

# MATRI

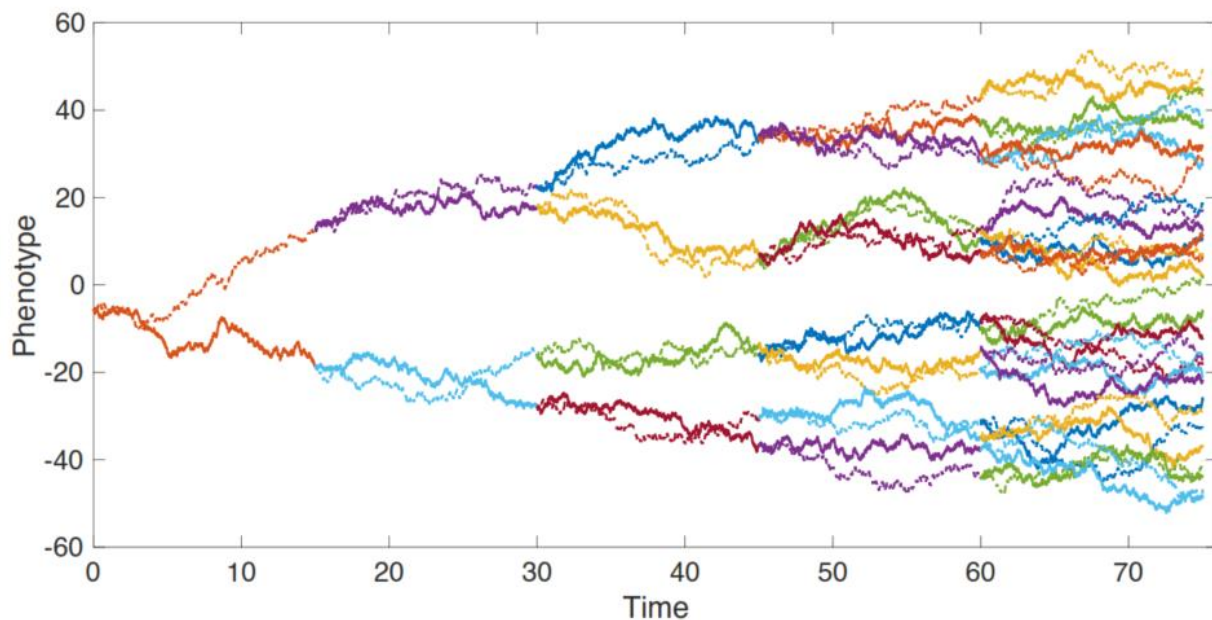


MATRIX RESEARCH PROGRAM

# Probabilistic Models in Evolutionary Biology and Game Theory

6 – 10 January 2025

CRESWICK, AUSTRALIA



<https://www.matrix-inst.org.au/>



<https://bsky.app/profile/matrix-inst.bsky.social>



<https://www.linkedin.com/in/matrix-australia>

## 1 Tentative Program

Time	Monday	Tuesday	Wednesday	Thursday	Friday
8-9am	Breakfast	Breakfast	Breakfast	Breakfast	Breakfast
9-11am	Opening & 5-min. introductions*	<b>Session 2</b> Three 40-min. talks	<b>Session 3</b> Three 40-min. talks	<b>Session 4</b> Three 40-min. talks	<b>Session 4 (cont.)</b> One 40-min. online talk  Final collaboration & discussion session
11am-12.30pm	Morning tea & Discussion	Morning tea & Discussion	Morning tea & Discussion	Morning tea & Discussion	Morning tea & Closing
12.30-2pm	Lunch	Lunch	Lunch	Lunch	Lunch (not provided)
2-5pm	<b>Session 1</b> Three 40-min. talks & Discussion	Collaboration & discussion session.  <i>Wine and cheese</i>	Collaboration & discussion session.  <i>Excursion</i>	Collaboration & discussion session  <i>Excursion</i>	Departure
7-8pm	Dinner	Dinner	Dinner	Dinner	

2

\*Each participant gives a brief 5-minute overview of themselves and their research

## 2 Talks sessions

Talks are 40 minutes long, including questions.

Day	Session Title	Speakers and Talks
Monday	<b>Session 1:</b> Stochastic Models in Population and Evolutionary Dynamics	<b>Simon Harris:</b> Genealogies of samples from stochastic populations <b>Robert Charles Griffiths:</b> Coalescence in Feller Branching Diffusions <b>Jie Yen Fan:</b> Measure-valued population processes
Tuesday	<b>Session 2:</b> Branching Processes with Competition and Statistical Inference	<b>James Kerlidis:</b> Linking Population-Size-Dependent and Controlled Branching Processes <b>Fima Klebaner:</b> Emergence of populations: $H(W)$ theory <b>Jesse Goodman:</b> Saddlepoint approximations for likelihoods
Wednesday	<b>Session 3:</b> Evolutionary Biology and Genetic Models	<b>Malgorzata O'Reilly:</b> Matrix-analytic methods for the evolution of species trees, gene trees, and their reconciliation <b>Nantawat Udomchatpitak:</b> The Accumulation of Beneficial Mutations and Convergence to a Poisson Process <b>Watthanan Jatuviriyapornchai:</b> Evolutionary Biology through the Inclusion Process
Thursday	<b>Session 4:</b> Game Theory and Decision-Making	<b>Rui Zhang:</b> Finding Nash equilibria in random win-loss games <b>Hugh Entwistle:</b> On generalisations of the secretary problem <b>Peter Taylor:</b> A continuous-time Markov chain model for the distribution of blockchain heights
Friday	<b>Session 4 (cont.)</b>	<b>Krzysztof Szajowski (online):</b> Rationality, Deviation, and Diagnosis: A Holistic Approach to Stopping Games

### 3 Abstracts

Ordered alphabetically by surname.

#### Hugh Entwistle

**Affiliation:** Macquarie University

**Talk Title:** *On generalisations of the secretary problem*

**Abstract:** The classic “secretary problem,” where one wishes to maximise their probability of selecting the best candidate with only one choice when observing a random permutation of  $N$  secretaries sequentially, is a well-studied problem. In this talk, we consider different variants of this problem, including when the number of objects is random, and when it is possible to make several choices.

#### Jie Yen Fan

**Affiliation:** Monash University

**Talk Title:** *Measure-valued population processes*

**Abstract:** Measure-valued stochastic processes are very useful in describing the evolution of complex structural systems, such as in population modeling. We can use measure-valued processes to study population structures that feature characteristics such as age, gender, and other attributes; and where the dynamics depend on the individual state as well as the overall structure of the population. Limiting behaviors in terms of the Law of Large Numbers and the Central Limit Theorem can be obtained in the form of PDE and SPDE, respectively. Some associated results will also be discussed.

## Jesse Goodman

**Affiliation:** University of Auckland

**Talk Title:** *Saddlepoint approximations for likelihoods*

**Abstract:** The saddlepoint approximation is a systematic method for converting a known generating function into an approximation for an unknown density function. Interpreted instead as an approximation to the unknown likelihood function, the saddlepoint approximation can be maximized to compute the saddlepoint MLE for a given observed value. This talk will explain how the saddlepoint approximation can be interpreted with a statistical lens and describe a class of models with theoretical guarantees for the effect of using the saddlepoint MLE as a substitute for the unknown true MLE. The talk will also demonstrate new tools to simplify and automate the computation of saddlepoint MLEs and to quantitatively assess the amount of approximation error. Based on joint work with Godrick Oketch and Rachel Fewster.

## Robert Charles Griffiths

**Affiliation:** Monash University

**Talk Title:** *Coalescence in Feller Branching Diffusions*

**Abstract:** A Feller branching diffusion  $X(t)$  has infinitesimal mean  $\alpha x$  and variance  $x$ . The process is subcritical if  $\alpha < 0$ , critical if  $\alpha = 0$ , and supercritical if  $\alpha > 0$ . It can be thought of as a limit from a Bienaymé-Galton-Watson process. This talk will discuss coalescence in the Feller diffusion. Coalescence is described by  $A_n(s, t)$ , the number of ancestors of a sample of  $n$  individuals taken at time  $t$ , at an earlier time  $t - s$ . The number of ancestors of the population  $A_\infty(s, t)$  has a distribution which is compound Poisson-Geometric.

The limit coalescent distribution as  $t \rightarrow \infty$ , conditional on non-extinction at  $t$ , in the subcritical case, has the property that the mean time between coalescent events is the same as in the Kingman coalescent, but the distribution is different. Another limit is to take  $s \rightarrow 0$  in a supercritical model. The asymptotics of the coalescent process, looking at  $s$  back to zero, are related to the reduced tree of a non-homogeneous birth and death process.

## Simon Harris

**Affiliation:** University of Auckland

**Talk Title:** *Genealogies of samples from stochastic populations*

**Abstract:** What does the family tree look like for a random sample of  $k$  individuals taken from some population? Surprisingly, until relatively recently this fundamental question remained an open problem even for one of the simplest of stochastic population models. We will discuss some recent research in this area, including the emergence of certain universal limiting genealogies when sampling individuals at random from large stochastically evolving populations, such as critical Galton-Watson stochastic branching processes conditioned to survive.

## Watthanan Jaturiyapornchai

**Affiliation:** Mahidol University

**Talk Title:** *Evolutionary Biology through the Inclusion Process*

**Abstract:** The inclusion process is a continuous-time stochastic particle system where particles perform independent random walks on a lattice and, in addition, interact via an inclusion mechanism, with rates proportional to the product of occupation numbers at departure and arrival sites. It can also be interpreted as a multi-allele version of the Moran model with mutation, describing generic dynamics in finite populations, where sites are interpreted as allele types and particles as individuals. A key phenomenon in the inclusion process is condensation, where stationary states typically concentrate on a single site, corresponding to species fixation as mutation rates vanish. In this talk, we present a complete picture of condensation in inclusion processes in the thermodynamic limit, covering all scaling regimes of the diffusion parameter. Using this approach, we study the structure of the condensed phase and hierarchical organization characterized by the Poisson-Dirichlet distribution. This is joint work with Stefan Grosskinsky and Paul Chleboun.

## James Kerlidis

**Affiliation:** The University of Melbourne

**Talk Title:** *Density-dependent branching processes and their consistent estimation*

**Abstract:** Population-size-dependent branching processes (PSDBPs) and controlled branching processes (CBPs) are two classes of branching processes commonly used to model biological populations living in restricted habitats. We develop connections between PSDBPs and CBPs. We first derive conditions for PSDBPs and CBPs to be equivalent. Then, for CBPs with linearly-divisible control functions, we establish an upper bound on the total variation distance between these CBPs and a non-equivalent PSDBP with matching first and second moments. We show that this bound tends to zero as the initial population size becomes large. We consider these CBPs with linearly-divisible control functions further in the context of consistent estimation. We establish conditions under which consistent parameter estimation may be possible: (1) when the CBP's control function has a known parametrisation, (2) when the control function is unknown, and (3) when the control function is unknown but progenitor counts are observed at each generation. We illustrate these results with numerical examples.

## Fima Klebaner

**Affiliation:** Monash University

**Talk Title:** *Emergence of populations:  $H(W)$  theory*

**Abstract:** We study how populations emerge when starting with just a few individuals, maybe only one, and then growing to its (large) carrying capacity  $K$ .

We prove an old conjecture that initially population can be approximated by a linear process and then follows a nonlinear deterministic path. We suggest a new approximation, which states that from the very beginning, the population behaves as a nonlinear transformation  $H$  of the linear process. In the Laws of Large Numbers (Fluid Approximation), we obtain the classical dynamics but with a random initial condition  $H(W)$ .

## Malgorzata O'Reilly

**Affiliation:** University of Tasmania

**Talk Title:** *Matrix-analytic methods for the evolution of species trees, gene trees, and their reconciliation*

**Abstract:** We consider the reconciliation problem, in which the task is to find a mapping of a gene tree into a species tree, so as to maximize the likelihood of such fitting, given the available data, which may be incomplete.

*Reference:* Matrix-analytic methods for the evolution of species trees, gene trees, and their reconciliation. Albert C. Soewongsono, Jiahao Diao, Tristan Stark, Amanda E. Wilson, David A. Liberles, Barbara R. Holland, Malgorzata M. O'Reilly. Preprint (2023) <https://arxiv.org/abs/2309.06447>

## Krzysztof Szajowski (online)

**Affiliation:** Wroclaw University of Science and Technology

**Talk Title:** *Rationality, Deviation, and Diagnosis: A Holistic Approach to Stopping Games*

**Abstract:** The holistic approach to modeling often involves considering the ontological aspects of mathematical concepts, making our research closely related to ontological studies in mathematics. It provides crucial insights into the nature of decisions and strategies characterized by Markov moments. In a stopping game, a holistic decision-maker would evaluate comprehensive information by assessing the probabilities of various outcomes and their associated payoffs, while understanding the current state, historical data, and potential future scenarios. Such a decision-maker should also consider strategic interactions by anticipating and accounting for the strategies of other players. Flexibility in adapting their strategy as the game evolves, along with the ability to integrate uncertainty by incorporating risk preferences and tolerances, is essential. Additionally, they would perform scenario analysis to evaluate the impact of different stopping times under varying conditions. The goal of such modeling and its implementation in psychological practice is to introduce a novel method for assessing the state of players, leveraging deviations from rational strategies as diagnostic indicators of their psychological and decision-making profiles. Details of other models will be subject to contributed papers.

## Peter Taylor

**Affiliation:** The University of Melbourne

**Talk Title:** *A continuous-time Markov chain model for the distribution of blockchain heights*

**Abstract:** In 2009, the pseudonymous Satoshi Nakamoto published a short paper on the Internet, together with accompanying software, that proposed an 'electronic equivalent of cash' called Bitcoin. At its most basic level, Bitcoin is a payment system where transactions are verified and stored in a distributed data structure called the blockchain. The Bitcoin system allows electronic transfer of funds without the presence of a trusted third party. It achieves this by making it 'very hard work' to create the payment record, so that it is not computationally feasible for a malicious player to repudiate a transaction and create a forward history with the transaction deleted. The Nakamoto paper contained a simple model used to show that the above-mentioned malicious player would be very unlikely to succeed. Unfortunately, this calculation contained an error, which I shall quickly discuss and show how to correct. As its name suggests, the blockchain is comprised of discrete blocks. Blocks are added to the blockchain by 'miners' working across a distributed peer-to-peer network

to solve a computationally difficult problem. With reference to historical data, I shall describe a series of models for the block mining process. They culminate in a recent model for the distribution of blockchain heights that has a number of interesting connections to some well-known stochastic models.

## Nantawat Udomchatpitak

**Affiliation:** Mahidol University

**Talk Title:** *The Accumulation of Beneficial Mutations and Convergence to a Poisson Process*

**Abstract:** We consider a model of a population with fixed size  $N$ , which is subjected to an unlimited supply of beneficial mutations at a constant rate  $\mu_N$ . Individuals with  $k$  beneficial mutations have the fitness  $(1 + s_N)^k$ . Each individual dies at rate 1 and is replaced by a random individual chosen with probability proportional to its fitness. We show that when  $\mu_N \ll 1/(N \log N)$  and  $N^{-\eta} \ll s_N \ll 1$  for some  $\eta < 1$ , large numbers of beneficial mutations are present in the population at the same time, competing against each other, yet the fixation times of beneficial mutations, after a time scaling, converge to the times of a Poisson process.

## Rui Zhang

**Affiliation:** Australian National University

**Talk Title:** *Finding Nash equilibria in random win-loss games*

**Abstract:** We consider random win-lose games, where the entries of the  $n \times n$  payoff matrices are independent Bernoulli random variables with parameter  $p = p(n)$ . We prove that for a wide range of values of the parameter  $p$ , there is an expected polynomial-time algorithm, depending on  $p$ , that computes a Nash equilibrium.