Workshop

Computational Methods for Networks

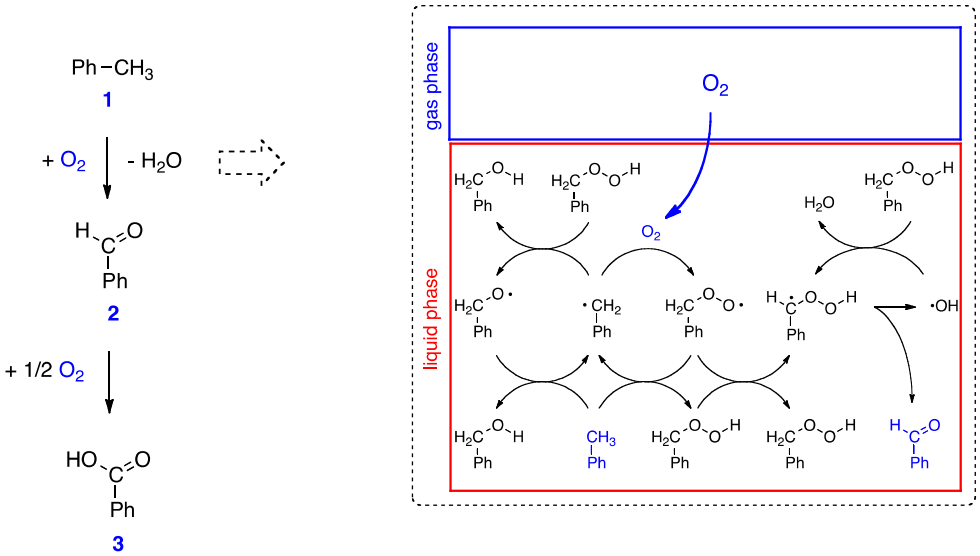
9 and 10 November 2015

Abstracts

Hendrik Zipse

Hydrocarbon (Aut)Oxidation –Reaction Networks in Seemingly Trivial Molecular Transformations

The (aut)oxidation of hydrocarbons plays a central role in a larger number of industrially important processes. Taking the oxidation of toluene (**1**) to benzaldehyde (**2**) and on to benzoic acid (**3**) as an example, the overall reaction can be expressed as a simple two-step sequence (Scheme 1). At microscopic level, however, already the first step involves a complex interplay of individual steps, many of which depend on each other. Moreover, progress of the reaction in the liquid phase is coupled to the rate of diffusive oxygen transport across the gas/liquid interface.



Scheme 1. Radical chain oxidation of toluene (**1**) to benzoic acid (**3**) at the gas/liquid interface.

Numerical simulations of this type of reaction (and thus the prediction of the reaction outcome as a function of experimental parameters) require accurate data for the temperature- and pressure-dependence of the rate of each of the microscopic steps. Little data of this type is available from experiment, and highly accurate quantum chemical calculations are therefore used to calculate these kinetic parameters directly.

Thomas Seidl

Big Data Stream Analytics

In our days, huge and still increasing amounts of data are collected from scientific experiments, sensor and communication networks, business processes and many other domains. Database and data mining techniques aim at efficiently handling these large and complex data to support daily transactions as well as strategic decision making based on the discovery of regular or irregular patterns hidden in the data. Current research trends in data analytics are driven by the high volume, velocity, and variety of Big Data.

The talk discusses some challenges in the field. In recent developments for dynamic stream data mining, anytime algorithms play an important role. Novel hierarchical, statistical indexing structures including BayesTree and ClusTree allow for obtaining high quality results "at any time" while adapting themselves to varying stream velocities. Particular challenges occur when supervised and unsupervised mining tasks are faced with multimodal streams of complex multimedia objects.

Paul Thurner

The Network of Arms Transfers 1950-2013: An Application of ERGMs and TERGMs

(Authors: Paul W. Thurner, Skyler Cranmer, Goeran Kauermann, Christian Schmid)

Which factors determine the trade on conventional arms between countries? The international trade of arms constitutes a network of exporting and of importing countries exhibiting highly complex structures and dynamics. Unfortunately, there are only very few studies investigating these relations quantitatively and in a long-term perspective. Focusing on the trade on major conventional weapons (MCW) we provide the first statistical network analysis on this subject. Based on SIPRI's exhaustive data set we apply cross-sectional and temporal exponential random graph models (ERGMS & TERGMS) to binarized networks. Our analyses demonstrate that both, economic as well as security consideration are relevant factors with some of them being time-specific. Thus, military alliances are losing their imprint at least in the period between 1995-2005, and intrastate conflict continuously attracts weapons imports since 1980. Politico-military considerations seem to regain importance in the last years, thus indicating the emergence of a new international security regime. In order to assess the predictive performance of our model we rely on out-of-sample forecasting.

Mike West

Dynamic Sparsity Modelling

Bayesian model developments in a range of approaches to sparsity modelling in multivariate time series have recently been demonstrably valuable in a number of topical application areas. Practical deployment of models and coupled, relevant decision analytic strategies, have defined advances in evaluating time-varying inter-relationships among series, dynamic dependency networks (with time-evolving predictor-response interconnections and lagged structures), and applications in prediction and decisions across a range of financial, economic and scientific areas.

I will review some of these recent innovations for dynamic sparsity modelling using the concept – and resulting methodology – of dynamic latent thresholding. The latent thresholding concept and methodology defines a framework for sparsity in dynamic regressions, dynamic latent factor models, time-varying vector autoregressions, and dynamic graphical models of multivariate stochastic volatility; this opens a path to new computational approaches for identifying and exploring empirical representations of complicated dynamic networks. Several studies in finance and econometrics, and from clinical neuro-psychiatry, illustrate the approach.

Nils Detering

Bootstrap Percolation in Directed, Inhomogeneous Random Graphs and Financial Contagion

We propose a directed version of the Chung-Lu random graph (cf. (Chung and Lu 2002)). Each realized graph is simple even if the sequences of expected degrees have unbounded second moment. The model thus offers a more widely applicable and analytical tractable alternative to the directed configuration model. The degree sequence is shown to be close to a multivariate mixed Poisson distribution, which offers enough flexibility to represent many observed real-world networks. In a second step we assign to each knot an individual threshold value, which determines its vulnerability to infection. A bootstrap percolation process, triggered by some initial infections and spread through the graph via its edges, is analyzed and we derive formulas for the asymptotic fraction of infected vertices. Furthermore, we show when the final fraction is bounded away from zero, independent of the size of the fraction of initially infected vertices. The motivation for our work comes from financial mathematics and we discuss some of the implications for default contagion in financial networks.

Sascha Rothe

Similarity Measures on Wordgraph

Similarity Measures on Wordgraph can be used to compute synonyms or translations. We present CoSimRank, a graph-theoretic similarity measure and show CoSimRank's close relationship to Personalized PageRank and SimRank. While the new similarity measure can be computed for a single node pair without relying on the similarities in the whole graph we also show how to take advantage of fast matrix multiplication algorithms to compute CoSimRank.

Ziga Avsec

Application of Protein Interaction Network in Causal Gene Prioritization

Identification of novel causal genes in rare Mendelian diseases helps to understand the disease's pathomechanism, solve more cases and reveal novel biological pathways. For proving the causality of a gene, experimental follow-up studies are needed. These are usually expensive and time-consuming, hence the genes sent for validation should be chosen carefully. Here, we propose different statistical models for prioritizing candidate genes that integrate the information from exome sequencing data, known causal genes, molecular pathways and protein-protein interaction networks. The models are applied to a whole-exome sequencing dataset of about 400 mostly unrelated cases diagnosed with a mitochondrial disease and 125 controls. We show how protein-interaction networks can be encoded as prior information in a Bayesian setting and demonstrate the efficiency of these approaches using a benchmark with 71 well established known causal genes. Application of Random Forest for combining the results of the developed methods is presented, confirming that a protein interaction network is an essential part of a gene prioritization pipeline.

Paul Embrechts

Model Risk, Solvency, and Risk Aggregation

Under both Basel II/III for banking as well as Solvency 2/SST for insurance, Model Risk (MR), especially for Risk Aggregation purposes, plays an important role. In this talk I will concentrate on Dependence Uncertainty and quantify MR from that point of view. Besides reviewing some of the main results obtained over the recent years, I will discuss several examples coming from the realm of Operational Risk, as well as the calculation of Economic Capital in a real banking example. A basic reference is A.J. McNeil, R. Frey, P. Embrechts (2015) Quantitative Risk Management: Concepts, Techniques and Tools. Revised Edition, Princeton University Press.

Ulrik Brandes, University of Konstanz

Measuring a Node's Position

Centrality indices are frequently referred to as measures of structural importance. We will discuss why the allusion to measurement is problematic and propose an alternative. The new approach to centrality is inspired by conjoint measurement and based on a common generalization of previous notions of position together with a novel concept of positional dominance. It has far-reaching implications for network analysis at large, some of which will be hinted at.

Christian Ochsenfeld

Schrödinger's World – Quantum Chemistry for Describing Molecular Systems

The Schrödinger equation provides the exact non-relativistic description of molecular matter, however, the exact solution has been a key-challenge since its introduction in 1926. Therefore a multitude of approximations have been introduced that allow to systematically approach the exact solution. This hierarchy of approximations allows in principle to systematically check the reliability of computational predictions. In this way, quantum chemistry has evolved over the last decades to become a versatile tool for studying structures and properties of molecular systems. Despite this success, the applicability to large molecules is hampered by the strong polynomial increase of the computational effort with molecular size M. Therefore the central goal of our work is to overcome this scaling wall and to develop linear- or for specific properties even sublinear-scaling methods for describing large and complex molecular systems.

The talk will provide an overview of quantum-chemical methods, new possibilities for describing complex systems, and examples of studying for example biochemical processes such as DNA repair. Furthermore, recent steps for exploiting highly parallel computing platforms such as graphic-processing units (GPUs) for quantum-chemical calculations are outlined.

Claudia Klüppelberg

Risk in a Bipartite Graph Structure

We introduce a random network model for business relationships exemplified for a re-insurance market. Using Pareto-tailed losses (as are observed for natural or man-made catastrophes) with a dependence structure introduced by the graph we study systemic risk measures, which are based on the Value-at-Risk and the Expected Shortfall. We show that the dependence on the network structure plays a fundamental role for the individual agent's risk as well as for the systemic risk. If the Pareto exponent is larger than 1, then for the individual agent diversification is beneficial, whereas when it is less than one, concentration on a few objects is the better strategy for individual agents. The situation changes, however, when systemic risk comes into play. We describe different network scenarios including a homogeneous model and a Rasch-type model, and explain the influence of the network structure on diversification in such models. This is joint work with Oliver Kley and Gesine Reinert.