Title: Scale-spaces from protein networks: how diffusion profiles reveal functional information in physical interaction topologies

Authors: Christos Dimitrakopoulos[1,*], Marc Hulsman[2,*], Jeroen de Ridder[2,#]

1) University of Patras, Computer Engineering and Informatics Department, University of Patras campus, Patras, Greece
2) Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands
*) Contributed equally
#) Corresponding author

Understanding the living cell as a system of interconnected components is one of the key contemporary challenges. This is a complex problem, in which different types of functional interactions play a role, each operating across multiple distinct scales. At the lowest scale, proteins physically interact with each other and with the DNA. At higher scales, more indirect interactions appear, such as co-expression, signaling cascades or genetic interaction. These arise as a result of the physical network topology.

This raises the question: how do these functional interactions emerge from the topology of the physical interaction network and how does the answer to this question contribute to a systems level understanding of a cell? To investigate this we explore *scale independent* descriptions of the topology of the physical protein-protein and protein-DNA interaction networks in yeast.

We address this question with graph diffusion, a method that we employ to construct scale-space descriptions of network structure. These descriptions characterize the topology connecting two proteins across all scales simultaneously. We call these descriptions *diffusion profiles*.

We explore the use of diffusion profiles in a predictive setting. Based on the physical interaction topology, characterized by the diffusion profiles, we show that we can reliably distinguish the three types of functional interaction from random interactions, and, interestingly, also from each other. Importantly, this requires a non-linear combination of topological features at multiple scales, demonstrating the usefulness of a scale-space approach.

Taken together, diffusion profiles clearly capture biologically meaningful features by exploiting, across multiple scales, graph topology in readily available physical interaction data.