Project Proposals: Final list

Project 1.
Title. Stability of ecosystems with a core-periphery structure.
Supervisors. Jacopo Grilli (The Abdus Salam ICTP), Fernando L. Metz (Federal University of Rio Grande do Sul).

Summary. Understanding which properties of ecological interactions make ecosystems stable to perturbations is of paramount importance. Stable stationary states are usually beneficial and associated with a well-functioning system. In a seminal paper, Robert May derived, using random matrix theory, universal criteria for the linear stability of fixed-points modeling ecological communities as randomly coupled differential equations. May's model relies on the assumption that species interact randomly with each other. In this setup, any ecological system is unstable, provided it contains a sufficiently large number of species. Although this random-matrix approach led to important insights, the results are clearly unrealistic, since real ecosystems are often large and stable.
In contrast to May's original model, ecological interactions are organized in a network structure, where nodes represent species and links stand for the pairwise interactions among them. In particular, interactions are sparse, which means that a given species influences or interacts with a finite number of others. The stability of large ecosystems, being the interactions sparse or not, boils down to the study of the leading eigenvalue of a certain large random-matrix encoding the interactions among different species – the community matrix.
The main goal of the present project is to explore how the core-periphery architecture of certain ecosystems affects their stability. Core-periphery networks are characterized by a densely connected subset of species (the core) containing a smaller number of links with the rest of the network (the periphery). Such models are important in ecology because many empirical interaction networks have this feature. The central idea of the project is to compute the leading eigenvalue of the random community matrix as a function of the structural parameters characterizing a simple model of a core-periphery ecosystem. The student will employ analytical techniques from statistical physics and numerical diagonalization routines to study this problem.

Bibliography.
[1] J. Grilli, T. Rogers, and S. Allesina, Modularity and stability in ecological communities, Nat. Communications 7, 12031 (2016).
[2] F. L. Metz, I. Neri and T. Rogers, Spectral theory of sparse non-Hermitian random matrices, J. Phys. A: Math. Theor. 52, 434003 (2019).

==

Project 2.
Title. The evolution of income or wealth distribution with higher order autoregressive processes.
Supervisors. Yonatan Berman (London Mathematical Laboratory) and Ravi Kanbur (Cornell University).

Summary: The standard reduced model for the evolution of income (or wealth) over time is the Gibrat process, where the income of the next (children's) generation depends on the income of the

current (parent's) generation plus an error term. When the error term is i.i.d., this leads to a first order autoregressive process. Under certain conditions on the intergenerational elasticity of income (IGE), a steady state distribution of income exists. Furthermore, the IGE can be backed out from the appropriate second moment of the distribution, the steady state inequality. The steady state snapshot in effect encodes the dynamics of the process.
When the error term in the above set up is not i.i.d. but is itself first order serially correlated, the income process becomes a second order autoregressive process. The economic microfoundations of the above reduced form with second order autoregression have been developed by Solon (2018) following on from the classic work of Becker and Toomes, based on first order serial correlation in transmission of parental genetic and family advantages. Again, under certain conditions a steady state distribution exists, and is determined by the IGE and the first order serial correlation coefficient. Clearly, more information is now needed about the steady state distribution to infer these two parameters of the dynamic process. One question is how information on the third moment of the steady state distribution, its skewness, could be used to back out the IGE and the serial correlation coefficient.
There is a further argument that the grandparents' incomes influence grandchildren's incomes. In other words, income follows a third order autoregressive process, which could be rationalized by second order serial correlation in the error term in a Gibrat model (see Solon, 2018, equation 5). The evolution of income distribution under third and still higher order autoregressive processes remains understudied in the literature. How does the steady state depend on the parameters of the dynamic process, and what information is needed on the steady stat snapshot to back out theses parameters? A thorough exploration of the reduced form is of interest in its own right but also as a possible route to providing insight into elements of economic microfoundations.
A thorough grounding in autoregressive time series analysis and in numerical simulation of the evolution of the distributions emerging from these processes is needed for participation in this project.

Bibliography.
Berman, Yonatan and Ravi Kanbur. 2020. "Two Types of Income Mobility." In preparation.
Solon, Gary. 2018. "What Do We Know So Far About Intergenerational Mobility?" Economic Journal, Volume 128, Issue 615, pp. F340–F352.

==

Project 3.
Title. Exploring the Connections between Information Theory and Ergodicity Economics Through the Lens of Inequality Measurement.
Supervisors. Ravi Kanbur (Cornell University), Yonatan Berman and Alex Adamou (London Mathematical Laboratory).

Summary. Ergodicity economics leads to the Mean Log Deviation (MLD) as a measure of income inequality (Adamou, Berman and Peters, 2020). Information theory also leads to the MLD as a measure of inequality. In fact MLD is one member of the "Generalized Entropy" class of inequality measures, which contains the Theil index and MLD, also known as "Theil's second measure" (Theil, 1967; Cowell, 1980, 2003).

This exploratory project will investigate the possible deep structural reasons why such seemingly different perspectives lead to the same measure of inequality. This is analogous to the reasons for decomposability restrictions on inequality measures leading to the Generalized Entropy class of inequality indices (Shannon's information axioms have decomposability built into them).
If a connection can be established between the entropy perspective from information theory and the ergodicity economics perspective through the lens of inequality measurement, this opens up the possibility of relating generalized entropy from information theory to potential generalizations of ergodicity economics.
The project is exploratory and speculative, with an uncertain but potentially exciting outcome of providing new insights by investigating the intersections between two the two perspectives of information theory and ergodicity economics to inequality measurement.
A thorough understanding of entropy, particularly as applied to information theory, and of ergodicity economics, will be required for participation in this project.

Bibliography.
Adamou, Alexander, Yonatan Berman and Ole Peters. 2020. "The Two Growth Rates of the Economy."
Cowell, F. A. 1980. "Generalized entropy and the measurement of distributional change". European Economic Review 13, 147—159.
Cowell, F. A. 2003. "Theil, Inequality and the Structure of Income Distribution" Distributional Analysis Discussion Paper, 67, STICERD, London School of Economics.
Kanbur, Ravi. 2019. "Some New (Old) Directions in Entropy, Information and Measurement of Income Inequality." Keynote Address, Latin American Meetings of the Econometric Society, Puebla, Mexico.
Theil, H. (1967). Economics and Information Theory. Amsterdam: North Holland.

==

Project 4.
Title. Growth incidence curves in reallocating geometric Brownian motion.
Supervisors. Yonatan Berman and Alex Adamou (London Mathematical Laboratory).

Summary. There are two ways of evaluating changes in the distribution of wealth in a given population, depending on whether initial wealth levels are taken into account or not. If they are ignored, the new distribution is compared to the original one without considering the identity of wealth owners. The comparison is thus 'anonymous'. Poorest (richest) in the initial period are compared to poorest (richest) in the final period whoever their identity is. In the other case, the comparison is made between the two distributions conditionally on initial wealth. The comparison thus is 'non-anonymous'.
It has become common to represent the anonymous distributional change brought by economic growth by the Growth Incidence Curve (GIC) [1]. This very convenient tool simply shows the rate of growth of successive quantiles of the distribution in the initial and final periods. It turns out that there is a relationship between this curve sloping downward and the Lorenz curve of the marginal distribution shifting upward. Non-anonymous Growth Incidence Curves (NAGICs) are the non-anonymous equivalent [2]. Interestingly, in many periods in which the anonymous GICs were generally upward sloping, non-anonymous GICs were largely flat.

With these results in mind we study reallocating geometric Brownian motion (RGBM). This is a simple model of an economy in which wealth undergoes noisy exponential growth. We add to that framework a reallocation mechanism in which a fraction of everyone's wealth is pooled and shared [3]. The goal of this project is to study the anonymous and non-anonymous growth incidence curves predicted by the RGBM model, and their properties. We will compare these predictions to empirical evidence on growth incidence curves.
The student will have to be comfortable working with data and coding simple models.

Bibliography.
[1] Martin Ravallion and Shaohua Chen. Measuring pro-poor growth. Economics Letters, 78(1):93– 99, 2003.
[2] Franc ̧ois Bourguignon. Non-anonymous growth incidence curves, income mobility and social welfare dominance. The Journal of Economic Inequality, 9(4):605–627, 2011.
[3] Yonatan Berman, Ole Peters, and Alexander Adamou. Wealth inequality and the ergodic hypothesis: Evidence from the United States. Journal of Income Distribution, 2020. Available at SSRN.

==

Project 5.
Title. Optimal allocation and risk parity.
Supervisors. Mark Kirstein, Yonatan Berman and Alex Adamou (London Mathematical Laboratory).

Summary. Common investment advice has it that stock markets outperform less volatile investments in the long run. If this is true, why not borrow money and leverage stock market investments? The answer to this question can be found in a 2011 publication [1]. Here, it was predicted that the leverage that optimizes time-average growth in a portfolio containing a riskless and a freely traded risky asset should be close to 1, meaning borrowing won't help in the long run. Another common investment advice, especially popular in the last decade, has it that when selecting assets in a portfolio it is advisable to allocate risk, usually defined as volatility. This approach is called risk parity [2]. It asserts that when asset allocations are adjusted to the same risk level, the risk parity portfolio is more resistant to market downturns than in other situations. For example, traditional portfolio allocation of 60% stocks and 40% bonds carries about 90% of its risk in the stock portion of the portfolio. In a risk parity portfolio volatility is allocated equally between assets. How is such allocation compatible with the optimally leveraged portfolio? Answering this question is the goal of this project.
First, we will extend the optimal leverage concept of [1] to an allocation between several risky assets that still optimizes time-average growth. We will extend this result to a case in which the assets are cross-correlated in time. We will then compare the prediction of this approach to risk parity and study how they are related. We will also test the performance of the two approaches with data.
The student will have to be comfortable working with data and coding simple models. As preparation the following material should be studied – chapter 5 in [3] and [1].

Main goal.

In this project we will extend the optimal leverage criterion to optimal allocation (with more than two assets) and compare it to the risk parity approach in portfolio management.

Bibliography.

[1] Ole Peters. Optimal leverage from non-ergodicity. Quantitative Finance, 11(11):1593–1602, November 2011.
[2] S´ebastien Maillard, Thierry Roncalli, and J´erˆome Te¨iletche. The properties of equally weighted risk contribution portfolios. The Journal of Portfolio Management, 36(4):60–70, 2010.
[3] Ole Peters and Alexander Adamou. Ergodicity economics, 2018. 5.0, 2018/06/30.

==

Project 6.
Title. Time-average based saving versus Ensemble-average based saving and the evolution of wealth distribution.
Supervisors. Alexander Adamou (London Mathematical Laboratory), Yonatan Berman (London Mathematical Laboratory), Ravi Kanbur (Cornell University) and Ole Peters (London Mathematical Laboratory).

Summary. For multiplicative growth processes, time-averages differ from ensemble-averages. This feature is central in the differential between GDP and DDP as developed in Adamou, Berman and Peters (2020).
A key element of any economic growth process, particularly the evolution of the wealth distribution, is savings behavior. In mainstream economics there are theories of savings in stochastic environments, all based on Expected Utility and thus on an ensemble-average perspective.
However, a central tenet of Ergodicity Economics is that individuals respond behaviorally to the time-average of a stochastic process rather than to its ensemble average. The central question of this project is: What does the evolution of the wealth distribution look like under different specifications of savings behavior as dependent on time-averages rather than ensemble averages?
A first step in this project would be to develop a very simple model of the evolution of wealth distribution where the rate of return is stochastic and the savings rate out of the return from wealth (there are no other sources of income) depends on the time-average of income (for different time horizons). Different types of dependency can be parametrized (e.g., constant elasticity, allowing experiments with the elasticity value, etc.). This evolution can then be compared with a conventional mainstream economics savings model, and variants on this model, for example models with borrowing constraints (see Nobel Committee, 2015, Section 2).
Familiarity with the concept of ergodicity and with conventional economics literature on savings under uncertainty (as in the Nobel prize winning work of Angus Deaton, see Nobel Committee, 2015, Section 2), and the ability to program simple numerical simulations, will be needed for participation in this project.

Bibliography.

Adamou, Alexander, Yonatan Berman and Ole Peters. 2020. “The Two Growth Rates of the Economy.”

Berman, Yonatan, Ole Peters and Alexander Adamou. 2020. "Wealth Inequality and the Ergodic Hypothesis: Evidence from the United States."
Nobel Committee. 2015. "Angus Deaton: Consumption, Poverty and Welfare", Microsoft Word - 2015_final3.docx (nobelprize.org)
Peters, Ole. "The Ergodicity Problem in Economics." The ergodicity problem in economics | Nature Physics

==

Project 7.
Title. Learning random dynamical systems from data.
Supervisors. Jeroen Lamb (Imperial College), Yuzuru Sato (Hokkaido University), Stefano Luzzatto (Abdus Salam ICTP)

Summary. An important aspect of the modelling of physical and societal processes, is to find models that are conceptually meaningful and able to represent the reality as it is observed. Random dynamical systems form a class of models where the evolution is governed by equations of motion that intrinsically depend on a random variable, which is usually denoted "noise". Mathematical models of systems with noise have become increasingly popular, motivated by application areas in the context of complex systems, arising naturally in the context of the life sciences and socio-economical sciences, for instance. The approach taken in the large majority of these models, is that the understanding of averaged quantities is emphasized at the expense of dynamical properties of (typical) trajectories. However, in many applications, knowledge about trajectories, beyond their averages, are very important.

It is important to be able to decide from observations of a dynamical system, which model best represents the system. While this question has been addressed extensively for deterministic models as well as for some classes of stochastic models (ensuring good fits with averages), there has been relatively little work on the how to determine the suitability of a random dynamical system. The objective of the project is to develop methods to learn random dynamical systems from observations (data), building on existing methods in the deterministic and stochastic settings.

==

Project 8.
Title. Explainability and rationalization in decision theory: a coding theory approach.
Supervisors. Matteo Marsili (Abdus Salam ICTP) and Isaac Pérez Castillo (Universidad Autónoma Metropolitana).

Summary. The necessity to explain or communicate our decisions may impose limits to rationality, in cases where the explanation of a decision process is too complicated. For example, politicians need to gather consensus by motivating their choices about complex issues, and even at the individual level, we may prefer to rationalise our choices in simple ways in order to minimise our cognitive load (also in terms of remembering why we took a particular choice). Explainability in AI is a hot topic. Each decision an expert system takes can be explained in terms of the algorithm used and of the data in input, yet such a microscopic justification is too complex for most users.

The project aims at investigating how the necessity to have a simple rationalisation distorts rationality, taking a coding theory approach.
Let us consider a generic decision problem. Think of the choice of the treatment s for a particular patient. There are S possible choices and the probability that the optimal choice of *s* is [1,2]

$$p_s = \frac{1}{Z} e^{u_s} \qquad Z = \sum_{s=1}^{S} e^{u_s} \tag{1}$$

A rationalisation of the choices is equivalent to a code that maps each outcome s to a binary codeword, each bit of which specifies why a particular binary choice was taken. Let ls be the length of codeword that corresponds to s. An optimal rationalisation is one that minimises the expected length of the explanation. Coding theory [3] implies that the solution of the problem

$$\min_{\ell} \sum_{s=1}^{S} p_s \ell_s$$

is given by

$$\ell_s = -\log_2 p_s$$

(neglecting integer constraints). The minimal expected code length is given by the entropy

$$\min_{\ell} \sum_{s=1}^{S} p_s \ell_s = H[p] = -\sum_{s=1}^{S} p_s \log p_s$$

Yet, taking rational choices in this way may lead to choices which are hard to explain. What is the best strategy to take decisions whose rationalisation is at most of length l, in expected terms?

Bibliography.
[1] Marsili, M. (2019). The peculiar statistical mechanics of optimal learning machines. Journal of Statistical Mechanics: Theory and Experiment, 2019(10), 103401.

[2] The variable *s* can be thought of as the observable part of a larger decision problem

$$\max_{s,t} U(s,t)$$

where *U (s, t)* = u_s + *V (s, t*) and both the variables *t* and the function *V (s, t)* are not observable. In particular, if *V (s, t)* is a random variable, drawn independently for all *s, t*, with all finite moments, then [1] shows that the probability that s is chosen is given by Eq. (1). In the example of the choice of a treatment, recommended by deep learning, *t* can be related to the state of hidden variables and *V (s, t)* can depend on the data on input, for a particular patient.

[3] See Cover, Thomas M. Elements of information theory. John Wiley & Sons (1999), Chapter 5.

===

Project 9.
Title. Beyond Rational Herding.
Supervisors. Matteo Marsili (Abdus Salam ICTP) and Isaac Pérez Castillo (Universidad Autónoma Metropolitana).

Summary. Rational herding is the process by which individuals follow the majority opinion of their peers, instead of following their independent signals or seeking further information, when they have to form an opinion about a subject (e.g. whether the 2020 US Elections were stolen or not). The main insight is that, if the peers of an individual follow independent signals they receive, the majority aggregates that information and provides a much more accurate signal. This is known as the wisdom of the crowd phenomenon. Yet, if every- body follows their peers, no-one seeks independent opinion and there is no information to aggregate. As a result, the crowd behaves in a random manner.
This phenomenon has been studied by many authors, in particular in Why does individual learning endure when crowds are wiser? by de Courson et al. (arXiv:2012.14524 [physics.soc-ph]) who extend he model in Phase coexistence in a forecasting game by P. Curty and M. Marsili (J. Stat. Mech. (2006) P03013). de Courson et al. explore the conditions and mechanisms under which the the wisdom of the crowd may be preserved. In particular they observe that:

1. The network of the peers is adaptive. Whom everyone listens to is part of the strategy.
2. Agents should account for the diversity of the signals they receive, because the signals they receive from different neighbours may be highly correlated. Agents should favour independent signals.
3. Agents care about reputation. They care about how many other agents listen to them.
4. Agents may weigh their peers' opinions by their accuracy.

With the advent of social networks, we're witnessing an explosion of new collective phe- nomena, such as fake news or the formation of information bubbles. There's a growing and rapidly expanding literature of empirical research but yet an unsatisfactory theoretical un- derstanding. The exploration of how these phenomena are linked or promoted by different mechanisms is best elucidated in the simplest possible models. For this reason, the project aims at identifying the simplest possible modification of the basic models that can still be in the reach of analytical approaches, but that can shed light on these phenomena.
A possible direction is that of studying "likes-games" where the utility function of agents depends on how many "likes" he receives from his peers. In a simple model that takes this into account, if A's prediction is right and B listens to A, then B will put a "like" on A's page. Then A would receive likes in proportion to the number of neighbours s/he has and in proportion to his probability to be right. Yet, even if A is right all the time, if no-one follows him, he would not get any payoff in this game. Agent B who has many followers, may be better off even if he's wrong most of the time. This already shows the tension between accuracy and popularity that we'd like to explore.
A further dimension to explore in these models are the rules by which the network of peers is updated1: neighbours who are i) least liked and/or ii) too correlated with others may be replaced by a) random agents or b) the most liked neighbour of the most liked neighbour. Results from The

Social Climbing Game, by Bardoscia et al.)Journal of Statistical Physics 151, p. 440457 (2013)) can be used to study the networkk dynamics analytically, in some cases.

==

Project 10.
Title. Verifying interaction among different types of discrete events by using the multivariate Hawkes process.
Supervisors. Jiancang Zhuang (ISM, Japan) Max Werner (U. Bristol, UK)

Summary. A multivariate Hawkes process (MHPs) is a point process with several sub-processes of discrete events where each subprocess affects the occurrence rates of the other subprocesses. The MHP's ability to capture mutual excitation among different types of phenomena makes it a popular model in many fields. For a p-dimensional MHP, the intensity function of its ith subprocess takes the following form:

$$\lambda_i(t) = \mu_i(t) + \sum_{j=1}^{p} \int_{-\infty}^{t_-} g_{ij}(t-s) dN_j(s)$$

where the constant $\mu_i(t)$ is the background intensity of the ith subprocess, $N_j(t)$ counts the number of arrivals in the jth subprocess, and $g_{ij}(t)$ is the excitation effect caused by an event at time 0 in the jth subprocess to the occurrence rate at time t $(t > 0)$ in the ith subprocess. Therefore, estimating the excitation response function is the key to learning an MHP model.

In this project, we intend to use the MHP to analyze real datasets. The supervisors have prepared the data of earthquake hypocenters from different geographical regions and are interested in finding the correlations between deep and shallow earthquakes. This would reveal exciting new scientific insights into what earthquake processes. Depending on time and interest, we may consider extensions of the multivariate Hawkes processes, including (1) complicated background rates and (2) inhibition effects among some types of events. The student may also bring their own data.

Requirements.
1. Good knowledge in probability and statistics.
2. Programming skill in Python, Matlab, R, Fortran, or C/C++.

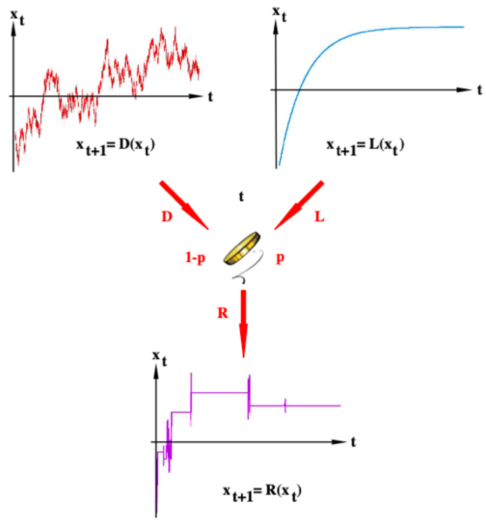
Bibliography.
1. Ogata Y., Akaike H. and Katsura K. (1982). The application of linear intensity models to the investigation of causal relations between a point process and another stochastic process, Annals of the Institute of Statistical Mathematics, Vol.34, No.2, B, pp.373-387
2. Zhuang J., Ogata Y. and Vere-Jones D. (2002). Stochastic declustering of space-time earthquake occurrences. Journal of the American Statistical Association, 97: 369-380.
3. Zhuang J., Ogata Y., Vere-Jones D. (2004). Analyzing earthquake clustering features by using stochastic reconstruction. Journal of Geophysical Research, 109, No. B5, B05301, doi:10.1029/2003JB002879.
4. Mohler G. O., Short M. B., Brantingham P. J., Schoenberg F. P. & Tita G. E. (2011) Self-Exciting point process modeling of crime, J. Amer. Statist. Assoc., 106:493, 100-108.

5. Zhuang, J. and Mateau, J. (2019). A semi-parametric spatiotemporal Hawkes-type point process model with periodic background for crime data. Journal of the Royal Statistical Society, Ser. A. 182(3), 919-942. doi:10.1111/rssa.12429.

==

Project 11.
Title. Anomalous diffusion in random dynamical systems
Supervisors. Rainer Klages (Queen Mary University of London), Yuzuru Sato (Hokkaido University) and Stefano Ruffo (SISSA).

Summary. Scientific Fields: dynamical systems theory, stochastic theory, anomalous diffusion, computer simula- tions, nonequilibrium statistical physics
Background: Sampling randomly in time between different types of deterministic dynamics defines a random dynamical system. The figure below illustrates the basic idea. This topic thus combines the fields of dynamical systems theory and stochastic processes. The offered project is motivated by very recent and ongoing research in which it was shown that random dynamical systems can generate anomalous diffusion [1]. While ordinary Brownian motion is characterized by a mean square displace- ment that grows linearly in the long time limit, for anomalous diffusion the spreading of particles grows either slower than linear in time (subdiffusion) or faster (superdiffusion) [2,3].
Project Description: There are a number of very interesting open questions to work on within this context, which are: (i) calculating mean square displacements; (ii) calculating correlation functions; (iii) constructing and testing superdiffusive models; (iv) exploring strong anomalous diffusion; (v) occupation time statistics. The first two problems can be explored analytically supported by numerics [4] while the latter three require computer simulations along the lines of [1]. Thus ample numerical and programming skills are required for this project, some familiarity with simple time-discrete dynamical systems and with basic stochastic processes. As a warm-up the student should familiarize her/himself with the model of [1]. This entails understanding analytical results for calculating the invariant density of the underlying map of the model [5] as well as learning about stochastic anomalous diffusion [2,3].



Bibliography.
[1] Y.Sato, R.Klages, Phys.Rev.Lett. 122, 174101 (2019)
[2] R.Klages, G.Radons, I.M.Sokolov (Eds.), Anomalous transport: Foundations and Applications (Wiley-VCH, Weinheim, 2008)
[3] J.Klafter, I.M.Sokolov, First steps in random walks (Oxford, 2011) [4] G.Knight, R.Klages, Nonlin. 24, 227 (2011)
[5] S.Pelikan, Trans.Am.Math.Soc. 281, 813 (1984)

==

Project 12.
Title. The trade-off between profit and time constraints and market impact.
Supervisors. Matteo Marsili (Abdus Salam ICTP) and Isaac Pérez Castillo (Universidad Autónoma Metropolitana).

Summary. The Glosten-Milgrom model describes how markets aggregate private information held by informed traders, by setting prices optimally in order to respond to their trading activity. The model is simple enough to allow for a full analytic solution. This shows that an informed trader can extract an expected gain that satisfies the inequality

$$\mathbb{E}[G] \leq T(\nu) H[Y]$$

where H[Y] is the entropy that quantifies the information held by the trader on future returns, whereas

$$T(\nu) = \left[\frac{1+\nu}{2} \log \frac{1+\nu}{1-\nu}\right]^{-1}$$

is the market temperature, that depends on the frequency ν with which the informer traders submits orders.

We refer to L. Touzo et al. for more details (see https://arxiv.org/abs/2010.01905). In brief, when the informed trader exploit his information to extract a gain from the market, he also reveals information to the market maker that adjusts prices to as to take the revealed information into account, thereby decreasing the gains of future trades by the informed trader.
In particular, a patient trader can extract an infinite gain by setting ν arbitrarily small, because $T(\nu) \to \infty$ as $\nu \to 0$. Yet this gain would accrue over an infinite time. At the same time, in this limit, the informed trader will leak very little information to the market trader, so the impact of trading activity on the prices is very weak.
The project explores the tradeoff between time constraints of informed traders (impa- tience) and the market impact (i.e. how fast prices react to trading) in the simple setting of the Glosten-Milgrom model, building on the results of the paper by L. Touzo et al. mentioned above.

==

Project 13.
Title. Large deviation theory in diluted random matrices.
Supervisors. Isaac Perez Castillo (Universidad Autónoma Metropolitana)

Summary. We will explore a general method to obtain the exact rate function controlling the large deviation probability that an sparse random matrix has an extensive number of eigenvalues inside a closed interval [a,b]. We will then use this method to study the eigenvalue statistics in two distinct examples: (i) the shifted index number of eigenvalues for an ensemble of Erdös-Rényi graphs and (ii) the number of eigenvalues within a bounded region of the spectrum for the Anderson model on regular random graphs. The theoretical findings will be to numerical diagonalization. Time permitting, we will extend this study to diluted asymmetric matrices.

Bibliography.
[1] Large Deviation Function for the Number of Eigenvalues of Sparse Random Graphs Inside an Interval, Fernando L. Metz and Isaac Pérez Castillo, Phys. Rev. Lett. **117**, 104101(2016).
[2] Analytic approach for the number statistics of non-Hermitian random matrices, Antonio Tonatiúh Ramos Sánchez, Edgar Guzmán-González, Isaac Pérez Castillo, Fernando L Metz (arXiv:2007.10526).

==

Project 14.
Title. Tackling Quantum Many-Body Systems with Artificial Neural Networks.
Supervisors. Joe Bhaseen (King's College London) and Isaac Perez Castillo (Universidad Autónoma Metropolitana)

Summary. A key challenge in quantum many-body systems is to obtain the wavefunction that encapsulates all the non-trivial quantum correlations. Apart from a few cases, for which the wavefunction can be obtained exactly, one must rely on approximation methods, or study a reduced set of correlation functions. Lately, artificial neural networks have been used to obtain variational approximations to the quantum many-body wave function. The aim of this project is to investigate the utility of artificial neural networks for applications in quantum physics, with a view towards novel states of quantum matter. The project will involve the theoretical analysis of machine learning techniques, and numerical investigation of paradigmatic quantum many-body systems.

Bibliography.

[1] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science 355, 602 (2017).
[2] S. Das Sarma, D.-L. Deng, and L.-M. Duan, Machine learning meets quantum physics, Physics Today 72, 3, 48 (2019).

==

Project 15.
Title. Experiments in ergodicity.
Supervisors. Ollie Hulme (Copenhagen University Hospital Hvidovre), Ole Peters (LML), Alex Adamou (LML), Yonatan Berman (LML), Mark Kirstein (LML)

Summary. As a relatively new field, ergodicity economics (EE) makes a number of falsifiable predictions which deviate both quantitatively and qualitatively from other theories of decision making (e.g. EUT, Prospect theory, Isoelastic utility and their variants). Though some experimental work has begun, very few of these behavioral predictions have been adequately tested, replicated, or generalised. The existence of conflicting predictions, and a paucity of evidence to adjudicate the conflict, is motivation for a series of behavioral experiments that are planned in collaboration between LML and Danish Research Centre for Magnetic Resonance. The experiments will be designed with the aim of discriminating between these candidate theories, as rigorously as possible. Central to EE is the premise that agents should modulate their decision-making strategies as a function of the dynamics that they face. Though other theories are typically not predicated explicitly on dynamics, many will make predictions that depend on dynamics, implicitly via other features that may depend on such dynamics. Thus, the task of deriving predictions from non-EE theories under different dynamics, is not always straightforward. In this project you will work with theorists and experimentalists, to design behavioral experiments that are capable of discriminating between such theories. The focus will be on risk or time preferences at the level of the individual agent. The risk preference project will largely centre around the protocol of Meder et al. 2019, and will focus primarily on the replication in a larger sample, and generalisation to new dynamics which predict qualitatively different risk preferences. Particular attention will be paid to addressing issues raised by the community, in terms of controlling for moments, the consequentiality of the design, and end-wealth computations. The time preference project will focus on testing the experimental predictions of Adamou et al. 2019, using equivalent strategies as used in the risk preference project. Other experimental ideas can be discussed and developed, subject to the interests of the student, but the main development work will centre on one or both of these experimental questions. Time permitting there will be scope for considering different experimental scales, from smaller cases studies comprising tens of subjects, larger cohorts of hundreds, and larger scale app-based paradigms sampling potentially thousands of subjects.

Goal of the project. The overarching goal of this project is to develop the experimental designs such that they are optimised to maximally discriminate between the key theories of interest. The experimental design, along with all code, as well as synthetic data will be pre-specified as part of registered reports that will be submitted to a journal prior to data collection.

Details. There are several challenges embedded in this endeavour which will shape the work of the project: *Model specification.* If a decision theory is widely used in the problem domain then we will take it seriously, regardless of theoretical considerations. The challenge is to specify all candidate models in their strongest form, taking measures to strengthen them if necessary. This may be achieved via a close reading of the literature, through your existing experience, and through open dialogue with domain specific experts who ideally, though not necessarily, will be sceptical of EE. This is mirrored by the EE experts on this team specifying the EE model in its strongest form. A working knowledge of theories in these domains will be advantageous, but not strictly necessary; *Modeling.* The candidate models will be primarily specified as hierarchical Bayesian models, estimated via MCMC techniques, though other techniques can be considered. Working knowledge of Bayesian statistics, models, and their variants will be advantageous; *Simulation.* Theoretical predictions must be pre-registered and coded prior to data collection, as far as possible. As such, experiments should be iteratively designed with the models, maximising the probability of obtaining compelling evidence for one way or another. Simulations of synthetic agents, derived from candidate theories, can be used to evaluate the effect of experimental design choices on the evidence accumulation rates under plausible assumptions. A strong background in coding will be essential. Expertise in app development will be considered a bonus; *Timeframe.* Only some of the work described above will be completed during the summer school, so care will be taken to ensure that the project undertaken is achievable under the time constraints.

Bibliography.
EE lecture notes: https://ergodicityeconomics.com/lecture-notes/
Meder et al 2019: https://arxiv.org/abs/1906.04652
Adamou et al 2019: https://arxiv.org/abs/1910.02137
Bayesian Cognitive Modelling: http://faculty.sites.uci.edu/mdlee/files/2011/03/BB_Free.pdf

==

Project 16.
Title. Epidemic spreading models in the presence of feedback
Supervisors. Fabio Caccioli (University College London, UK) and Daniele De Martino (Biofisika Institute – Ikerbasque, Spain)

Summary. The project will consider simple epidemic spreading models with a time-dependent infection rate. More specifically, the infection rate at time t will depend on the estimation of time-dependent quantities such as the effective reproduction number Rt. We will consider different dependencies that mimic different strategies aimed at containing the spreading, and we will compare their performance. We will in particular study the onset of endogenous oscillations induced by the feedback mechanism.

The project will be carried out by means of numerical simulations. The successful candidate should have a good knowledge of coding.

Bibliography.

- Caccioli F. and De Martino D. (2020) “Epidemic oscillations induced by social network control: the discontinuous case” arXiv:2012.02552
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). "Epidemic processes in complex networks." Reviews of modern physics, 87(3), 925.
- De Martino, D. (2019). Feedback-induced self-oscillations in large interacting systems subjected to phase transitions. *Journal of Physics A: Mathematical and Theoretical*, *52*(4), 045002.