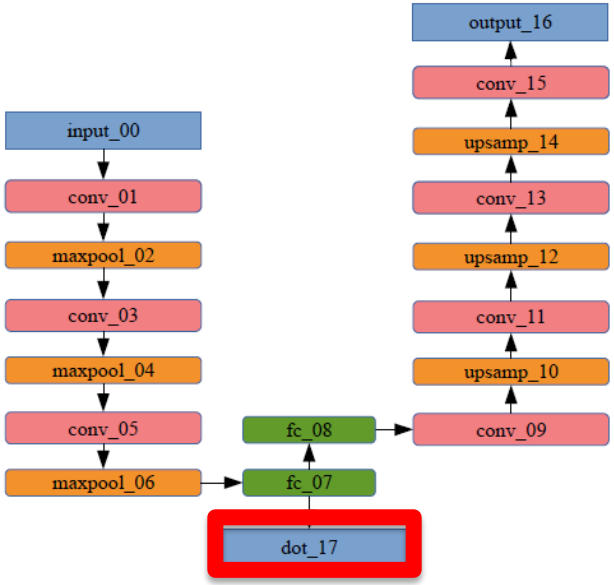
… as feature extractors

- We use the representation derived internally by Deep Learning architectures as input dimensions

c_5 in VGGNet



final encoding layer from (adapted) sparse autoencoders



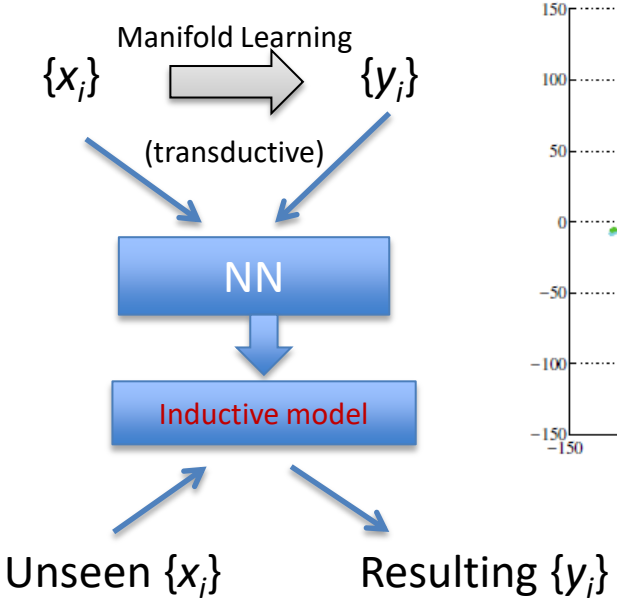
E. Roman-Rangel & S. Marchand-Maillet. COLD: Linearly Aggregated Convolutional Orthogonal Descriptors. *Submitted to the Int. Conference on Comp. Vision. 2017.*

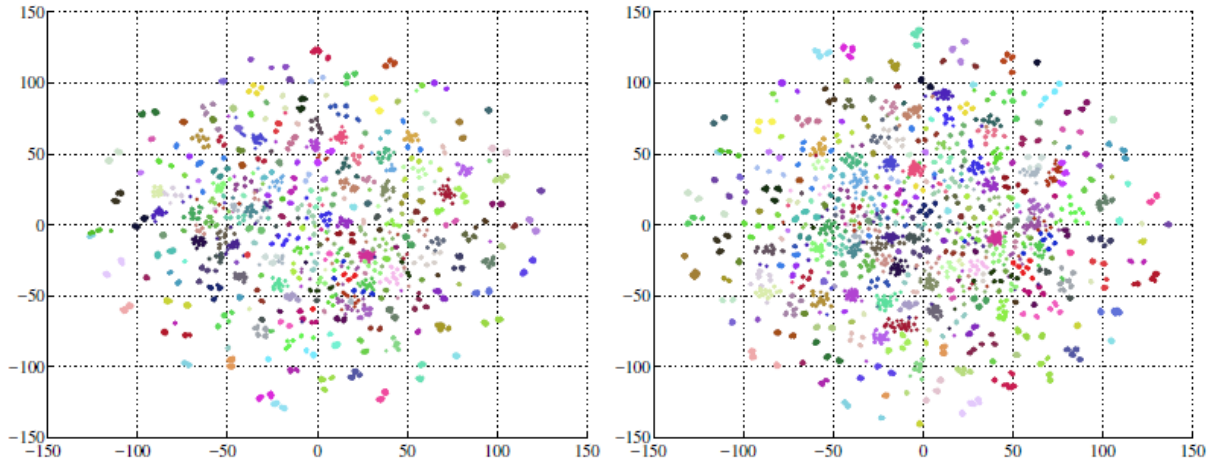
Embarking Neural Architectures

… as mappers

- Manifold Learning techniques are transductive
  - No absolute mapper learnt

→ We use Neural Architectures to make them inductive





(a) Training set only. (b) Training and test set together.

E. Roman-Rangel & S. Marchand-Maillet. Assessing Deep Learning Architectures for Visualizing Maya Hieroglyphs. *MCPR 2017.*

A Visual Analytics platform for Big Data Case of Neuroscience

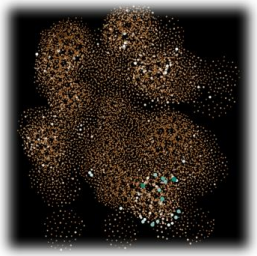
A. Agocs, D. Dardanis, R. Forster, J.-M. Le Goff, X. Ouvrard

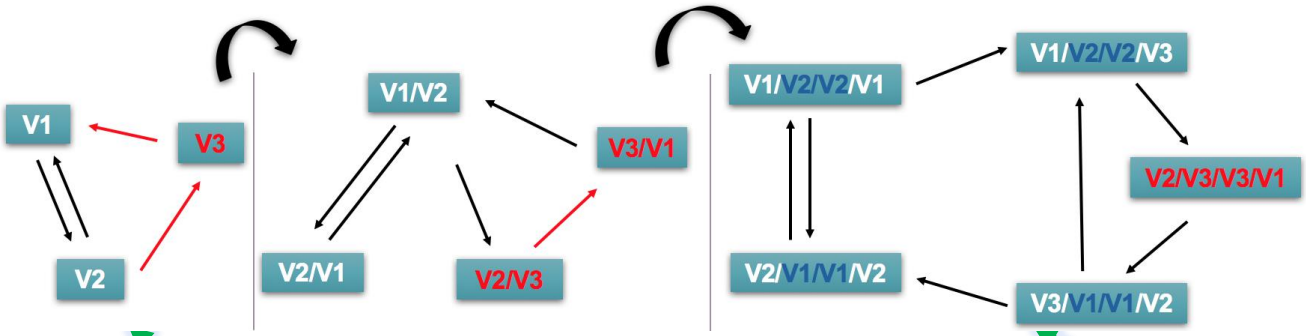
CERN

The macaque case

Linear graph model of the network of cortical interactions

- Derived a second linear graph of the visio-tactile network
- Projection of the derived graph back to the original network
- Characterize the nodes which are responsible for the information transmission





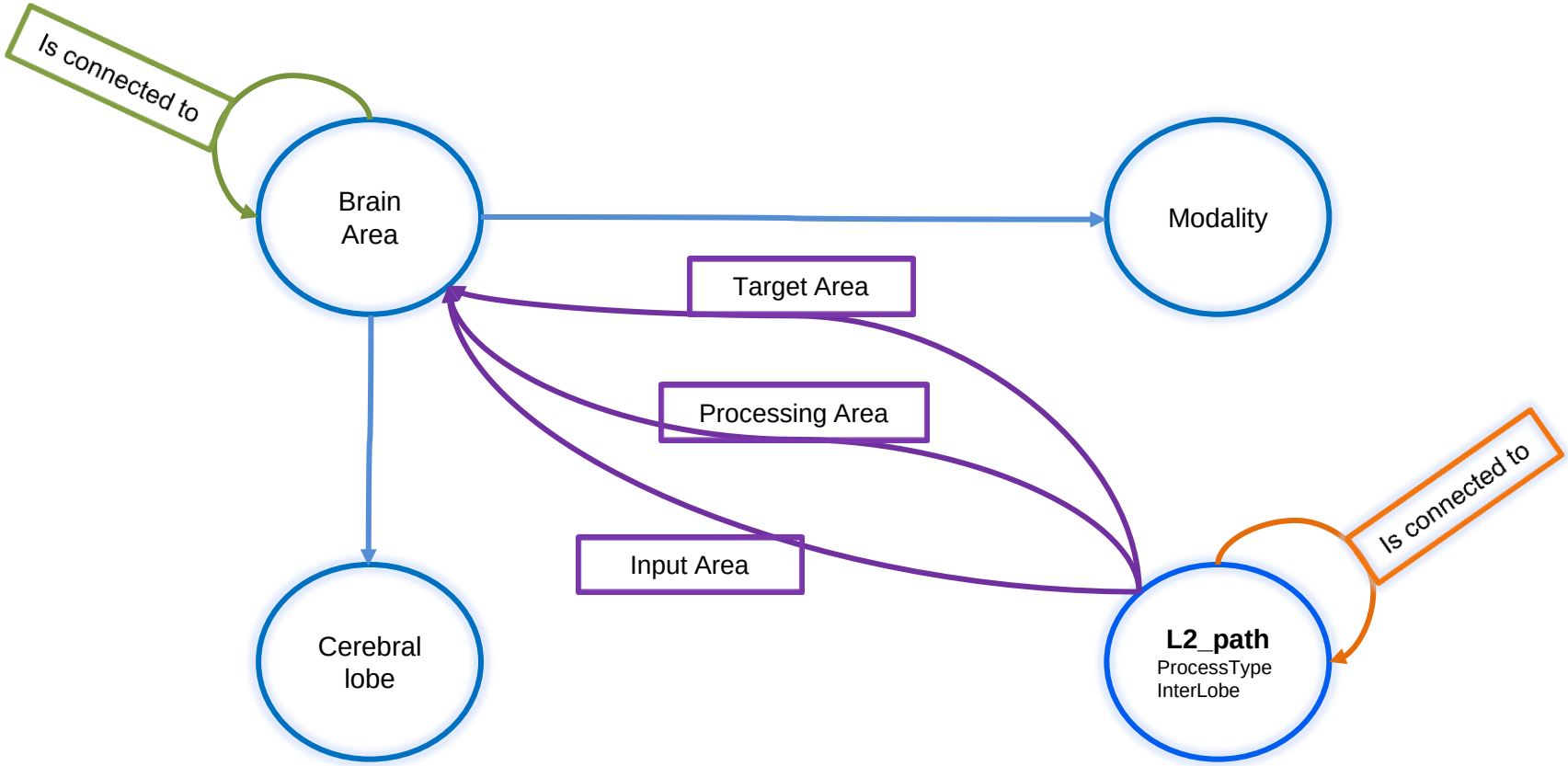
g2 is too large for visual perception
➔ Communities
172 clusters
10668 edges

g_0: directed graph of brain area interconnectivity*
(**42** vertices = areas, **601** edges= interactions)

g_2: directed graph of cortical interactions*
(Input/Processing/Target)
(**9869** vertices = IPT flows, **166219** edges = common interactions)

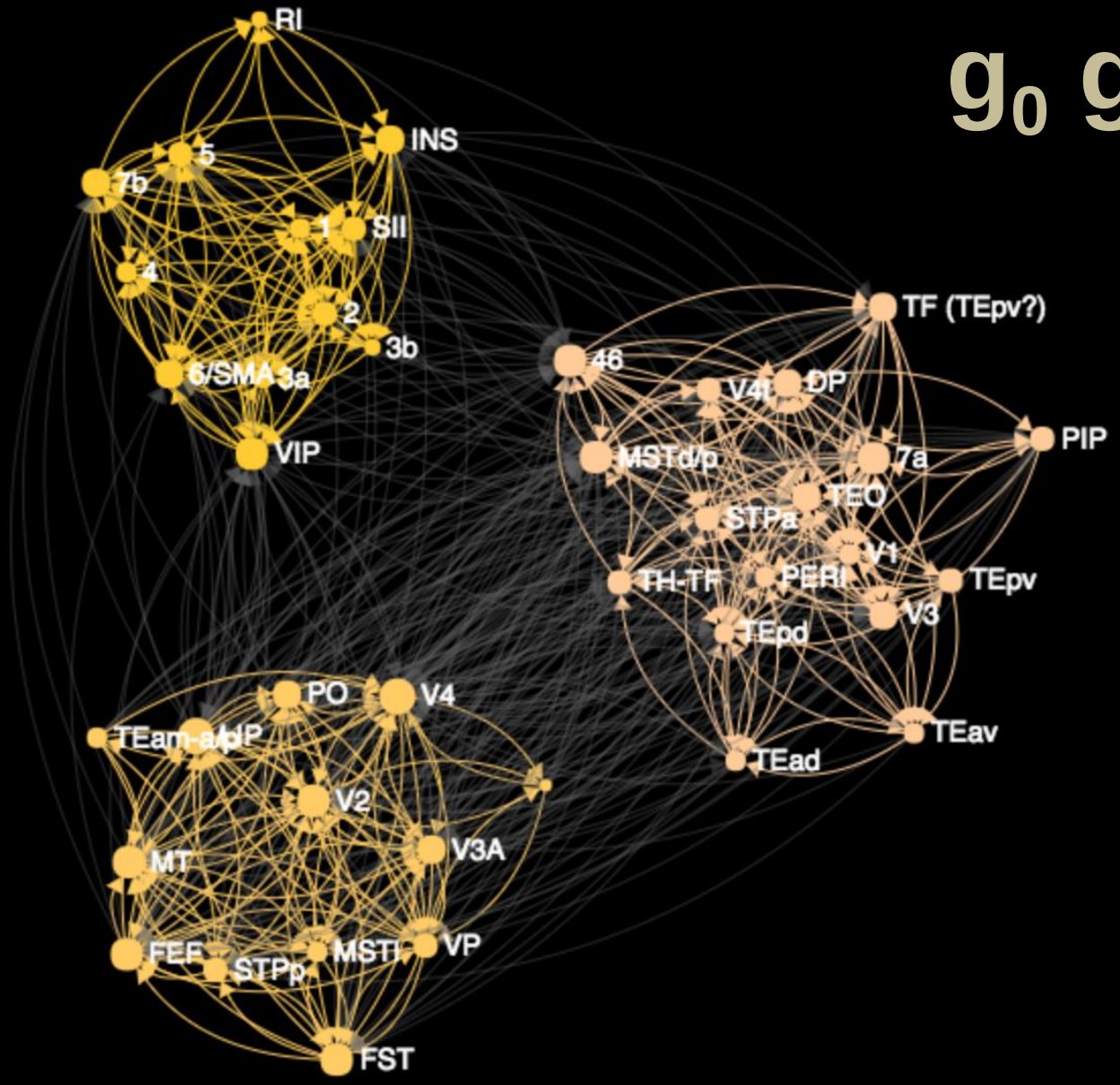
*Data/slide: L. Négyessy, A. Fülöp/Wigner Institute, Budapest

Constructed Reachability Graph

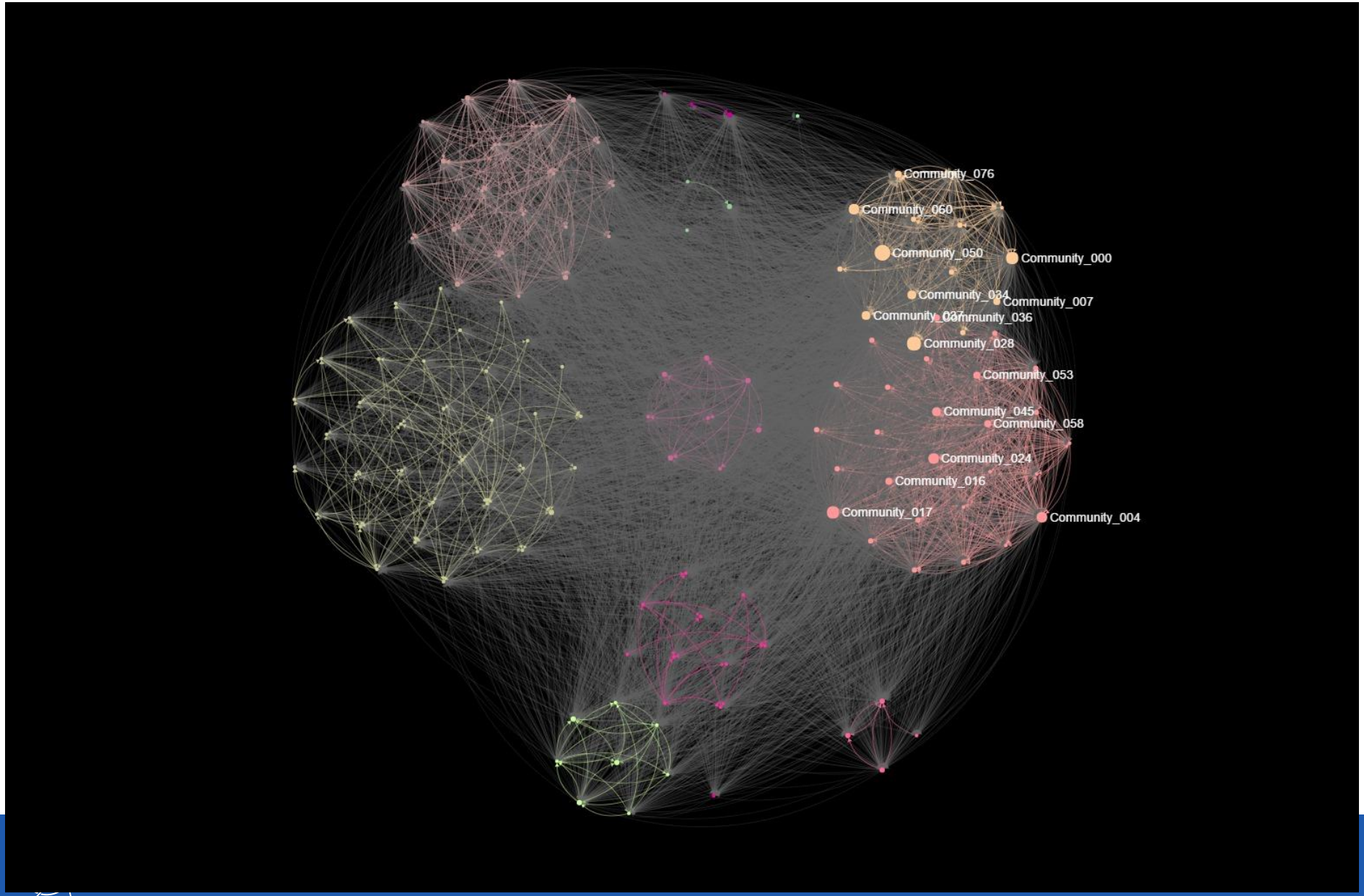


→ g_0 edges
→ g_2 edges
→ g_2 → g_0 connections

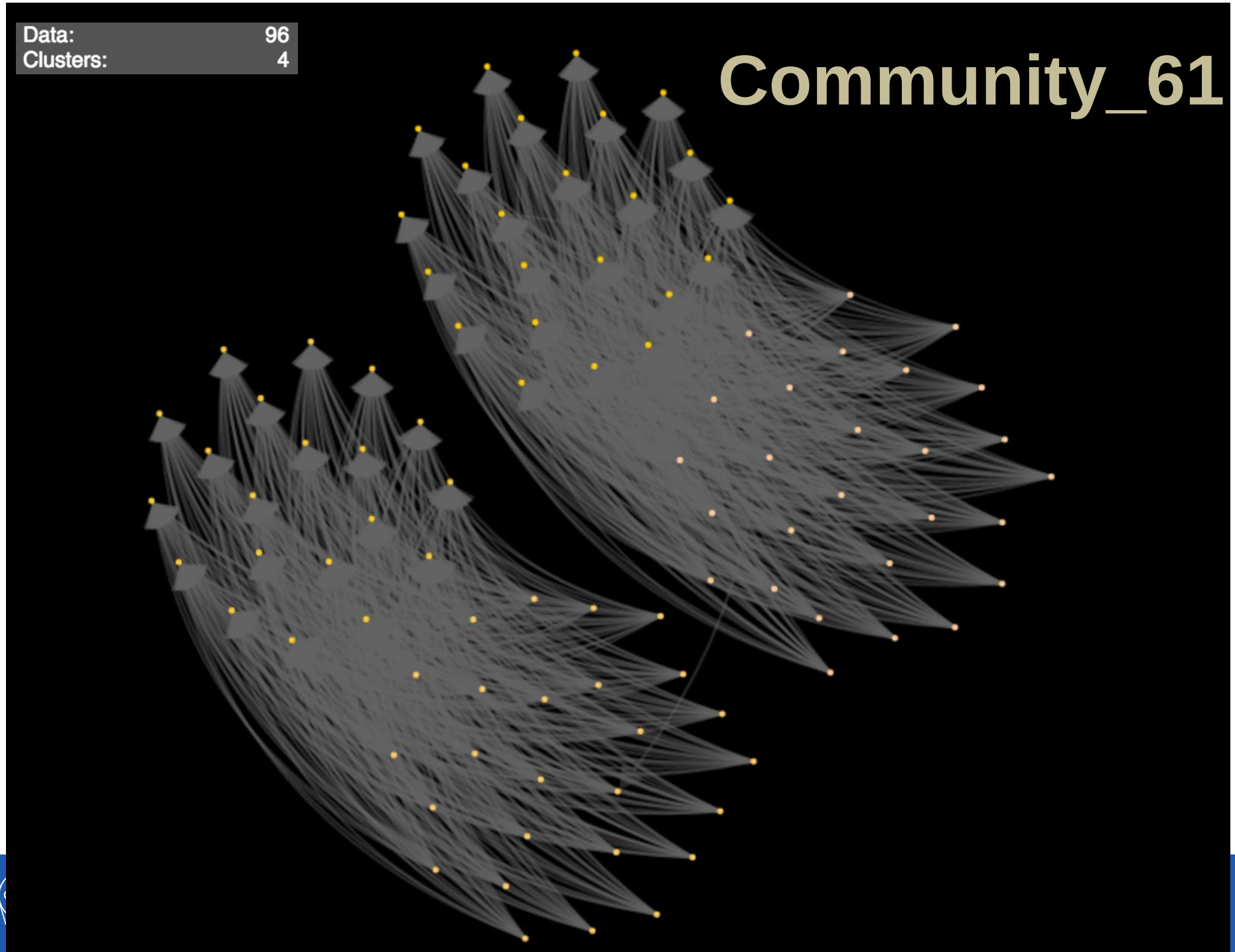
g_0 graph


RI
INS
5
7b
1
SII
4
2
3b
6/SMA
3a
VIP
TF (TEpv?)
46
V4t
DP
PIP
MSTd/p
7a
TEO
STPa
V1
TH-TF
PERI
TEpv
V3
TEpd
PO
V4
TEam-a/dP
TEav
TEad
V2
V3A
MT
VP
FEF
MSTI
STPp
FST

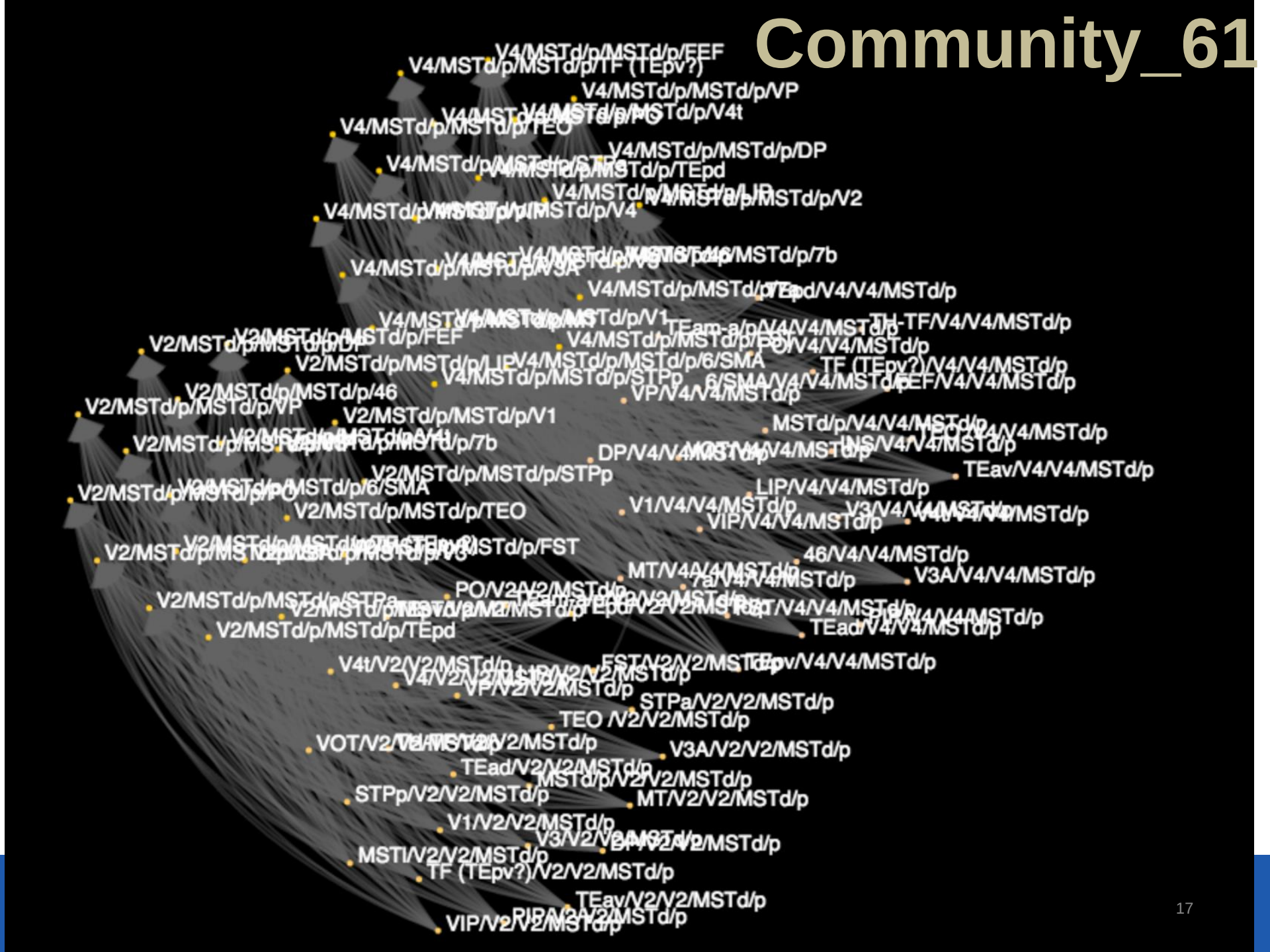
g_2 (with Quotient graph)

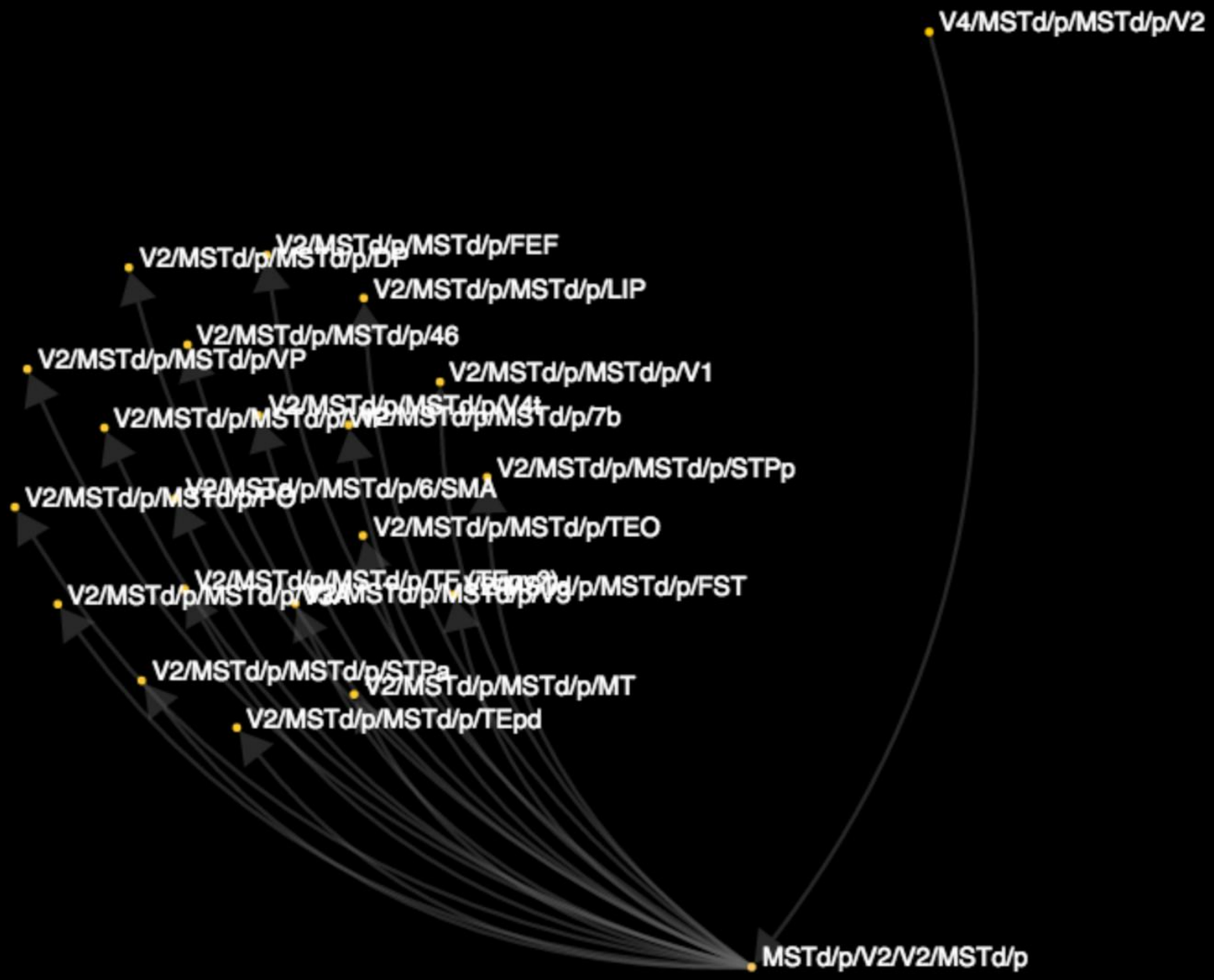


Community_61



Community_61




V4/MSTd/p/MSTd/p/V2
V2/MSTd/p/MSTd/p/FEF
V2/MSTd/p/MSTd/p/DP
V2/MSTd/p/MSTd/p/LIP
V2/MSTd/p/MSTd/p/46
V2/MSTd/p/MSTd/p/VP
V2/MSTd/p/MSTd/p/V1
V2/MSTd/p/MSTd/p/7b
V2/MSTd/p/MSTd/p/STPp
V2/MSTd/p/MSTd/p/6/SMA
V2/MSTd/p/MSTd/p/TEO
V2/MSTd/p/MSTd/p/FST
V2/MSTd/p/MSTd/p/STPa
V2/MSTd/p/MSTd/p/MT
V2/MSTd/p/MSTd/p/TEpd
MSTd/p/V2/V2/MSTd/p

Conclusion / Outlook

- Visual Analytics both inherits from and complements Machine Learning

- Neural Architectures are flexible tools to learn non-linear processes
- Their integration in Learning processes can be diverse

- The parallel with understanding neurological processes may still have a lot to offer