



Danish  
Data Science  
Academy

## Phase-type distributions with a view to population genetics

### Two-day symposium at the Technical University of Denmark

### Wednesday and Thursday April 26-27, 2023

**Aim.** In this symposium we bring together researchers and students from applied probability, statistics, data science and bioinformatics to discuss and – hopefully – develop new and efficient approaches for analysing population genetic data using phase-type distributions.

Phase-type distributions are traditionally used in risk, actuarial science and queuing theory. Phase-type distributions have been a major research theme for Bo Friis Nielsen (professor at the Technical University of Denmark) and Mogens Bladt (professor at Copenhagen University) for more than 25 years. They have described the theory in the monograph 'Matrix-exponential distributions in applied probability' (Bladt and Nielsen, 2017).

Recently, we showed that phase-type distributions can be used to describe the ancestral process for DNA sequences (Hobolth, Jegousse and Bladt, 2019; Hobolth, Bladt and Andersen, 2021). In particular, Kulkarni's reward structure is useful (Kulkarni, 1989). We are currently developing a computational framework for understanding genetic variation (R package `PhaseTypeR` by Rivas-Gonzales, Andersen and Hobolth, 2023). These efforts are complemented by a computational framework for risk (R package `matrixdist` by Martin Bladt and Yslas, 2022).

A major research area is to move from general and theoretical properties of phase-type distributions to procedures for inference and estimation from population genetic data. The aim of the symposium is to bridge this gap.

**Plan for the two days.** The symposium kicks off on Wednesday afternoon, April 26, with the PhD defence of Nicolai Siim Larsen. The title of the thesis is 'New results on multivariate phase-type distributions', the supervisor is Bo Friis Nielsen and the opponents are Anders Stockmarr, Qi-Ming He and Asger Hobolth. Qi-Ming He is a professor in operations research at University of Waterloo in Canada and is author of the book 'Fundamentals of matrix-analytical methods' (He, 2013). Qi-Ming He will give a key note lecture of the symposium on Thursday morning. The remaining part of that day consists of many short talks and opportunities for discussion.

The symposium will take place at the Technical University. The participants will arrive in due time to attend Nicolai Siim Larsen's PhD defence. The program for the two days is summarized below.

**Participation.** Please send an email to Asger (asger@math.au.dk) if you want to participate. Participation is free of cost, but we have a ceiling of 24 participants.

**Wednesday 26/4 in B324, R240:**

13.00-16.00	PhD defence Nicolai Siim Larsen: New results on multivariate phase-type distributions. Supervisor Bo Friis Nielsen Opponents Qiming He, Asger Hobolth and Anders Stockmarr
16.00-18.00	Reception
19.00	Dinner for people attending the symposium.

**Thursday 27/4:**

10.00 - 12.00	Session 1
12.00 - 13.00	Lunch
13.00 - 14.30	Session 2
14.30 - 15.30	Walk-and-talk
15.30 - 17.00	Session 3
17.00 - 20.00	Dinner
20.00	End of symposium.

**Session 1:** Theory and applications of PH-type in population genetics

Welcome and chair: Bo Friis Nielsen and Asger Hobolth

Key note lecture by Qiming He 45+15 (talk 45 minutes; questions 15 minutes)

Charlotte Brunnenberg (MSc student at AU): Coalescent with recombination) 20+10

Nicoline Sørensen and Andreas Osted (BSc students at AU): Coalescent with structure 20+10

**Session 2:** Software and computational methods for PH-type

Chair: Marta Pelizzola (PostDoc at AU)

Martin Bladt (Ass prof at KU: Insurance and `matrixdist`) 20+10

Iker Rivas-Gonzalez (PhD student at AU: Population genetics and `PhaseTypeR`) 20+10

Hans Martin Sandager and Tobias Overgaard (MSc student at DTU): Inverse Laplace transform for MPH-type 20+10

**Session 3:** Extensions of the basic PH-theory with a view to population genetics

Chair: Nicolai Siim Larsen (Ass prof at DTU)

Mogens Bladt (Prof at KU): Multivariate Inhomogeneous PH-type 20+10

Bo Friis Nielsen (Prof at DTU): Multivariate Discrete PH-type 20+10

Asger Hobolth (Prof at AU): Sampling formula for the coalescent 20+10

**References**

- Bladt, M and Yslas, J (2022). R package *matrixdist*: Statistics for matrix distributions. Available from CRAN.
- Bladt, M and Nielsen, B.F. (2017). *Matrix-Exponential distributions in applied probability*. Springer, New York.
- He, Qi-Ming (2013). *Fundamentals of Matrix-Analytic Methods*. Springer, New York.
- Hobolth, A., Bladt, M. and Andersen, L.N. (2021). Multivariate phase-type theory for the site frequency spectrum *Journal of Mathematical Biology*, **83**, 63-91.
- Hobolth, A., Siri-Jégousse and Bladt, M. (2019). Phase-type distributions in population genetics. *Theoretical Population Biology*, 127, 16-32.
- Kulkarni, V.G. (1989). A New Class of Multivariate Phase Type Distributions. *Operations Research*, **37** 151-158.
- Rivas-González, I., Andersen, L.N. and Hobolth, A. (2023). PhaseTypeR: an R package for phase-type distributions in population genetics. *Journal of Open Source Software* 8 (82), 50–54.