Probabilistic Structures in Evolution
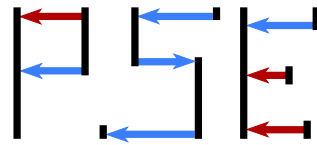
September 19–21, 2012

TU Berlin, Institut für Mathematik

Program

	September 19th
9:00–10:00	Registration
10:00–10:10	Welcome
10:10–11:00	AMANDINE VEBER Evolution in a spatial continuum
11:00–11:30	JOACHIM KRUG Evolutionary accessibility and genetic architecture
11:30–12:00	Break
12:00–12:30	SHIDONG WANG Multi-scale analysis of a locally regulated population
12:30–13:10	DIRK METZLER AND MEIKE WITTMANN Interacting populations and their genetic structure – stochastic models at the interface between ecology and evolution
13:10–15:00	Lunch break
15:00–16:00	YUN SONG Explicit transition density expansions of diffusion processes with general diploid selection and estimation of selection parameters from temporal DNA samples
16:00–16:30	ACHIM KLENKE Infinite rate mutually catalytic branching
16:30–17:00	Break
17:00–17:30	MICHAL KOMOROWSKI How cells do statistics? Inference, experimental design, signal transduction and evolution.
17:30–18:00	ANJA STURM On spatial coalescents with multiple mergers in two dimensions
18:00–19:00	GET-TOGETHER RECEPTION

	September 20th
10:00–11:00	TIBOR ANTAL Stochastic models of tumor progression
11:00–11:30	ANTON KLIMOWSKY Renormalization of hierarchically interacting Λ-Cannings processes
11:30–12:00	Break
12:00–12:30	BJARKI ELDON An ancestral recombination graph for diploid populations with skewed offspring distribution
12:30–13:10	MICHAEL BAAKE AND MAREIKE ESSER Single-crossover recombination and ancestral recombination trees
13:10–15:00	Lunch break
15:00–15:30	LUCA FERRETTI Height and shape of the spatial coalescent
15:30–16:00	GÖTZ KERSTING Beta-coalescent trees and their lengths
16:00–16:40	THOMAS WIEHE AND FILIPPO DISANTO Some combinatorial statistics on trees arising from population genetic problems
16:40–18:00	Break
18:00–19:00	JOHN WAKELEY Coalescent Theory: The Past and the Future
19:00–	RECEPTION AT THE MATHEMATICAL LIBRARY

	September 21st
10:00–10:30	MATTHIAS BIRKNER Ancestral lineages and local regulation
10:30–11:00	CLEMENT FOUCART M-Coalescents and Generalized Fleming–Viot processes with immigration
11:00–11:30	NINA GANTERT Conditioning Galton–Watson trees to have small martingale limits
11:30–12:00	Break
12:00–13:00	JEAN BERTOIN The cut-tree of large Galton–Watson trees and the Brownian CRT
13:00–13:05	CLOSING

Stochastic models of tumor progression

Tibor Antal

Edinburgh University

Stochasticity is essential when modeling initiation of tumors, progression of tumors from benign to malignant states, or metastasis formation. Many aspects of these phenomena can be modeled by simple multi-type branching processes, and the results compare fairly well with experimental and clinical data. Other aspects of tumor development needs more refined approaches. Models of developing resistance to chemotherapy, and modeling challenges related to spatial structure and sizes of tumors will also be discussed.

Single-crossover recombination and ancestral recombination trees

Micheal Baake and Mareike Esser

Universität Bielefeld

Determining explicit solutions to the dynamics of the genetic composition of a population evolving under the evolutionary force of recombination has been a challenge to theoretical population genetics for a long time. In continuous time, restricted to the special case of single-crossover recombination, the resulting system of non-linear differential equations could be solved in closed form. Inspired by this, we also investigate the stochastic counterpart. To this end, we trace back the history of a single individual from the present population, based on the Wright-Fisher model. Under the assumption of strong recombination, i.e. considering the process in the infinite population limit without rescaling parameters or time, the ancestry of the given individual can be described by a random binary tree which we call *ancestral recombination tree*. By decomposing these trees into subtrees, a closed-form expression for the probabilities of the ancestral trees could be found. Combining these results with the deterministic ones, we obtain an independent approach to solve the deterministic single-crossover equation in continuous time.

The cut-tree of large Galton–Watson trees and the Brownian CRT

Jean Bertoin

Universität Zürich

The talk is based on a joint work with Grégory Miermont (Orsay). Consider the edge-deletion process in which the edges of some finite tree are removed one after the other in the uniform random order. Roughly speaking, the cut-tree then describes the genealogy of connected components appearing in this edge-deletion

process. Our main result shows that after a proper rescaling, the cut-tree of a critical Galton–Watson tree with finite variance and conditioned to have size n, converges as $n \to \infty$ to a Brownian CRT in the weak sense induced by the Gromov-Prokhorov topology. This yields a multi-dimensional extension of a limit theorem due to Janson for the number of random cuts needed to isolate the root in Galton–Watson trees conditioned by their sizes.

Ancestral lineages and local regulation

Matthias Birkner

Universität Mainz

The spatial embeddings of genealogies in models with fluctuating population sizes and local regulation are relatively complicated random walks in a space-time dependent random environment. We use the supercritical discrete-time contact process as a simple non-trivial example of a locally regulated population model and study the dynamics of ancestral lineages sampled at stationarity, viz. directed random walk on a supercritical directed percolation cluster. We prove a LLN and a quenched CLT for such walks via a regeneration approach. Furthermore, we discuss possible extensions to more general models that allow multiple occupancy of sites.

Based on joint work with Jiří Černý, Andrej Depperschmidt and Nina Gantert.

Height and shape of the spatial coalescent

Luca Ferretti

Universitat Autonoma de Barcelona

Single recombination events are one of the major forces shaping coalescent trees along the genome, particularly important near selective sweeps. However, our understanding of their effect on the trees is still partial. We focus on drastic recombination events, i.e. events that change important features of the tree like the total height h and the size of the root branches Omega. We discuss the distribution on the coalescent trees along the genome and provide analytical results on the probability of height-changing and Omega-changing single recombination events in the standard neutral model. These results can also be used to estimate correlation lengths for tree statistics.

Joint work with Filippo Disanto and Thomas Wiehe.

M-Coalescents and Generalized Fleming–Viot processes with immigration

CLÉMENT FOUCART

Laboratoire de Probabilités et Modèles Aléatoires, Paris

The purpose of the talk is to introduce a new class of processes arising as the genealogy of some population models with immigration. In the same spirit as Bertoin and Le Gall, we will study a stochastic flow of bridges which encodes a population with immigration backward in time (leading to a process called M-coalescent) and forward in time (leading to a probability measure valued process, called M-generalized Fleming–Viot process). According to the time left, I would explain some results linking certain of these new processes with the stable branching processes with immigration.

Conditioning Galton–Watson trees to have small martingale limits

NINA GANTERT

TU München

We show that an infinite Galton–Watson tree, conditioned on its martingale limit being smaller than epsilon, agrees up to generation K with a regular μ-ary tree, where μ is the essential minimum of the offspring distribution and the random variable K is strongly concentrated near an explicit deterministic function growing like a multiple of $\log(1/\varepsilon)$. More precisely, we show that if $\mu \gg 1$, then with high probability as ε goes to 0, K takes exactly one or two values. This shows in particular that the conditioned trees converge to the regular μ-ary tree, providing an example of entropic repulsion where the limit has vanishing entropy.

The talk is based on joint work with Nathanael Berestycki, Nadia Sidorova and Peter Mörters.

Beta-coalescent trees and their lengths

GÖTZ KERSTING

Universität Frankfurt

n-coalescent trees are used to describe random genealogies belonging to n particles. If one considers n individuals within a large population, then usually the Kingman coalescent is used, whereas the Bolthausen–Sznitman coalescent arose from a physical context. Both coalescents belong to the larger (parametrized) class of Beta-coalescent trees. We discuss geometric properties of these trees, in particular their total lengths and total external lengths (both quantities having a clear biological meaning). A specific feature is a kind of phase transition, appearing within the range between the Kingman and Bolthausen–Sznitman coalescent.

Infinite rate mutually catalytic branching

ACHIM KLENKE

Universität Mainz

The mutually catalytic branching process is an example of a two-type branching process with self-regulation. We consider the limit of this model as the branching rate (constant of proportionality in the branching term) tends to infinity.

The aim of this talk is to survey what is known and outline possible future developments.

An ancestral recombination graph for diploid populations with skewed offspring distribution

BJARKI ELDON

University of Oxford

A large offspring number diploid biparental multilocus population model of Moran type is our object of study. At each timestep, a pair of diploid individuals drawn uniformly at random contribute offspring to the population. The number of offspring can be large relative to the total population size. Such 'heavily skewed' reproduction mechanisms have been considered by various authors recently, cf. e.g. Eldon and Wakeley (2008), and reviewed by Hedgecock and Pudovkin (2011). Each diploid parental individual contributes exactly one chromosome to the diploid offspring, and hence ancestral lineages can only coalesce when in distinct individuals. We observe a separation of timescales phenomenon; a result of Möhle (1998) is extended to obtain convergence of the ancestral process to an ancestral recombination graph necessarily admitting simultaneous multiple mergers of ancestral lineages.

Due to diploidy and large offspring numbers, novel effects appear. For example, the marginal genealogy at each locus admits simultaneous multiple mergers in up to four groups, and different loci remain substantially correlated even as the recombination rate grows large. Thus, genealogies for loci far apart on the same chromosome remain correlated. Correlation in coalescence times for two loci is derived and shown to be a function of the coalescence parameters of our model. Correlations in ratios of coalescence times between loci can be high, even when the recombination rate is high, in large offspring number populations, as suggested by simulations, hinting at how to distinguish between different population models. Correlations in F_{st} and heterozygosity across loci also results from a skewed offspring distribution mechanism.

Renormalization of hierarchically interacting Λ-Cannings processes

Anton Klimovsky

Leiden University

Consider a large system of individuals of multiple types, who live in spatially distributed colonies. Let the individuals move randomly around between the colonies (= migration). Assume additionally that the individuals are subject to occasional stochastic non-local reshuffling-resampling events with skewed offspring distribution (= reproduction under constraint amount of resources). What can we say about the ergodic behaviour of such systems? How does it depend on the intensities of the evolutionary forces? Under which circumstances shall we expect in the long run local coexistence of individuals of different types in the colonies? Under which circumstances do mono-type clusters of colonies appear? We suggest a class of stochastic models of such systems (based on the Cannings models from population genetics) and obtain a clear-cut criterion for the clustering vs. local coexistence dichotomy. This is joint work with Frank den Hollander, Andreas Greven, and Sandra Kliem.

How cells do statistics? Inference, experimental design, signal transduction and evolution.

Michal Komorowski

Institute of Fundamental Technological Research, Polish Academy of Sciences

We employed statistical tools to take a comprehensive view at stochastic biochemical models. Living cells are forced to solve inference problems as they measure constantly changing parameters of their environment. Their biochemical machineries encode certain experimental procedures which, at least to some extent, must have been evolutionary optimised. Statistical methodology of optimal experimental design, therefore, can provide insight about information processing and evolutionary adaptation of signal transduction networks. Having developed a method to calculate an optima design criterium (Fisher Information) for stochastic models of biochemical reactions we analysed the nitrogen assimilation circuit in *Escherichia Coli* to reveal its design principles.

Evolutionary accessibility and genetic architecture

JOACHIM KRUG

Universität zu Köln

The adaptive dynamics of a population in the space of genotypes is constrained by epistatic interactions between mutations at different genetic loci. Recent empirical studies have shown that this strongly reduces the number of evolutionary trajectories that are accessible under the common conditions of weak mutation and strong selection. After a brief overview of the experimental situation, I will present new results on evolutionary accessibility in Kauffman's NK-model, in which each of N loci interacts with a 'neighborhood' of K other loci in a random way. Different neighborhood choices are equivalent with respect to the fitness correlations and the Fourier decomposition of the landscape, but differ markedly in their accessibility properties. In particular, for the 'block' version of the NK-model introduced by Perelson and Macken, the distribution of the number of accessible paths can be explicitly computed.

Interacting populations and their genetic structure – stochastic models at the interface between ecology and evolution

DIRK METZLER AND MEIKE WITTMANN

LMU Munich

That ecological and evolutionary processes take place on different time scales is a widespread simplifying assumption in population biology. There are phenomena, however, that can only be understood by considering the direct interplay between the ecology of species interactions and the evolution or genetic structure of the interacting populations. Here we present stochastic models for two such cases.

First, we consider kin selection for parasite defense traits in spatially structured populations. To examine whether this can explain the evolution of slave rebellion in a system of two ant species (*Temnothorax longispinosus* and the slavemaking parasite *Protomagnatus americanus*) we analyze genetic data following an approximate Bayesian approach that allows us to account for detailed biological background information.

Second, we analyze for how long a population can persist in the face of repeated introductions of a new competing species. We address this question for two continuous-time Markov models: an ecological model and an eco-genetic model, which includes the genetics of the resident population at a locus involved in its adaptation to a changing environment. For both models, we compute the expected persistence time and its dependence on competition intensity. A comparison between the two models reveals that the introduced species' ecological and genetic effects engage in a feedback which can accelerate the extinction process.

Explicit transition density expansions of diffusion processes with general diploid selection and estimation of selection parameters from temporal DNA samples

Yun Song

UC Berkeley

The transition density function of the Wright-Fisher diffusion describes the evolution of population-wide allele frequencies over time. This function has important practical applications in population genetics, but finding an explicit formula under a general diploid selection model has remained a difficult problem. In this talk, I will describe a new computational method based on orthogonal functions to tackle this classic problem. Specifically, our method explicitly finds the eigenvalues and eigenfunctions of the diffusion generator associated with the Wright-Fisher diffusion with recurrent mutation and arbitrary diploid selection, thus allowing one to obtain an accurate spectral representation of the transition density function. As a byproduct of our work, we obtain the rate of convergence to the stationary distribution under mutation-selection balance. I will also describe an application of our transition density expansion in an HMM framework to estimate selection parameters from temporal sample allele frequency data.

Joint work with Matthias Steinrücken, Anand Bhaskar, and Rachel Wang.

On spatial coalescents with multiple mergers in two dimensions

Anja Sturm

Universität Göttingen

In this talk we consider the genealogy of a sample of individuals taken from a spatially structured population when the variance in the number of each individual's offspring is relatively large. The space is structured into discrete sites of a graph G. If the population size at each site is large, spatial coalescents with multiple mergers, for which ancestral lines migrate in space and coalesce according to some Lambda-coalescent mechanism, are shown to be appropriate approximations to the genealogy of a sample of individuals.

We then consider as the graph G the two dimensional torus with side length L and show that as L tends to infinity, and time is rescaled appropriately, the partition structure of such a spatial Lambda-coalescent of individuals sampled far enough apart converges to the partition structure of a non-spatial Kingman coalescent. From a biological point of view this means that in certain circumstances both the spatial structure as well as larger variances of the underlying offspring distribution are hard to detect from the sample.

This is joint work with Benjamin Heuer (Universität Göttingen).

Evolution in a spatial continuum

Amandine Veber

École Polytechnique, Paris

The spatial Lambda-Fleming–Viot process models the evolution of the genetic composition of a population spread over some continuous space. In this talk, we shall review some recent results and open questions on this process. In particular, we shall focus on the corresponding genealogical process, when the gene of interest is subject to natural selection.

Multi-scale analysis of a locally regulated population

Shidong Wang

University of Oxford

We consider a fitness-structured population model with competition and migration between nearest neighbors. Under constraints of relative strength between migration and mutation, we study the long time behavior of the total population partition on support trait space. For the population without mutation on a finite trait space we obtain the equilibrium configuration and characterize the right time scale for fixation. For the model with mutation on an infinite trait space, a measure-valued jump process, which we call trait substitution tree (TST) model, is established on the rarer mutation time scale against the rare migration one. The novelty here is that every phenotype, which may nearly die out on the migration time scale, has a chance to recover and further to be stabilized on the mutation time scale because of a change of fitness landscape due to a new-entering mutant. The genetic motivation will be mentioned.

Some combinatorial statistics on trees arising from population genetic problems

Thomas Wiehe, Filippo Disanto

Universität zu Köln

Binary rooted trees of size n are enumerated by the Wedderburn–Etherington numbers. This sequence can be computed exactly only recursively. Using generating functions techniques we are able to re-derive an approximation for this sequence and we generalize it in a such a way that we can compute the probability distribution of maximal un-balanced subtrees, where balance is measured by Colless index.

Unbalanced trees and subtrees are of high interest in population genetic modeling and data analysis, because positive selection induces distortions of tree topology such that unbalanced (sub-)trees become highly over-represented compared to neutral expectations.