EXCELENCIA
MARÍA
DE MAEZTU

IFISC

Universitat de les Illes Balears

CSIC

PROCEEDINGS OF THE 2023 SURF@IFISC FELLOWSHIPS

Proceedings of the 2023 SURF@IFISC Fellowships

The SURF (Summer Undergraduate Research Fellowships) Program is offered by IFISC since 2013 (only interrupted in 2020 due to the Covid-19 pandemic) with the goal of attracting excellent undergraduate students with majors compatible with IFISC research lines and offers Summer internships at IFISC, typically during the month of July, advised by members of IFISC Claustro.

SURF attracted for the Summer of 2023 a total of 50 candidates (30 male and 20 female), out of which 6 candidates were selected, 3 of them female and 3 male candidates. The candidates came from 30 different universities belonging to 13 different countries, while the candidates were citizens of 11 different countries. The universities with more candidates were the Universities of Barcelona and Valencia, with 5 each, and the U. of Salamanca with 4 applicants. Almost half of the applicants, 24, were Spanish citizens, and in second place comes China with 13 applicants.

These are the proceedings:

Stochastic Thermodynamics and Information

Iván Salvador, Gonzalo Manzano
Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB)
Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain

Abstract

In this work we start providing a short review of Stochastic Thermodynamics and discussing a recent realization of a Maxwell demon that appeared in the a recent article using the equivalence between Clausius and Shannon entropy. Then, we review the Gillespie algorithm that allow us to study Stochastic Thermodynamics at the trajectory level. We numerically apply it to two systems. First, a toy model consisting of a system with three energy levels coupled to three thermal baths. We numerically check the validity of some fluctuation theorems for this system. Finally, for the Maxwell demon model we verify some fluctuation theorems, discuss the régimes of operation and quantitatively discuss the improved lower bound on the total entropy production presented in the mentioned paper.

1 Introduction

Classical Thermodynamics is usually restricted to the description of equilibrium states of macroscopic systems, that undergo transformations involving exchange of heat, work and matter with an environment. Those transformations are governed by universal laws such the non negativity of the entropy production, which imposes limitations on the efficiency of thermal machines. Equilibrium Classical Thermodynamics of macroscopic systems got its justification from a microscopic point of view from equlibrium Statistical Mechanics, which basically states that for a system in contact with a heat bath, the probability of a given system being at a certain microstate is given by the Boltzmann factor.

Recently, these ideas have been extended to the description of non equilibrium states of small scale systems for which an average description in terms of ensemble quantities is not sufficient. These extensions has been carried out to different kinds of system, Markovian and non-Markovian, discrete and continuous... giving birth to a new discipline, Stochastic Thermodynamics. In that theory, the Second Law gets replaced by Integral and Detailed Fluctuation Theorems which constrain the probability distribution of quantities as entropy along individual trajectories as well as the average production of entropy.

Stochastic Thermodynamics can be applied to a wide range of systems, being some paradigmatic ones molecular motors in single molecule assays and thermoelectric devices involving single electron transport. On the other hand, one kind of interesting small scale systems are some realizations of the Maxwell demon mental experiment. In particular, there have appeared some models in the literature, which behave as a Maxwell demon by using the equal footing of thermodynamic and information entropy.

This manuscript is organised as follows. In section 2.1 we summarize the main results of Stochastic Thermodynamics, mainly following the approach in [1]. In section 2.2 we remember the definition of a Maxwell demon and present a realization that appeared recently in [9]. In section 2.3 we explain how can one simulate stochastic processes, using the Gillespie algorithm [5]. Results for the stochastic simulation of a toy model and for the previous Maxwell demon realization are presented and discussed in sections 3.1 and 3.2.

2 Theoretical Background

2.1 Stochastic Thermodynamics

2.1.1 Ensemble Stochastic Thermodynamics

We consider a system with discrete non-degenerate states (labeled by the index m) in contact with a single heat and particle reservoir at a temperature T and chemical potential μ. The energy of a

state m is given by ϵ_m and the number of particles in such a state by n_m. The system is also allowed to exchange work with a work source which controls the energy levels $\epsilon(\lambda)$ via a time dependent control parameter $\lambda(t)$.

In the (non)-equilibrium ensemble thermodynamics picture a system has a probability P_m of being in the state m, with $\sum_m P_m = 1$. This probability distribution is not the equilibrium one. In what follows we only consider Markovian dynamics, i.e: the evolution of a system only depends on its current state. In that case the time evolution of the probabilities P_m is controlled by the Markovian master equation:

$$\frac{dP_m}{dt} \equiv \dot{P}_m = \sum_{m'} W_{m,m'} P_{m'} \tag{1}$$

With $W_{m,m'}$ the probability per unit time of making a transition from state m' to m. From the normalization of P_m, one can easily obtain the following property for the transition rates $W_{m,m'}$:

$$\sum_m W_{m,m'} = 0 \tag{2}$$

Which implies $W_{m,m} = -\sum_{m' \neq m} W_{m',m}$. This means that the diagonal elements contain the probability rates of "leaving" a given state. With this (1) is easily understood. The change in the probability of a given state m is the result of a transition from any other state $m' \neq m$ to m minus the probability of jumping from state m to any $m' \neq m$.

In the steady state the system is in equilibrium with the reservoir and the probabilities are the grand canonical equilibrium ones:

$$P_m^{eq} = \exp\left[-\beta(\epsilon_m - \mu n_m - \Omega^{eq})\right] \tag{3}$$

Where $\beta \equiv 1/T$ (we use units where the Boltzmann constant equals 1 and has no dimensions) and Ω^{eq} is the equilibrium grand potential, which is obtained by normalization of P_m^{eq}.

In this work we also impose the more restrictive condition of detailed balance, i.e: at equilibrium the transitions from m to m' and from m' to m are equally likely, thus:

$$W_{m,m'} P_{m'}^{eq} = W_{m',m} P_m^{eq} \tag{4}$$

From (3) and (4) one obtains:

$$\frac{W_{m',m}}{W_{m,m'}} = \exp\left[-\beta(\epsilon_{m'} - \epsilon_m - \mu(n_{m'} - n_m))\right] \tag{5}$$

Now we introduce the state functions (quantities that do not depend on how P_m was achieved):

$$E = \sum_m \epsilon_m P_m \tag{6}$$

$$N = \sum_m n_m P_m \tag{7}$$

$$S = -\sum_m P_m \ln P_m \tag{8}$$

These are the non-equilibrium ensemble-averaged values of the energy, number of particles and entropy of the system (analogous to the familiar equilibrium expressions).

One can obtain a first and second law of stochastic thermodynamics at the ensemble level. We start by taking the time derivative of (6):

$$\dot{E} = \sum_m (\epsilon_m \dot{P}_m + \dot{\epsilon}_m P_m) \equiv \dot{Q} + \dot{W}_{chem} + \dot{W} \tag{9}$$

Where we have identified $\dot{W} = \sum_m \dot{\epsilon}_m P_m$, $\dot{Q} = \sum_m \epsilon_m \dot{P}_m - \dot{W}_{chem}$ and $\dot{W}_{chem} \equiv \mu \dot{N}$. Therefore, work is the result of an energy shift in an occupied state and heat and chemical work the result of transitions between states.

Moving to the second law, one obtains from (8):

$$\dot{S} = \underbrace{\frac{1}{2} \sum_{m,m'} (W_{m,m'} P_{m'} - W_{m',m} P_m) \ln \frac{W_{m',m}}{W_{m,m'}}}_{\dot{S}_e} + \underbrace{\frac{1}{2} \sum_{m,m'} (W_{m,m'} P_{m'} - W_{m',m} P_m) \ln \frac{W_{m,m'} P_{m'}}{W_{m',m} P_m}}_{\dot{S}_{TOT}} \tag{10}$$

The $\dot{S}_e$ term can be seen as an entropy exchange between the system and the reservoir. To see this, first we note that from (9) we can write $\dot{Q} = \sum_m (\epsilon_m - \mu n_m)\dot{P}_m$ and that:

$$\dot{S}_e = \sum_{m,m'} W_{m,m'} P_{m'} \frac{\epsilon_m - \epsilon'_m - \mu(n_m - n_{m'})}{T} = \sum_m \dot{P}_m \frac{\epsilon_m - \mu n_m}{T} = \frac{\dot{Q}}{T} \tag{11}$$

Where we have used (5) in the first equality and (2) and (1) in the second equality. Hence we can identify $\dot{S}_e$ as the entropy exchanged.

Obviously $\dot{S}_{TOT} \geq 0$ so then we identify it as the total entropy production. It is also easy to check that for quasistatic processes, where $P_m(t) = P_m^{eq}(t)$ using (8) we retrieve the result that in that case $\dot{S} = \dot{Q}/T$ and there is no entropy production.

2.1.2 Trajectory Stochastic Thermodynamics

For large systems with a number of particles $\sim N_A \sim 10^{23}$, fluctuations are negligible and the ensemble description is suitable. This is no longer the case for small scale systems, where the measured quantities vary from one experiment to another and fluctuations become important. Of course, if one averages the results for several experiments one expects to retrieve the ensemble averages. Neveretheless, as we explain below, by studying the system at the trajectory level (i.e: studying how the state m of a single system evolves in time) one obtains a deeper formulation of the second law.

We shall imagine a small system and focus on its trajectory on time. The state of the system at a time t will be $m(t)$ and its energy $e = e_{m(t)}(t)$, number of particles $n = n_{m(t)}$ and entropy $s = -\ln P_{m(t)}(t)$. The energy has a double dependence on time. First, it can vary as a result of a change in the state of the system $m(t)$. It can also change as a result of the external driving that changes the energy levels $e_m(t)$ for a fixed m. Note that this is not the case for n because we assume that the number of particle in a given state does not change with time.

For an arbitrary quantity of our system $f_{m(t)}(t)$, we can picture it as a continuous function of t except for the times t^* where there is a transition from state $m(t^* - dt)$ to $m(t^* + dt)$ and using properties of the distributional derivative one obtains:

$$\frac{df_{m(t)}(t)}{dt} = \sum_{t^*} \left[\delta(t - t^*)(f_{m(t^*+dt)}(t) - f_{m(t^*-dt)}(t))\right] + \dot{f}_{m(t)}(t) \tag{12}$$

Where $\delta(t)$ is the Dirac delta distribution and the sum is for the times where there is a transition. For simplicity, from here on we follow the notation in [2]:

$$\frac{df_{m(t)}(t)}{dt} = \sum_m \left[\dot{\delta}^{Kr}_{m,m(t)} f_m(t) + \delta^{Kr}_{m,m(t)} \dot{f}_m(t)\right] \tag{13}$$

Where the first term is the one related with transitions and the second with the driving (δ^{Kr} is the Kronecker's delta). Applying (13) to e one obtains the first law at the trajectory level:

$$\dot{e} = \dot{q} + \dot{w}_{chem} + \dot{w} \tag{14}$$

Where $w = \sum_m \delta^{Kr}_{m,m(t)} \dot{e}_m(t)$ (work), $q = \sum_m \left[\dot{\delta}^{Kr}_{m,m(t)} e_m(t)\right] - \dot{w}_{chem}$ (heat) with $\dot{w}_{chem} = \sum_m \mu n_m \dot{\delta}^{Kr}_{m,m(t)}$ (chemical work). The interpretation is the same as before. Work is related with a shift in the energy of a given level and heat to the transition between levels.

For the second law, using (13) we obtain:

$$\dot{s} = \underbrace{-\sum_m \left[\dot{\delta}^{Kr}_{m,m(t)} \ln \frac{P_m(t)}{P_m^{eq}(t)} + \delta^{Kr}_{m,m(t)} \frac{\dot{P}_m(t)}{P_m(t)}\right]}_{\dot{s}_{TOT}} \underbrace{- \sum_m \dot{\delta}^{Kr}_{m,m(t)} \ln P_m^{eq}(t)}_{\dot{s}_e} \tag{15}$$

Where P_m^{eq} is given by (3). This decomposition was introduced in [3]. First, we see that $\dot{s}_e$ is the usual entropy exchange:

$$\dot{s}_e = -\sum_m \dot{\delta}^{Kr}_{m,m(t)} \ln P_m^{eq}(t) = \frac{1}{T} \sum_m \dot{\delta}^{Kr}_{m,m(t)} [\epsilon_m(t) - \mu n_m] = \frac{\dot{q}}{T} \tag{16}$$

And therefore the remaining part must be the entropy production. What is remarkable now is that $\dot{s}_{TOT}$ is not semipositive defined, and therefore, at the trajectory level we no longer need

$\dot{s}_{TOT} \geq 0$. One can find a more physically meaningful expression for the entropy production. Let $M_{[0,\tau]} = \{m(t_0 \equiv 0), m(t_1), \ldots, m(\tau = t)\}$ be a trajectory followed by the system. This trajectory has a probability given by:

$$\mathcal{P}(M_{[0,\tau]}) = P_0(m_0)P_1(m_1|m_0)P_2(m_2|m_1)\ldots P_t(m_t|m_{t-1}) \tag{17}$$

Where $P_1(m_1|m_0)$ is the probability that at time t_1 we have a transition from m_0 to m_1 and so on. If we consider $t_i - t_{i-1} \equiv dt$ infinitesimal, that probability is simply $W_{m_1,m_0}dt$. The claim is that the entropy production along $M_{[0,\tau]}$ is given by:

$$S_{TOT}(M_{[0,\tau]}) = \ln \frac{\mathcal{P}(M_{[0,\tau]})}{\tilde{\mathcal{P}}(\tilde{M}_{[0,\tau]})} \tag{18}$$

Where $\tilde{M}_{[0,\tau]}$ is the time reversal trajectory. Namely, reversing the time-dependence of the driving while using as starting probabilities for the time reversal experiment ($\tilde{P}_0(m_t)$) the final ones for the direct experiment $P_t(m_t)$:

$$\tilde{\mathcal{P}}(\tilde{M}_{[0,\tau]}) = P_t(m_t)P_t(m_{t-1}|m_t)\ldots P_2(m_1|m_2)P_1(m_0|m_1) \tag{19}$$

One can separate (18) as:

$$S_{TOT}(M_{[0,\tau]}) = \ln \frac{P_0(m_0)}{P_t(m_t)} \frac{\prod_{t=t_1}^{\tau} P_t(m_t|m_{t-1})}{\prod_{t=t_1}^{\tau} P_t(m_{t-1}|m_t)} \tag{20}$$

Using (5) we have $\ln \frac{P_t(m_t|m_{t-1})}{P_t(m_{t-1}|m_t)} = \frac{-q_{m_t,m_{t-1}}}{T}$ and using the properties of ln:

$$S_{TOT}(M_{[0,\tau]}) = -\ln P_t(m_t) - (-\ln P_0(m_0)) + \sum_{t=0}^{t=\tau} -\frac{q_{t+dt,t}}{T} = \Delta S_{sys}(M_{[0,\tau]}) + \frac{-Q(M_{[0,\tau]})}{T} \tag{21}$$

Where $Q(M_{[0,\tau]})$ is the heat exchanged by the system during the trajectory and $\Delta S_{sys}(M_{[0,\tau]})$ is the total entropy change in the system. Equation (18) has a clear meaning. The more likely the forward trajectory is respect to the inverse trajectory, the more irreversible the evolution is, and hence the larger the entropy production is. In the next subsection we introduce and derive some fluctuations theorems for the evolution of a system along trajectories.

2.1.3 Fluctuation Theorems

Each trajectory ($M_{[0,\tau]}$) has a given probability $\mathcal{P}(M_{[0,\tau]})$. Therefore, the probability that there is a given entropy production ($P(S_{TOT})$) is given by the following path integral:

$$P(S_{TOT}) = \int_{M_{[0,\tau]}} \mathcal{D}M_{[0,\tau]} \delta \left(S_{TOT} - \ln \frac{\mathcal{P}(M_{[0,\tau]})}{\tilde{\mathcal{P}}(\tilde{M}_{[0,\tau]})} \right) \mathcal{P}(M_{[0,\tau]}) \tag{22}$$

Where δ is the Dirac delta distribution. By using that $\delta(ax) = \frac{1}{|a|}\delta(x)$, that the jacobian for the transformation from $M_{[0,\tau]}$ to $\tilde{M}_{[0,\tau]}$ is 1 and (18):

$$P(S_{TOT}) = \exp S_{TOT} \int_{\tilde{M}_{[0,\tau]}} \mathcal{D}\tilde{M}_{[0,\tau]} \delta \left(-S_{TOT} - \ln \frac{\tilde{\mathcal{P}}(\tilde{M}_{[0,\tau]})}{\mathcal{P}(M_{[0,\tau]})} \right) \tilde{\mathcal{P}}(\tilde{M}_{[0,\tau]}) = \exp S_{TOT} \tilde{P}(-S_{TOT}) \tag{23}$$

And we obtain:

$$\frac{P(S_{TOT})}{\tilde{P}(-S_{TOT})} = \exp S_{TOT} \tag{24}$$

Which is the Detailed Fluctuation Theorem ([1]). It has a clear meaning, the probability of having an entropy production S_{TOT} in the forward trajectory is exponetially more likely than having a decrease $-S_{TOT}$ in the reverse trajectory.

From (18) we can also derive the so-called Integral Fluctuation Theorem ([3]):

$$\langle \exp(-S_{TOT}) \rangle \equiv \int_{M_{[0,\tau]}} \mathcal{D}M_{[0,\tau]} \mathcal{P}(M_{[0,\tau]}) \cdot \exp \left(-\ln \frac{\mathcal{P}(M_{[0,\tau]})}{\tilde{\mathcal{P}}(\tilde{M}_{[0,\tau]})} \right) = \int_{M_{[0,\tau]}} \mathcal{D}M_{[0,\tau]} \tilde{\mathcal{P}}(\tilde{M}_{[0,\tau]}) = 1 \tag{25}$$

Where we have used the fact that that the probability of any path happening is equal to 1.

It has already been mentioned that in a given trajectory the entropy production can be negative, i.e: $S_{TOT} < 0$. However, by applying the Jensen inequality (i.e: $f(\langle x \rangle) \leq \langle f(x) \rangle$) to the exponential function which is a convex function:

$$1 = \langle \exp(-S_{TOT}) \rangle \geq \exp(-\langle S_{TOT} \rangle) \rightarrow \langle S_{TOT} \rangle \geq 0 \tag{26}$$

Which is nothing but the Second Law of Thermodynamics. Thus, even though a single trajectory can violate the Second Law of Thermodynamics, on average one has that the entropy cannot decrease.

2.2 A realization of a Maxwell Demon

2.2.1 Maxwell Demons

Let us picture two chambers filled with gases at different temperatures and imagine a tiny intelligent creature acting as a gatekeeper between the two chambers (see Figure 1). It lets the fast-moving particles move from the cold to the hot chamber and conversely, the slow-moving particles move from the hot to the cold chamber. In such a way, the system acts as a refrigerator by transferring heat from a cold to a hot chamber and doing so without work involved. Such a thought experiment is known as the Maxwell Demon (see [6]), and it violates the Second Law of Thermodynamics.

Such a device could exist without violating the first law if [8] it was at the same time allowed to write information in a physical memory. In such a way the increase of information entropy compensates the decrease of entropy of the refrigerator, and the Second Law is no longer violated if the total entropy increase is positive. Such a system obviously requires treating the Clausius entropy and the Shannon entropy on equal footing.

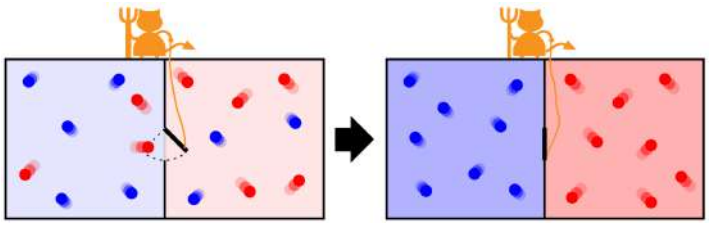


Figure 1: Depiction of a Maxwell Demon. Extracted from [**?**]

2.2.2 The model

Recently, [9], an analytically solvable model that uses the equivalence between Shannon and Clausius entropy has been introduced.

The system under consideration is shown in Figure 2.

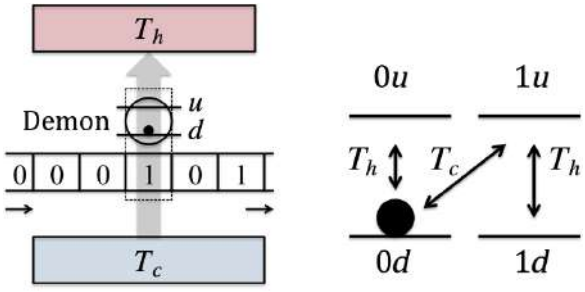


Figure 2: Realization of a Maxwell demon. Extracted from [9]

There are four elements. Two thermal baths, at temperature T_h (hot bath) and T_c (cold bath) with $T_h > T_c$. The demon in this case is a two level system, u and d with an energy difference $\Delta E \equiv E_u - E_d$. Then there is a memory register, consisting of a sequence of bits (with equal energy for the 0 and 1 states) equally spaced along a tape that slides past the demon.

The joint state of the demon and bit can undergo two types of random transitions. First, the state of the demon can change without a change in the state of the bit ($0d \leftrightarrow 0u$ and $1d \leftrightarrow 1u$),

coupled to the T_h bath. We shall call these the intrinsic transitions. The rates verify (5), i.e (we continue with the notation introduced in Section 2.1):

$$\frac{W_{u,d}}{W_{d,u}} = \exp\left(-\beta_h \Delta E\right) \tag{27}$$

As in [9] we parameterize the rates as $W_{u,d} = \gamma(1-\sigma)$ and $W_{d,u} = \gamma(1+\sigma)$ with $\sigma = \tanh\frac{\beta_h \Delta E}{2}$ and γ a parameter that characterizes the rate for the transitions.

The other type of transitions, the corporative transitions, involve a simultanous change in the state of the bit and demon (see Figure 2 right). The allowed transitions are $0d \leftrightarrow 1u$, coupled to the cold bath. Again that rates verify:

$$\frac{W_{1u,0d}}{W_{0d,1u}} = \exp\left(-\beta_c \Delta E\right) \tag{28}$$

And a similar parametrization $W_{1u,0d} = 1-\omega$ and $W_{0d,1u} = 1+\omega$ with $\omega = \tanh\frac{\beta_c \Delta E}{2}$. It is also convenient to introduce:

$$\epsilon = \tanh\frac{(\beta_c - \beta_h)\Delta E}{2} \tag{29}$$

Which controls the temperature difference between the two baths. The value of the incoming bits has probabilities p_0 and p_1. For later convenience we introduce $\delta \equiv p_0 - p_1$.

We first provide an heuristic explanation of how this system behaves as a Maxwell demon. First, assume that $\delta = 1$, i.e: the incoming bits are all 0's. The demon interacts with the incoming bit during a time τ. The initial joint state can be either $0u$ or $0d$. If the final state is also one of these two, it means no neat energy was exchange with the cold bath as every $0d \leftrightarrow 1u$ was compensated by the opposite one. On the other hand, if the joint state after the time τ is $1u$ or $1d$ it means there was an energy transference ΔE from the cold bath to the system. As the demon cannot "keep" absorbing energy, the conservation of energy implies that there is an effective transference of energy from the cold to the hot bath, thus acting as a refrigerator. If the incoming bits were all 1's, the opposite situation takes place. The cold bath absorbs energy from the hot bath. Thus, heuristically an excess of 0s ($\delta > 0$) favors the flow of heat from the cold to the hot bath and an excess of 1s ($\delta < 0$) favors the opposite.

Obviously, that competes with the usual thermodynamic bias due to the temperature difference, so a quantitative description is necessary.

Let p'_0 and p'_1 be the probabilities of the outgoing bits once the demon has reached its periodic state, with $\delta' \equiv p'_0 - p'_1$. $\Phi \equiv p'_1 - p_1 = \frac{\delta - \delta'}{2}$ represents the average production of 1s per interaction interval. As a transition $0 \to 1$ is accompanied by the absorption of energy ΔE from the cold reservoir, the average transfer of energy from the cold to the hot reservoir per interaction interval is given by:

$$Q_{c\to h} = \Phi \Delta E \tag{30}$$

Therefore a positive value of Φ implies that the system behaves as a refrigerator. The Shannon entropy of the incoming bits is given by:

$$S(\delta) = -\frac{1-\delta}{2}\ln\frac{1-\delta}{2} - \frac{1+\delta}{2}\ln\frac{1+\delta}{2} \tag{31}$$

And the outgoing bits the same just changing δ by δ'. Therefore the increase in information due to the process is:

$$\Delta S_B \equiv S(\delta') - S(\delta) = S(\delta - 2\Phi) - S(\delta) \tag{32}$$

A positive value, means the demon writes information in the tape. A negative value implies the demon erases information. By (30) and (32) ΔS_B and $Q_{c\to h}$ are determined if one knows $\Phi(\omega, \sigma, \gamma, \tau, \delta)$. For that, we need to find the periodic steady state of the system.

The statistical state of the demon is $\mathbf{P}^D = \begin{pmatrix} p_u \\ p_d \end{pmatrix}$, that of the incoming bit $\mathbf{P}^B = \begin{pmatrix} p_0 \\ p_1 \end{pmatrix}$ and the joint state $\mathbf{P} = \begin{pmatrix} p_{u0} \\ p_{d0} \\ p_{u1} \\ p_{d1} \end{pmatrix}$. By introducing the following matrix $\mathcal{M} = \begin{pmatrix} p_0 & 0 \\ 0 & p_0 \\ p_1 & 0 \\ 0 & p_1 \end{pmatrix}$, we clearly have $\mathbf{P} = \mathcal{M}\mathbf{P}^D$. The joint probabilities evolve acording a master equation $\dot{\mathbf{P}} = \mathcal{W}\mathbf{P}$, where $\mathcal{W}$ contains the rates of (27) and (28). Solving it, the joint state of the system after the interaction time τ is:

$$\mathbf{P}_\tau = \exp\left(\mathcal{W}\tau\right)\mathcal{M}\mathbf{P}_0^D. \tag{33}$$

Projecting the marginals for the state of the demon after an interaction:

$$\mathbf{P}_\tau^D = \mathcal{P}^D \exp(\mathcal{W}\tau)\mathcal{M}\mathbf{P}_0^D \equiv \mathcal{T}\mathbf{P}_0^D. \tag{34}$$

$\mathcal{T}$ is the transition matrix and $\mathcal{P}^D = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$. As $\mathcal{T}$ is a positive transition matrix, the demon evolves to a periodic state, $\mathbf{P}^{D,ps}$ defined by:

$$\mathcal{T}\mathbf{P}_0^{D,ps} = \mathbf{P}_0^{D,ps} \tag{35}$$

Once the periodic state has been obtained, the probability of the outgoing bits will be:

$$\mathbf{P}_\tau^{B,ps} = \mathcal{P}^B \exp(\mathcal{W}\tau)\mathcal{M}\mathbf{P}_0^{D,ps} \tag{36}$$

With $\mathcal{P}^B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ the matrix that projects out the state of the demon. Now the procedure for an analytical or numerical calculation is clear. First, we find the eigenstate of (35). Then, we use it in (36) and then we simply use that $\Phi = p_1' - p_1$. In the next section we will show some results we obtained by simulating that system numerically. For future comparison we write the analytical result derived using symbolic calculus [10]:

$$\Phi = \frac{\delta - \epsilon}{2}\eta(\Lambda) \tag{37}$$

The explicit form of $\eta(\Lambda)$ is complicated and it can be found in [10]. However, what is relevant is that $\eta(\Lambda) \geq 0$ and therefore, the sign of Φ (that determines if the system behaves as a refrigerator) is fully determined by δ (information bias of the input bits) and ϵ (thermodynamics bias of the temperature difference between the baths).

Now, we show how this system verifies a generalized Clausius inequality. If the interaction time $\tau \to \infty$, the system reaches a steady state, $\bar{\mathbf{P}}$ defined by $\mathcal{W}\bar{\mathbf{P}} = 0$. The result is $\bar{\mathbf{P}} = \frac{1}{\mathcal{N}} \begin{pmatrix} 1 \\ \mu \\ \mu\nu \\ \mu^2\nu \end{pmatrix}$: Where $\mu = \frac{1+\sigma}{1-\sigma}$, $\nu = \frac{1-\omega}{1+\omega}$ and $\mathcal{N} = (1+\mu)(1+\mu\nu)$. This is the product of the marginal distributions $\bar{p}_{ij} = \bar{p}_i^D \bar{p}_j^B$ with $i \in \{0, d\}$ and $j \in \{0, 1\}$ and $\bar{\mathbf{P}}^D = \frac{1}{1+\mu} \begin{pmatrix} 1 \\ \mu \end{pmatrix}$ and $\bar{\mathbf{P}}^B = \frac{1}{1+\mu\nu} \begin{pmatrix} 1 \\ \mu\nu \end{pmatrix}$ Let $\mathbf{P}$ be the joint probabilities at time t. We introduce the Relative Entropy or Kullback divergence ([11]):

$$D(\mathbf{P}||\bar{\mathbf{P}}) = \sum_\alpha p_\alpha \ln \frac{p_\alpha}{\bar{p}_\alpha} \geq 0 \tag{38}$$

Where the sum over α is for the joint states ($\alpha = \{0u, 0d, 1u, 1d\}$). It can be shown ([12]) that $\dot{D}(\mathbf{P}||\bar{\mathbf{P}}) \leq 0$. That means that $\mathbf{P}$ evolves monotonically to $\bar{\mathbf{P}}$. If $\mathbf{P}_0$ and $\mathbf{P}_\tau$ are the joint distributions at the start of the interaction and right after the interaction, then:

$$D(\mathbf{P}_0||\bar{\mathbf{P}}) - D(\mathbf{P}_\tau||\bar{\mathbf{P}}) \geq 0 \tag{39}$$

By expanding and using that $\bar{\mathbf{P}}$ factorizes, we have:

$$\underbrace{\sum_\alpha p_{0,\alpha} \ln p_{0,\alpha}}_{-S_0} \underbrace{- \sum_\alpha p_{0,\alpha} \ln p_{0,\alpha}}_{S_\tau} + \sum_{i=u,d} (p_{\tau,i}^D - p_{0,i}^D) \ln \bar{p}_i^D + \sum_{i=0,1} (p_{\tau,i}^B - p_{0,i}^B) \ln \bar{p}_i^B \geq 0 \tag{40}$$

The third term is 0 by considering the steady periodic state. Here S_0 and S_τ are the joint Shannon entropies of the initial and final states. The joint entropy is usually written as $S = S^D + S^B - I(D;B)$ where S^D is the entropy of the demon, S^B of the bit and $I(D;B)$ is the mutual information [1]. In our case $S_0^D = S_\tau^D$ by periodicity and $I_0(D;B) = 0$. Therefore:

$$\Delta S_B - I_\tau(D;B) + \Phi \ln \frac{\bar{p}_1^B}{\bar{p}_0^B} \geq 0 \tag{41}$$

[1] Let X and Y be two discrete random variables. Let $P_X(x)$ and $P_Y(y)$ be the marginal probabilities and $P_{XY}(x,y)$ be the joint probability. By definition, the mutual information of both variables is $I(X;Y) \equiv \sum_{x\in\mathcal{X}} \sum_{y\in\mathcal{Y}} P_{XY}(x,y) \ln \frac{P_{XY}(x,y)}{P_X(x)P_Y(y)}$

And finally:

$$\Delta S_B + Q_{c\to h}(\beta_h - \beta_c) \geq I_\tau(D;B) \geq 0 \tag{42}$$

This is the generalization of Clausius inequality. In this case, only the sum of the information and thermodynamic entropy is positive. Not only that, but the "total" entropy has a lower bound given by the mutual information.

From (37) and (42) we can plot the phase diagram of the system (see Figure 3). First, from (37) it is obvious that for $\delta > \epsilon$ the system acts as a refrigerator. In that case, from (42) $\Delta S_B > 0$ and the system writes information on the memory. Thus, the cost of refrigeration is the increase in entropy on the memory. This corresponds to the darker region in Figure 3. For $\delta < \epsilon$ the system no longer acts as a refrigerator and it can erase/write information depending on the parameters. In Figure 3 it is shown for certain values of τ the boundary of such behaviours. To the right (for a given τ) the system acts an eraser. To the left it writtes information. When $\tau \to \infty$ the boundary is the line $\delta = -\epsilon$. To the left of that (white region) it neither acts as a refrigerator nor as an eraser.

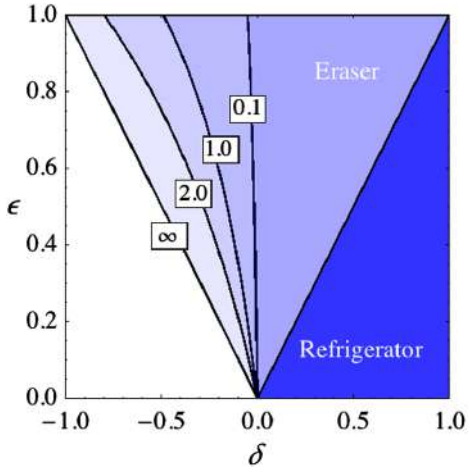


Figure 3: Different modes of operation for $\gamma = 1$ and $\omega = 0.5$. Extracted from [9].

2.3 Simulating Trajectories: Gillespie Algorithm

In this section we explain two methods to simulate the stochastic trajectory of a given system. For simplicity we assume that the probabilities for the different states at $t = 0$ are the the stationary ones $\mathbf{\Pi}$, such that equation (1) in matrix form becomes $\dot{\mathbf{\Pi}} = \mathcal{W}\mathbf{\Pi} = 0$. From the fact that $det(\mathcal{W}) = 0$ we have an infinite number of solutions to the previous equation, until we impose normalization.

2.3.1 Naive Approach

The straightforward approach is the following. We start at time $t_0 \equiv 0$. From the probabilities Π_m, we build the cumulative probabilities. Then, we generate a random number $Y \sim U(0,1)$ (uniformly distributed between 0 and 1) and we pick the state whose cumulative probability is immediately above the random number. This is implemented in Python by the function *random.choices(list,probabilities,N)* which returns a list with N random elements from *list* distributed according to *probabilities*.

Once we have the initial state $m_0 \equiv m(t_0 = 0)$ we divide the time τ in small intervals, $t_{i+1} = t_i + dt$. Let $m_i \equiv m(t_i)$. There are two possibilities. First, there can be no transition in the interval dt. That has a probability $P_{stay} = 1 - dt \cdot \sum_{m' \neq m_i} W_{m',m_i} = 1 + W_{m_i,m_i} \cdot dt$. Then, there can be transitions to any other state m' with probabilities $P(m'|m_i) = W_{m',m_i}dt$. So then we pick the state at t_{i+1} according to these probabilities.

2.3.2 Gillespie Algorithm

For small dt the previous procedure is inefficient because at each step the most likely event is not having a transition. This is solved by the Gillespie Algorithm [5], which was historically introduced to simulate coupled chemical reactions.

Let m be the state of the system at time t and let $\{\alpha\}$ be the possible transitions from state m to any other state. The key is to consider the joint probability $P(\tau, \alpha_0)dt$ that the system remains a time τ at the state m and then in a time dt experiences the transition $\alpha_0 \in \{\alpha\}$. Let us divide τ in K infinitesimal intervals, then we have:

$$P(\tau,\alpha_0)dt = \underbrace{\left(1-\sum_\alpha P_\alpha\right)\dots\left(1-\sum_\alpha P_\alpha\right)}_{Ktimes} P(\alpha_0) = \lim_{K\to\infty}\left(1-\frac{\tau}{K}\sum_\alpha W_\alpha\right)^K W_\alpha dt = \exp\left(-\tau\sum_\alpha W_\alpha\right)W_\alpha dt \tag{43}$$

Here W_α is the transition rate for the α transition and $P_\alpha = dtW_\alpha$. Multiplying and dividing by $\lambda \equiv \sum_\alpha W_\alpha$:

$$P(\tau,\alpha_0)dt = \lambda \exp\left(-\lambda\tau\right)\cdot P(\alpha_0|\{\alpha\}) \tag{44}$$

Where $P(\alpha_0|\{\alpha\}) = \frac{W_{\alpha_0}}{\lambda}$ is the probability that the transition that takes place is α_0 given that there was some transition taking place. From (44) we see how does the algorithm work. First, we generate a random number $Y \sim Exp(\lambda)$ (following an exponential distribution). That gives us the time τ required for the next transition to happen. Then, we pick the particular transition α_0 with probabilities W_α.

3 Results and discussion

3.1 A toy model

In this section we introduce a toy model that allows us to numerically test the theoretical results introduced in section 2.1 with a code we prepared to implement Gillespie Algorithm.

The model consists of (see Figure 4) a system with three energy levels ϵ_i, i = 0,1,2 (with $\epsilon_0 \equiv 0$), no external driving and coupled to three thermal baths at temperatures T_i, $i = a, b, c$. Each thermal bath couples a different transition, T_a the transition $\epsilon_0 \leftrightarrow \epsilon_1$, T_b the transition $\epsilon_1 \leftrightarrow \epsilon_2$ and T_c the transition $\epsilon_0 \leftrightarrow \epsilon_2$.

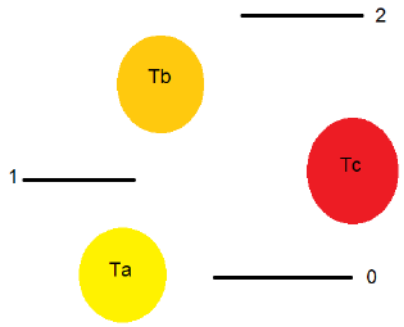


Figure 4: Schematic of our toy model.

We parameterize the transition rates $W^r_{j,i}$ as:

$$W^r_{j,i} = \Gamma_0 \exp -\frac{\epsilon_j - \epsilon_i}{2T_r} \tag{45}$$

Where r is the bath controlling the transition $i \leftrightarrow j$. It is easy to check that these rates verify (4), where Γ_0 is a characteristic time scale. This system is interesting, because for a suitable set of parameters ϵ_i and T_j the system acts as a refrigerator, pulling heat from the cold bath and transferring it to a hotter bath. Notice that this is allowed and it does not violate the Second Law because we have three baths.

First, we check the validity of the IFT (equation (25)). In Figure 5 we plot the result of averaging the exponential of minus the entropy production, $\exp\left(-S_{TOT}\right)$, for different number of trajectories

(N_{tra}). Note that here we have more than one bath, so we use the generalization of (21) (see [2] for a detailed description of systems with multiple reservoirs):

$$S_{TOT} = \Delta S_{sys} - \sum_r \frac{Q_r}{T_r} \tag{46}$$

Where Q_r is the heat exchanged by bath r.

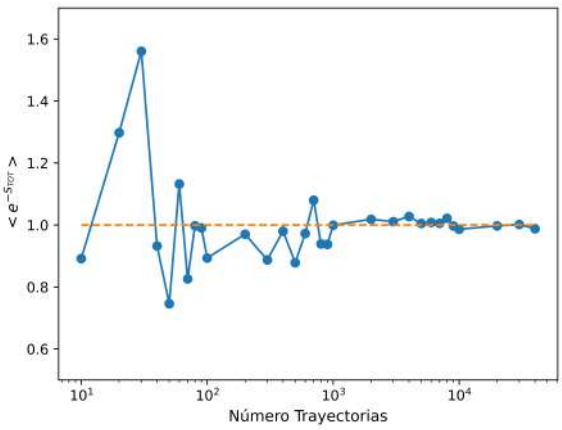


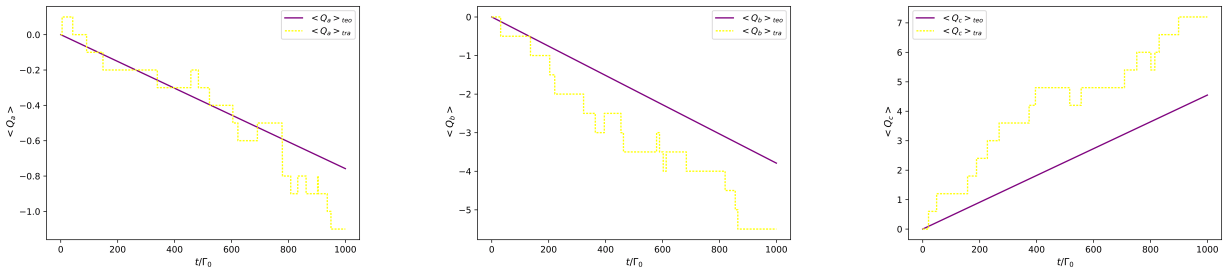
Figure 5: Blue Dots: numerical average value of $\exp(-S_{TOT})$ for different number of individual trajectories. Orange line: ensemble analytical result (=1). Parameters: $\epsilon_1 = 1$, $\epsilon_2 = 6$, $T_a = 1$, $T_b = 2$, $T_c = 20$, $\Gamma_0 = 1$, $dt = 0.01\Gamma_0$ (time interval between points in our simulation) with $\tau = 1000\Gamma_0$ (total time for a single trajectory).

Now (see Figures (6)-(8)), we check that the (as we start the simulation with the steady probabilities Π given by $\mathcal{W}\Pi = 0$), the average heat rate exchanged by a bath is constant given a set of parameters when a sufficiently large number of trajectories is taken, and that this rate coincides with the emsemble ones, found by solving the eigenvalue equation for Π (either analytically or numerically) and using that:

$$\dot{Q}_a = \epsilon_1 \left(W_{1,0}\Pi_0 - W_{0,1}\Pi_1\right) \tag{47}$$

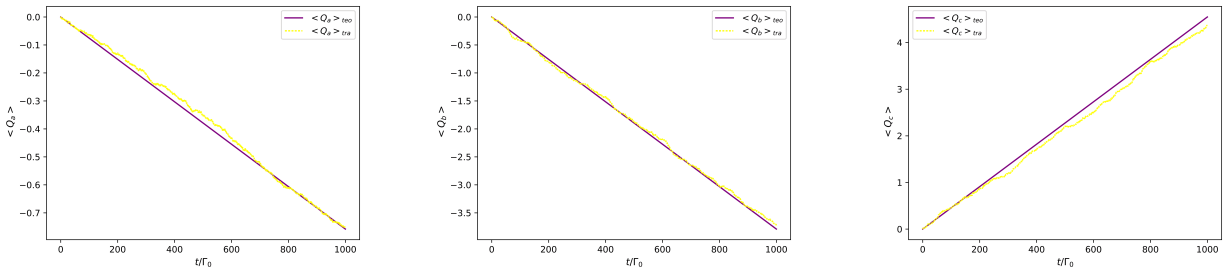
$$\dot{Q}_b = (\epsilon_2 - \epsilon_1)\left(W_{2,1}\Pi_1 - W_{1,2}\Pi_2\right) \tag{48}$$

$$\dot{Q}_c = \epsilon_2 \left(W_{2,0}\Pi_0 - W_{0,2}\Pi_2\right) \tag{49}$$



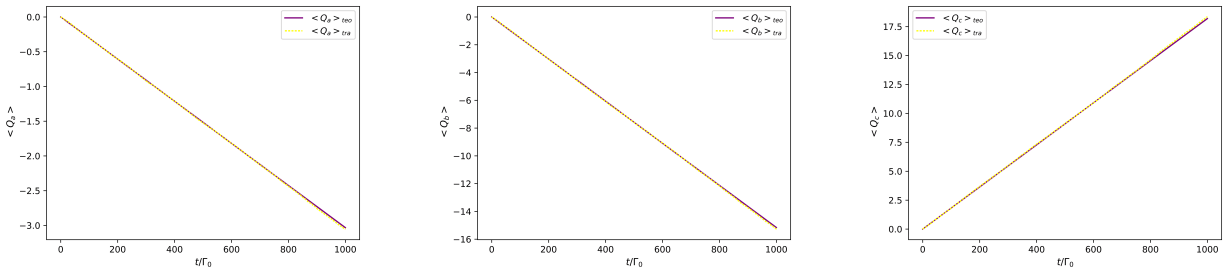
(a) Bath a (b) Bath b (c) Bath c

Figure 6: Parameters as in Figure 5. Results for $N_{tra} = 10^1$



(a) Bath a (b) Bath b (c) Bath c

Figure 7: Parameters as in Figure 5. Results for $N_{tra} = 10^3$



(a) Bath a (b) Bath b (c) Bath c

Figure 8: Parameters as in Figure 5. Results for $N_{tra} = 10^5$

In all three Figures, the yellow line corresponds to the simulation over single trajectories, and the purple line to the exact ensemble result (47)-(49). We clearly see that for $N_{tra} \sim 10^5$ the average over single trajectories coincides with the ensemble results.

Finally, we also plot in Figure 9 the distribution of the different values for the entropy production S_{TOT}, by ploting the cumulative distribution $F(S) \equiv \sum_{S_{TOT} \leq S} P(S_{TOT})$ as a function of S_{TOT}. From (24) it can be shown (see [4]) that $F(S_{TOT}) \sim exp(S_{TOT})$ for $S_{TOT} \leq 0$:

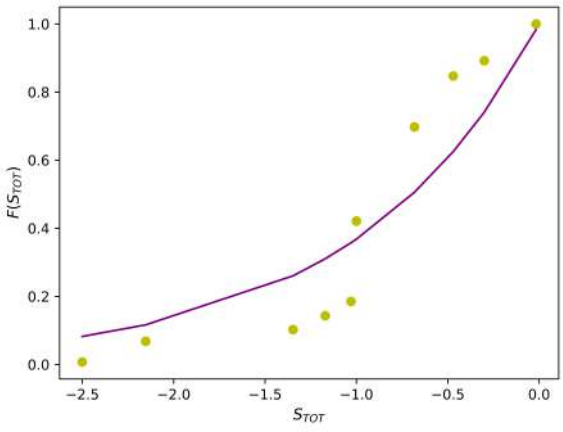


Figure 9: We plot the cumulative probability distribution $F(S_{TOT})$ of entropy production S_{TOT}. Parameters as in Figure 5 and $N_{tra} = 10^5$. Yellow dots: Simulation results. Purple line: $\sim exp(S_{TOT})$.

We approximately see the expected behaviour. Note that $P(S_{TOT}) = \frac{N_{S_{TOT}}}{N_{tra}}$ where $N_{S_{TOT}}$ is the number of trajectories with a entropy production S_{TOT} and that in our case S_{TOT} can only take discrete values.

3.2 Simulating the Maxwell's demon model

In this section we present and discuss the results we obtained by simulating the model in section 2.2.2 at the trajectory level using the Gillespie Algorithm. In this simulation we start with an arbitrary distribution for the incoming bits (parametrized by δ) and for the initial distribution of the demon we chose the steady periodic one $\mathbf{P}_0^{D,ps}$.

First, in Figures 10 and 11 we show the average heat exchanged by the cold bath as a function of the number of individual trajectories taken into account for two different set of parameters.

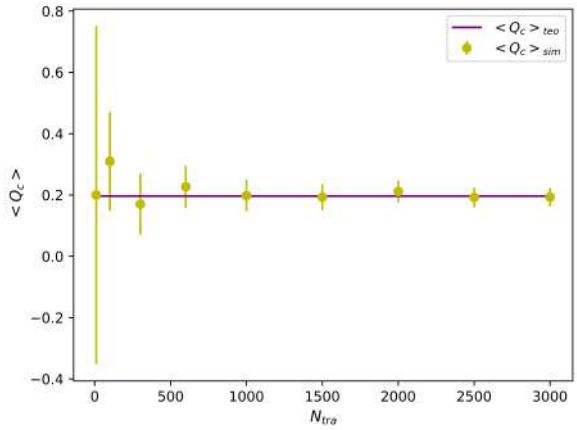


Figure 10: Average heat exchanged ($\langle Q_c \rangle$) by the cold bath as a function of the number of individual trajectories (N_{tra}) taken. Parameters: $\gamma = 1.0, \omega = 0.5, \epsilon = 0.7, \tau = 2.0$ and $\delta = 0.2$. The yellow dots are the simulation results, whereas the purple line comes from equations (30) and (37).

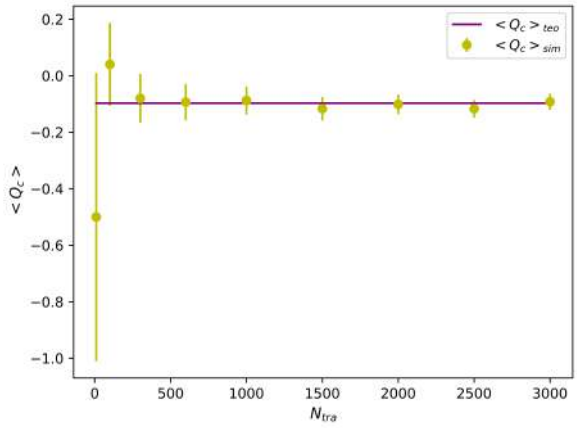


Figure 11: Same as in the previous Figure but with the following Parameters: $\gamma = 1.0, \omega = 0.5, \epsilon = 0.2, \tau = 2.0$ and $\delta = 0.7$.

We clearly see that the simulation reproduces the ensemble result for the heat exchanged by the cold bath whenever we take $N_{tra} \gtrsim 10^3$. We also recover that for $\delta < \epsilon$ (Figure 10) the information bias is not strong enough to overcome the thermodynamic bias and the cold bath absorbs heat from the hot bath (in particular in that case the system acts an eraser). On the other hand, when $\delta > \epsilon$ (Figure 11) the information bias is strong enough so that the system acts as a refrigerator at the cost of writting information on the memory.

As in the previous section, here we also check the validity of the Integral Fluctuation Theorem for the model. The result is shown in Figure 12:

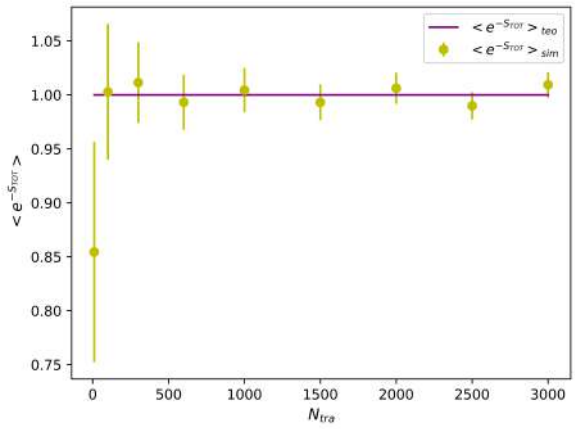


Figure 12: Yellow Dots: numerical average value of $\exp(-S_{TOT})$ for different number of individual trajectories. Purple line: analytical result (=1). Same parameters as in Figure 11.

S_{TOT} is calculate as follows. From section 2.1.2 we know that the total change in entropy in the system ΔS along a trajectory is:

$$\Delta S = -\ln p_\alpha + \ln p_{\alpha_0} \tag{50}$$

Where α is the final joint state, α_0 the initial joint state of the demon and bit and p_α are the corresponding probabilities. S_{TOT} is then calculated as:

$$S_{TOT} = -\ln p_\alpha + \ln p_{\alpha_0} - \frac{q_c}{T_c} - \frac{q_h}{T_h} \tag{51}$$

Where q_c and q_h are the heat exchange by the cold and hot bath during a single trajectory. We stress that S_{TOT} has contributions from both the Shannon and Clausius entropy.

Finally, if we go back to (42), we see the entropy production is lower bounded by the mutual information of the final probability distributions of the demon and bit. In Figure 13 we check how much does this improve the usual well-known lower bound (equals to 0).

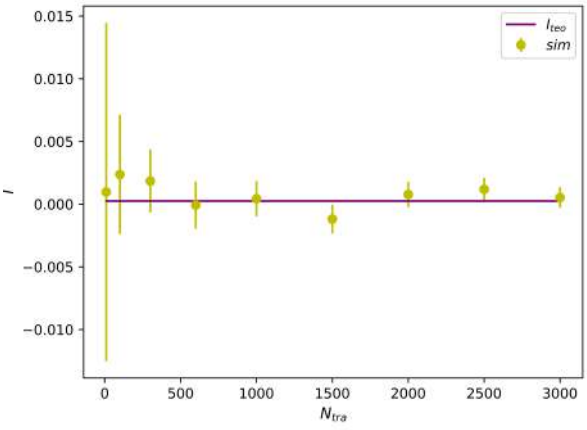


Figure 13: Yellow Dots: average values for different number of individual trajectories for the mutual information. Purple line: exact result for the mutual information of the demon and bit. Same parameters as in Figure 11

.

We clearly see the mutual information does not difer much from 0, so (at least for the particular parameters we chose) we can forget about the mutual information term in (42). This in agreement with the numerical findings in [9]. As we start from the periodic distribution for the demon, $\mathbf{P}_0^{D,ps}$, after an interaction time τ (by construction) the distribution of the demon will be the same. For

the outgoing distribution of the bits $\mathbf{P}_\tau^B$ we simply use (36) and for the joint final state $\mathbf{P}_\tau$ we simply use (33) with $\mathbf{P}_0^{D,ps}$. For all three distributions we use the numerical results by finding the autovectors and so numerically. Then the average mutual information is given by:

$$I_\tau(D;B) = \sum_{i=u,d} \sum_{j=0,1} p_{\tau,ij} \ln \frac{p_{\tau,ij}}{p_{0,i}^{D,ps} p_{\tau,j}^B} \tag{52}$$

The mutual information at the trajectory level is simply defined as $\ln \frac{p_{\tau,ij}}{p_{0,i}^{D,ps} p_{\tau,j}^B}$ where i, j are the states after an interaction of the Demon and bit.

4 Conclusions

We have learned how one can extend the classical Laws of Thermodynamics to microscopic systems driven out of equilibrium. At the ensemble level, one retrieves laws analogous to the macroscopic Classical Thermodynamics ones. However, at the individual trajectory level, one finds deviations from those well known laws. For instance, a single trajectory can have a negative entropy production. In addition to retrieve the ensemble results when averaging respect to several trajectories, a series of Fluctuation Theorems emerge, which characterize these averages beyond the general ensemble results and also impose restrictions on quantities as the probability of a certain entropy being produced along a single trajectory.

One can check these theorems numerically by using the Gillespie Algorithm, which is optimized to simulate stochastic trajectories. We have implemented such an algorithm, and used it to study two systems. We have started with a "toy model" consisting of three energy levels and three heat baths. We have verified the IFT theorem, the heat rate production and the distribution of the entropy production along single trajectories.

We have explained the idea of Maxwell Demons and a recent realization that appeared in an article, that uses the equivalence between Clausius and Shannon entropy. We have simulated that system at the trajectory level, verifying again the IFT, the analytical results for the heat exchange and also studying quantitatively an improvement of the lower bound to the total entropy production that appeared in the aforementioned article, finding, as stated there, that such a bound does not really depart much from 0.

As a final conclusion, we have learned the fundamentals of Stochastic Thermodynamics, which is a recent theory applicable to thermal machines and biological systems to name a few, and we have focused on numerically studying an information system, using the Gillespie Algorithm. Some work can be done beyond this introductory work. For instance, one idea can be to explore and quantify in which cases (values for parameters such the interaction time τ and energy differences ϵ) does the lower bound production to the total entropy given by the mutual information in the Maxwell demon realization is optimal, in such a way that in those cases it represents a good improvement respect to the positivity restriction.

Acknowledgments

First, I would like to thank my supervisor, Gonzalo Manzano, for planning this project, solving my doubts and providing material beyond the covered here. Second, I would like to thank the organizers/speakers of the SURF program, specially Manuel Matias for organising most of the activities. Last but not least, I would to mention the good times shared with the other fellows, Gala, Jacob, Leyla, Marina and Nacho.

This work was supported by the SURF@IFISC fellowship.

References

[1] M. Esposito and C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).

[2] C. Van den Broeck, C. and M. Esposito. Phys. A **418**, 6–16 (2015)

[3] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).

[4] U. Seifert, Rep. Prog. Phys. **75**, 126001 (2012)

[5] Daniel T. Gillespie Journal of Computational Physics, **22**, 403-434, (1976).

[6] J. C. Maxwell, Theory of Heat (Longmans, London, 1871)

[7] https://mappingignorance.org/2020/10/05/maxwells-demon-and-the-relationship-between-information-and-irreversibility/

[8] H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).

[9] D. Mandal, H. T Quan and C.Jarzynski, C. Phys. Rev. Lett. **111**, 030602 (2013)

[10] D. Mandal, H. T Quan and C.Jarzynski, C. Supplemental Material http://link.aps.org/ supplemental/10.1103/PhysRevLett.111.030602

[11] Michael A. Nielsen and Isaac L. Chuang. Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press (2011)

[12] N. G. van Kampen, Stochastic Processes in Physics and Chemistry Elsevier, Amsterdam, (2007)

Semiconductor laser network dynamics with inhomogeneous elements

Author: Jacob Rifà Estol
Supervisors: Miguel C. Soriano, Apostolos Argyris
Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB)
Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain

Abstract

This report presents a numerical investigation into the effects of diverse couplings between a network of lasers with inhomogeneous elements. A comprehensive model, based on the connectivity matrix, is developed to analyze specific scenarios. The study focuses on the multinodal star all-optical network topology. Notably, the results reveal the occurrence of spontaneous partial synchronization for specific coupling values. By examining these phenomena, we aim to deepen our understanding of semiconductor laser network dynamics and their behavior in the presence of inhomogeneous elements. The results contribute to the advancement of optical network design and synchronization analysis in complex systems.

1 Introduction

Chaotic signals have become integral components in various applications, including optical communication systems [1], high-speed photonic data processing [1, 2] amongst other applications. As the study of network dynamics has gained momentum, the collective behavior of networks composed of mutually coupled semiconductor lasers has become a subject of great interest. Understanding the intricate dynamics of these networks, particularly in the presence of inhomogeneous elements and diverse coupling mechanisms, offers new avenues for exploring complex dynamical systems and their applications [3]. Previous research has paved the way for investigating the dynamics of semiconductor laser networks. Notably, studies have examined the sensitivity analysis of star optical networks based on mutually coupled semiconductor lasers [4]. By exploring the robustness and stability of such systems, valuable insights have been gained regarding the effects of varying parameters on network dynamics. Experimental investigations have also contributed to our understanding of semiconductor laser networks in coupled configurations. These studies have observed synchrony in networks of coupled semiconductor lasers, shedding light on the underlying mechanisms and opening up possibilities for synchronization-based systems [5].

The emergence of chimera states in coupled semiconductor laser networks has been a topic of recent interest. These states, characterized by spatially localized coexistence of coherent and incoherent laser dynamics, have been observed without the presence of multistability [6]. This finding challenges previous assumptions and deepens our understanding of the intricate spatiotemporal patterns that arise in such networks. Motivated by these significant developments, our research aims to further investigate the dynamics of semiconductor laser networks with inhomogeneous elements and diverse coupling schemes. By considering different coupling strategies, we aim to uncover the rich behaviors exhibited by these networks. To accomplish this, we perform numerical simulations and analyze in detail the obtained results. By systematically varying coupling strengths, time delays, and network topology, we aim to elucidate the effects of inhomogeneities and diverse coupling mechanisms on network dynamics. Our research endeavors to provide new insights into the emergence of collective behaviors, identifying optimal coupling strategies that provide partial or total synchronization that is needed for different applications.

In summary, this report presents an investigation into the dynamics of semiconductor laser networks with inhomogeneous elements and diverse coupling schemes. Building upon previous studies, our research aims to deepen our understanding of the collective behaviors exhibited by these networks. Our findings are intended to contribute to the ongoing progress in the field of semiconductor laser network dynamics.

2 Theoretical model

Our model consists of a generalisation of previously studied network topologies. Due to computational limitations we have used $N = 23$ lasers varying the coupling strength and studying separately two different topologies. The mathematical model that describes the behaviour of networks with an arbitrary topology of the semiconductor lasers is based on the Lang Kobayashi equations [4]. Being E_j and N_j the complex electrical field and the number of carriers of the j-th laser respectively, these equations read as

$$\frac{dE_j(t)}{dt} = i\Delta\omega_j E_j(t) + \frac{1}{2}\left(G_j(t) - \frac{1}{t_{ph}}\right)(1+ia)E_j(t) + \sum_{i=1}^{n} k_{ij}E_i(t-\tau_{ij})e^{-i\omega_0\tau_{ij}} + \sqrt{D}\xi_j(t) \quad (1)$$

$$\frac{dN_j(t)}{dt} = \frac{I_j}{e} - \frac{N_j(t)}{t_{s,j}}G_j(t)|E_j(t)|^2 \quad (2)$$

being

$$G_j(t) = \frac{g_n(N_j(t) - N_0)}{1 + s|E_j(t)|^2} \quad (3)$$

where $a = 3$ is the linewidth enhancement factor, $g_n = 1.2 \cdot 10^{-5}$ ns^{-1} is the differential gain coefficient, t_{ph} is the photon lifetime, $t_s = 1.54$ ns is the carrier lifetime, $s = 5 \cdot 10^{-7}$ is the saturation gain coefficient, I_j is the injection current for each laser, $\Delta\omega_j$ is the detuning frequency with respect to the free-running reference frequency $\omega_0 = 2\pi c\lambda_0$, with $\lambda_0 = 1550$ nm. $\xi_j(t)$ is a complex Gaussian white noise for each laser, with $D = 5 \cdot 10^{-5}$ ns^{-1} being the noise strength. Through the connectivity matrix k_{ij} and the delay matrix τ_{ij} we can implement the different topologies we want to study.

The integration of the delayed differential equations is performed with a Runge-Kutta of 4th order method using an integration step $dt = 0.8$ ps. For large values of the coupling strength k, this method does not converge to a physical solution of the problem, so a detailed analysis of the data was needed in order to guarantee the validity of the results.

In order to check the functioning of our model we study the particular case where 22 lasers are connected to one single laser, the HUB laser, (Fig.1) through a connectivity matrix given by $k_{iH} = k$ $\forall i \neq H$, $k_{Hj} = \beta k$ $\forall j \neq H$ and rest of the elements 0.

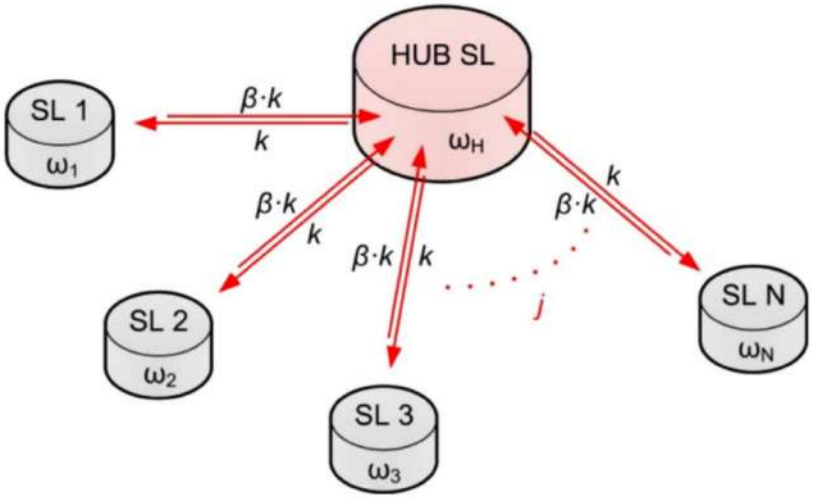


Figure 1: Schematic diagram of our network topology. Source: [4]

β is the asymmetry parameter representing an attenuation factor in the injected fields in the hub. By employing this approach, the collective injection received by the hub laser becomes independent of the number of lasers involved. For our purpose we take $\beta = 0.5$ and investigate the results for different values of $k \in [0, 70]$ ns^{-1}, which is between the limits of the converging regime. For the delay matrix we take $\tau_{iH} = \tau_{Hj} = \tau = 2.9$ ns $\forall i, j \neq H$ and rest of the elements are 0. Taking the delay time smaller facilitates the surpassing of the transient state, so that the total integration time can be smaller making it more viable computationally. Half of the non-HUB lasers have negative detuning frequencies uniformly distributed in steps of 3 GHz and centered at -20 GHz, the other

Laser #	Detuning Frequency (GHz)	Laser #	Detuning Frequency (GHz)
1	-35	12	5
2	-32	13	8
3	-29	14	11
4	-26	15	14
5	-23	16	17
6	-20	17	20
7	-17	18	23
8	-14	19	26
9	-11	20	29
10	-8	21	32
11	-5	22	35

Table 1: Distribution of detuning frequencies for the different lasers.

half follow the same structure but centered at 20 GHz (see Tab.1).
We study separately the case where all lasers have the same photon lifetime $t_{ph} = 2$ ps and the case where the photon lifetimes vary between $t_{ph} \in [1.9, 2.1]$ ps. This variation affects significantly the spontaneous synchronization of the lasers.

3 Results

3.1 Constant t_{ph}

The time evolution of $|E_j(t)|^2$ presents a transient state before transitioning into a chaotic behavior, which constitutes a significant subject of interest for our study. However, the integration of the delayed differential equations encountered computational limitations, restricting the analysis to a duration of only 300 ns. This timeframe allows us to observe the transient state and the fully developed dynamics, as evidenced in Fig.2. Despite the relatively short length of the simulated timeframe with respect to the length of the transient, the findings remain compelling.

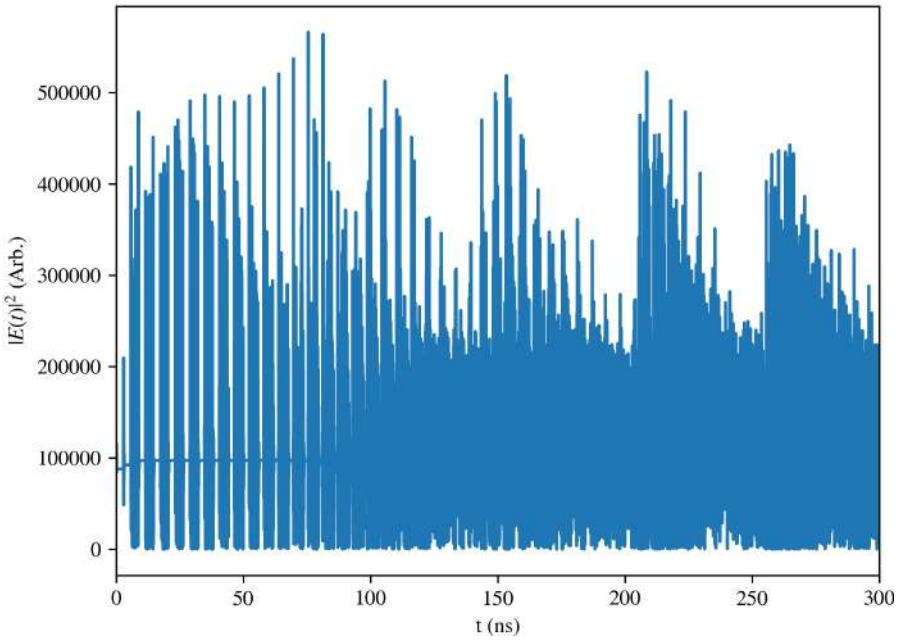


Figure 2: Time evolution of the electrical field for the 4th laser for $\beta = 0.5$ and $k = 50$ ns^{-1} for $I = 25$ mA

During the transient phase, a periodic state emerges with a period close to the delay time $\tau = 2.9$ ns, foreshadowing the subsequent appearance of a larger periodic cycle lasting approximately 80 ns. The interplay between these periodicities and the ensuing chaotic behavior calls our attention,

raising essential questions about the underlying dynamics. Furthermore, even within the transient regime, some elements of chaos are already discernible, indicating the presence of a complex and intricate system. This observation provides a valuable opportunity to examine the underlying mechanisms contributing to the emergence of chaos amidst the transient state.
Despite the challenges posed by the limited integration duration, the signals under study exhibit characteristics that warrant further investigation. The coexistence of periodic and chaotic behavior in $|E_j(t)|^2$ highlights the intricacies of the system, underscoring the need for deeper analysis to comprehend its behavior fully.
In order to comprehensively investigate the synchronization of the lasers, a critical step involves computing the correlation matrix for each laser within a specific time interval, ensuring that the transient state has already been surpassed. This essential data analysis technique enables us to gauge the degree of coherence and alignment among the lasers. By quantifying the interrelation between their emission patterns, we gain valuable insights into the collective behavior of the laser array. The correlation matrix is obtained through a precise mathematical expression (4) tailored to capture the intricate dynamics of the lasers. This matrix serves as a powerful tool in revealing underlying patterns and synchrony within the system, aiding us in identifying any potential correlations or dependencies that might exist. By delving into the synchronized behavior of the lasers, we can pave the way for enhanced laser applications, advanced optical technologies, and deeper understanding of complex dynamical systems.

$$C_{ij} = \frac{\left\langle (P_i(t) - < P_i(t) >) \cdot \left(\left\langle P_i(t) - < P_i(t) > \right) \right\rangle}{\sqrt{\left\langle (P_i(t) - < P_i(t) >)^2 \cdot (P_i(t) - < P_i(t) >)^2 \right\rangle}} \tag{4}$$

where $P_i(t) \propto |E_i(t)|^2$ is the optical power.
We start by showing the zero-lag correlation matrix for some intermediate values of the coupling, which exhibit partial synchronization. In Fig.3 we observe that the lasers with the most negative detuning (lasers 1 to 10) are almost fully synchronized for $k = 45$ ns^{-1}.

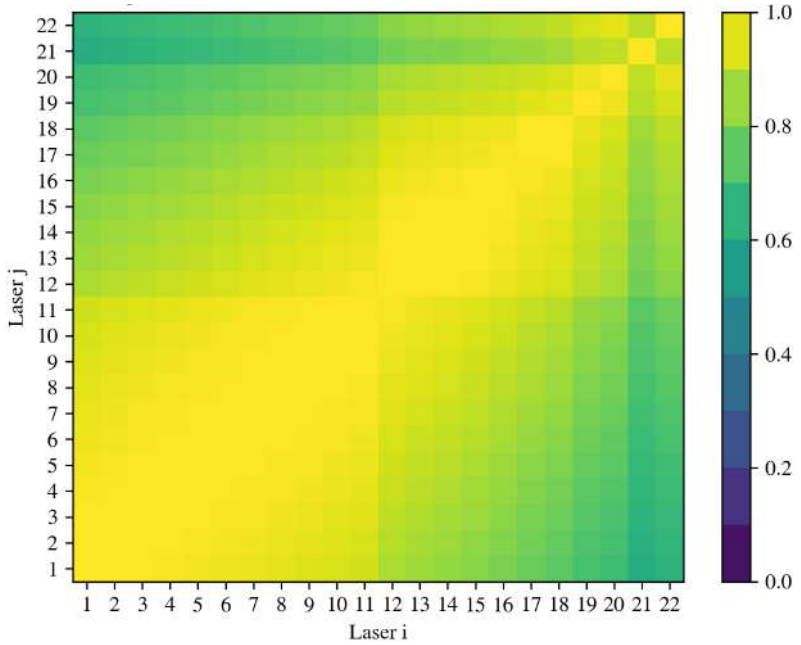


Figure 3: Cross-correlation mapping of the numerically built network with 22 lasers and 1 hub for $\beta = 0.5$ and $k = 45$ ns^{-1}

It becomes evident that the cross-correlation mapping displays two distinct regions with large correlation values, precisely centered at the lasers that have -20 GHz and 20 GHz detuning, respectively. A noticeable asymmetry is observed, predominantly favoring the synchronization of the negative frequencies, which aligns with our expectations based on the system's characteristics related to the linewidth enhancement factor [7]. This asymmetrical pattern remains consistent across nearly all values of k, underscoring its robustness and significance. Moreover, the compelling effect of this asymmetry is not limited to the cross-mapping alone; it also prominently manifests in the optical spectra depicted in Fig.5. These spectra offer further corroborative evidence of the phenomenon,

emphasizing its presence and its potential impact on the overall laser dynamics. This compelling discovery enhances our understanding of the intricacies at play in the system, shedding light on the interplay between detuning frequency distributions and laser synchronization, thus opening new avenues for research and applications in the realm of optical sciences. However, we cannot take any strong conclusions from this point since for large k, $t \in [280, 300]$ ns is still not far away in time from the transient state. Despite this limitation, it is interesting to notice that spontaneous synchronization may appear for certain coupling strengths (Fig.4), even though these states are short-lived in the transient state.

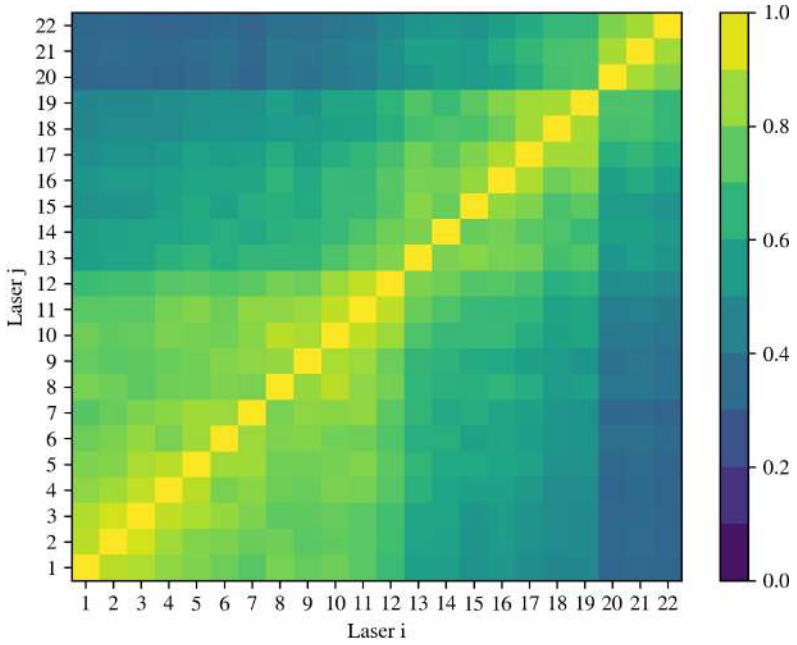


Figure 4: Cross-correlation mapping of the numerically built network with 22 lasers and 1 hub for $\beta = 0.5$ and $k = 45$ ns^{-1} at the transient state

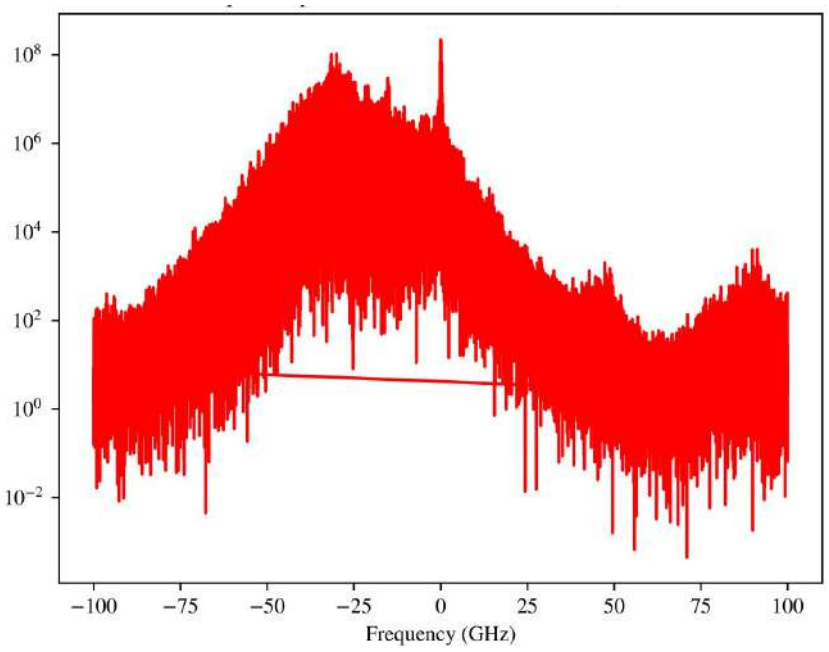


Figure 5: Optical spectrum of the 4th laser for $\beta = 0.5$ and $k = 50$ ns^{-1}

For $k = 65$ ns^{-1} we observe that all lasers have reached full synchronization (Fig.6).

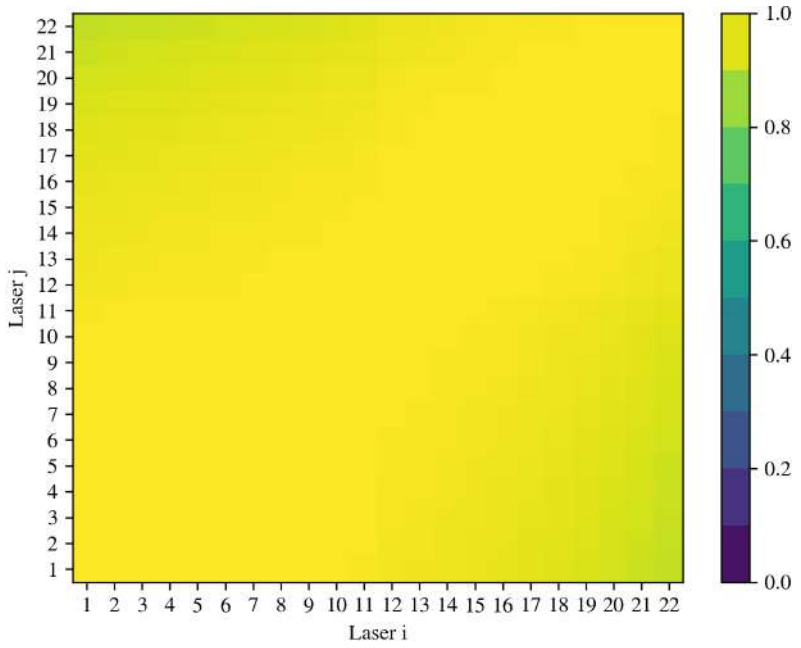


Figure 6: Cross-correlation mapping of the numerically built network with 22 lasers and 1 hub for $\beta = 0.5$ and $k = 65$ ns^{-1}

Under such strong coupling conditions the network emits chaotic dynamics. The high correlation indicates chaotic synchronization, a very useful attribute for chaos-encrypted communications. An effective indicator of chaos' presence is the bifurcation diagram (Fig.7), which is obtained by taking the relative maxima and minima of the optical emission's amplitude, in a certain window of time. The lasers' amplitude oscillates rapidly, in multiple GHz. In order to unravel a possible periodic oscillation with period ~ 10 ps, we perform a smothering of the data up to this value and then look for extremes of the time evolution of $|E(t)|^2$. We observe an almost complete chaotic behaviour for all values of k. We do notice a small periodic regime for some values of k, however, we would have to look closer to this zone in order to analyze it properly, which is beyond the goals of the project.

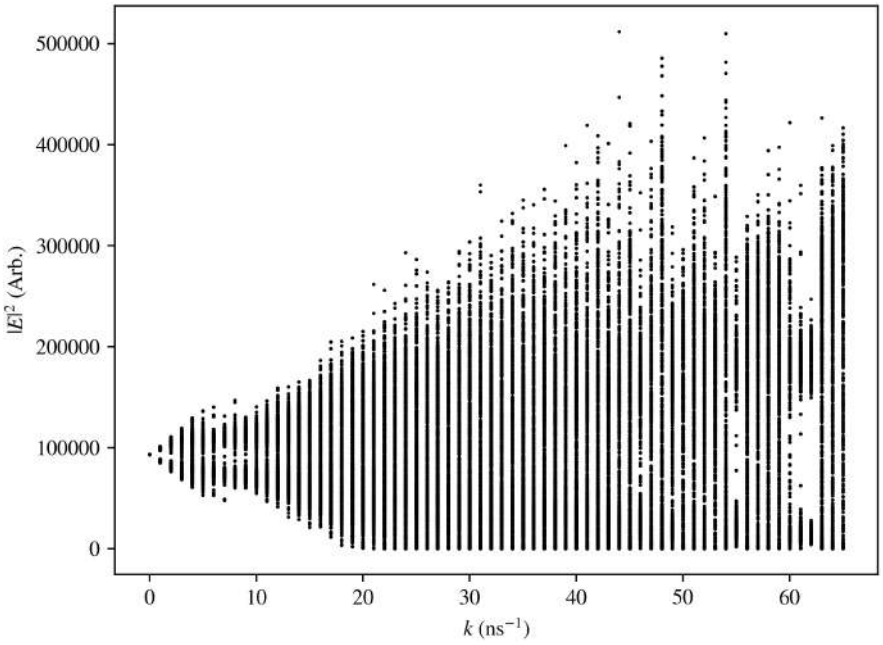


Figure 7: Bifurcation diagram for the 22nd laser

3.1.1 Threshold current

The threshold current, denoted as I_{th}, plays a pivotal role in semiconductor lasers' emission as it defines the minimum current necessary for generating coherent light through stimulated emission. This fundamental value demarcates the transition from a non-lasing state to a lasing one, wherein the laser emits coherent light. To delve deeper into the behavior of lasers near the threshold value, which is approximately $I_{th} = 17.6$ mA, researchers find great interest in studying the transition from the non-lasing to the lasing mode, alongside investigating the laser's efficiency, stability, and various other properties. Although this analysis does not constitute the primary focus of this particular paper, its relevance cannot be overlooked.
Examining Fig.8, we can readily distinguish that the behavior in proximity to the threshold current differs significantly from the scenario where $I = 25$ mA $> I_{th}$, as depicted in Fig.2. Notably, at $t = 300$ ns, it is evident that the emission is still bounded to transient states. This prolongation can be attributed to the relatively low external injection of current, which requires more time to disengage from the initial conditions. By meticulously investigating the laser's performance close to the threshold, we gain valuable insights into the intricate dynamics governing its transition from a non-lasing to a lasing regime. Moreover, comprehending the efficiency and stability under these conditions contributes to a comprehensive understanding of the semiconductor laser's operational principles and capabilities. The ability to precisely control the laser's behavior near the threshold current is of great importance for numerous cutting-edge technologies, such as optical communications, medical devices, and various industrial applications. In these aspects, further research is needed, since this analysis is beyond the goals of our project.

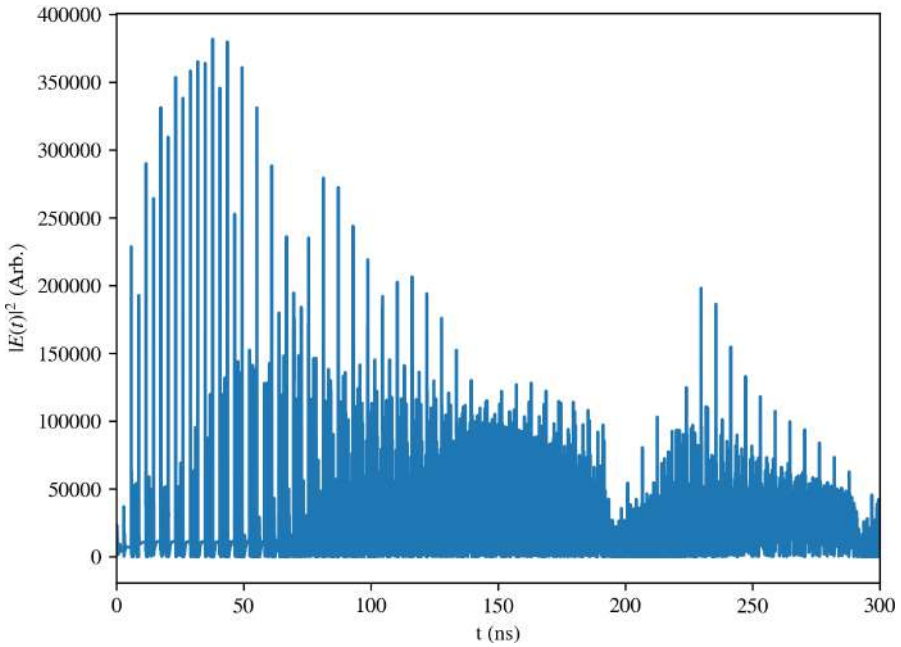


Figure 8: Time evolution of the electrical field for 4-th laser for $\beta = 0.5$ and $k = 50$ ns^{-1} for $I = 17.8$ mA

Comparing Fig.8 and Fig.2, we also observe a conspicuous difference in the typical value of $|E(t)|^2$ when contrasted with that displayed in Fig.2. This variation is not surprising, as lower values were expected in the vicinity of the threshold current.

3.2 Variation of t_{ph}

Upon delving further into the effects of changing the photon lifetime t_{ph} on the laser system, intriguing insights emerge. As expected, the altered t_{ph} brings about an inherent inhomogeneity, influencing the rate of energy loss in each laser with $\frac{dE_j}{dt} \propto -\frac{1}{t_{ph}}$. This variation might naturally lead one to anticipate a diminished level of partial synchronization compared to the homogeneous t_{ph} scenario observed in Fig.3. The outcome depicted in Fig.9 showcases a more localized pattern of partial synchronization.
To interpret this phenomenon, we consider the impact of varying t_{ph} values on individual lasers.

Notably, those lasers endowed with smaller t_{ph} exhibit a faster rate of energy loss. Consequently, these early stages of energy depletion seem to exert a more dominant influence on the overall synchronization dynamics, leading to the observed pronounced focus of partial synchronization.
This observation presents an avenue for further exploration of the complex interplay between photon lifetime and the emergence of partial synchronization. Understanding the mechanisms that govern this behavior can potentially provide valuable insights into optimizing laser systems and harnessing their synchronization properties for diverse applications, ranging from communication networks to precision metrology and beyond. As we continue to unravel the intricacies of this dynamic phenomenon, this study paves the way for advancements in laser technology and synchronization dynamics, with implications extending across various fields of science and engineering.

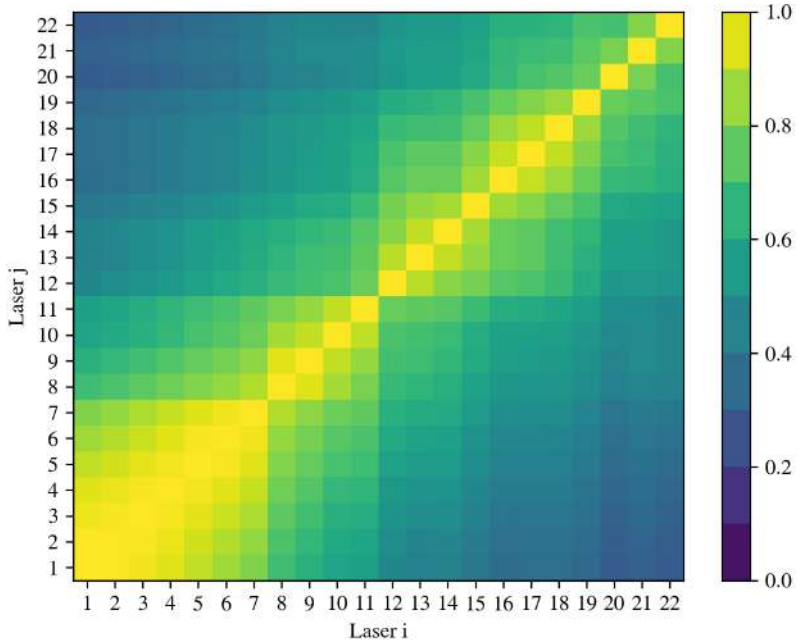


Figure 9: Cross-correlation mapping of the numerically built network with 22 lasers and 1 hub for $\beta = 0.5$ and $k = 45$ ns^{-1} for the variation of t_{ph}

4 Conclusions

We developed a comprehensive numerical investigation into the dynamics of semiconductor laser networks with inhomogeneous elements and diverse coupling schemes. By studying the behavior of a network of mutually coupled lasers, several observations can be made.
The behavior of the network strongly depends on the coupling strength k. For small k, the system exhibits unsynchronized chaotic behavior due to the network's size and complexity. As k increases, the lasers undergo a progressive synchronization process, leading to coordinated chaotic behavior. This synchronization phenomenon holds great promise for a multitude of applications [8].
The detuning frequency also plays a significant role in the synchronization process. We observed that lasers with negative detuning converge faster towards synchronization compared to positively detuned lasers, as evidenced by the cross-correlation mappings. This finding suggests the possibility of tuning the detuning frequency to optimize the synchronization speed for specific applications.
The introduction of inhomogeneities, particularly varying the photon lifetime t_{ph}, poses challenges to achieving synchronization. However, even in the presence of these inhomogeneities, for sufficiently large k, synchronization is eventually attained. This resilience to inhomogeneities expands the potential practicality of semiconductor laser networks.
These findings hold important implications for various applications in different fields. In communication systems, the observed chaotic behavior and synchronization phenomena in laser networks offer opportunities for secure and efficient data transmission. Chaotic signals can enhance encryption techniques, providing resistance to eavesdropping and improving the security of optical communication systems. For high-speed data processing, the chaotic and synchronized behavior of coupled lasers can be harnessed for tasks such as signal modulation and encryption. This can lead to faster data processing and transmission rates, enabling advancements in high-speed data communication. In sensing and metrology, the synchronized behavior of lasers in a network can

enhance the sensitivity and accuracy of sensors. The coordinated response of the lasers to external perturbations can enable precise measurements in fields like environmental monitoring, biomedical sensing, and industrial metrology.
Moreover, the study of inhomogeneous elements and diverse coupling schemes in laser networks contributes to the understanding of nonlinear dynamics and complex systems. The observed chaotic behavior and synchronization phenomena provide valuable insights into the collective behaviors of complex systems, aiding researchers in modeling and analyzing various real-world systems with intricate dynamics.
Additionally, the findings of this research have implications for optical computing, where the synchronized behavior of lasers in a network can be utilized for parallel computing and solving complex computational tasks. Optical computing offers the potential for higher processing speeds and energy efficiency compared to traditional electronic computing methods.
Furthermore, the study of synchronization in laser networks assists in network design and optimization. Understanding the effects of different network topologies and coupling strengths on synchronization can lead to the development of more robust and efficient optical networks for various applications.
In summary, the findings of this research open up possibilities for applying chaotic synchronization and collective behaviors of semiconductor laser networks in a wide range of applications, including communication systems, high-speed data processing, sensing, nonlinear dynamics, optical computing, and network optimization. As further investigations continue, the potential of laser networks in complex systems and their real-world applications is expected to grow, paving the way for advancements in technology and scientific understanding. The continuous exploration of semiconductor laser networks and their behavior promises to revolutionize multiple industries and contribute to the advancement of modern science and technology.

Acknowledgments

I am deeply thankful to my project tutors, Miguel C. Soriano and Apostolos Argyris, as well as the IFISC, for their invaluable guidance and support throughout this project. Their expertise and resources have been instrumental in its success, and I am grateful for the opportunity to work with them, which has enriched my academic journey significantly.
This work was supported by the SURF@IFISC fellowship.

References

[1] G. D. Van Wiggeren and R. Roy, “Communications with chaotic lasers”, Science, Vol. 279, pp. 1198–1200, (1998). https://doi.org/10.1007/978-3-642-04117-4_25

[2] D. Brunner, M. C. Soriano, C.R. Mirasso and I. Fischer, ”Parallel photonic information processing at gigabyte per second data rates using transient states”, Nature Communications, Vol.4, art. 1364, (2013). https://doi.org/10.1038/ncomms2368

[3] L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, “Information processing using a single dynamical node as complex system”, Nature Communications, Vol.2, p.468, (2011). https://doi.org/10.1038/ncomms1476

[4] M. Bourmpos, A. Argyris and D. Syvridis, “Sensitivity analysis of a star optical network based on mutually coupled semiconductor lasers”, Journal of Lightwave Technology, Vol. 30, pp. 2618-2424, (2012). https://doi.org/10.1109/JLT.2012.2204860.

[5] A. Argyris, M. Bourmpos, and D. Syvridis, ”Experimental synchrony of semiconductor lasers in coupled networks”, Optics Express 24, 5600-5614, (2016). https://doi.org/10.1364/OE.24.005600

[6] A. Röhm, F. Böhm and K. Lüdge, ”Small chimera states without multistability in a globally delay-coupled network of four lasers”. Physical Review Letters, 042204, (2016). https://doi.org/10.48550/arXiv.1606.07685

[7] J. Ohtsubo, ”Semiconductor Lasers: Stability, instability and chaos”, Springer, Vol.111, (2013). https://doi.org/10.1007/978-3-540-72650-0

[8] M. C. Soriano, J. Garca-Ojalvo, C. R. Mirasso and I. Fischer, ”Complex photonics: Dynamics and applications of delay-coupled semiconductors lasers”, Reviews of Modern Physics, 85(1), 421, (2013). https://doi.org/10.1103/RevModPhys.85.421

Compartmental models and their application to study phytopathologies

Gala Tomás-Portalés, Àlex Giménez-Romero, Manuel A. Matías
Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB)
Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain

Abstract

Pierce's Disease (PD) poses a significant threat to global viticulture, particularly in coastal and Mediterranean regions. Our study reveals the critical interplay of PD dynamics with climate factors, highlighting the emergence of new risk areas as temperatures rise. Furthermore, we emphasise the pivotal role of vectors in disease transmission. To safeguard vineyards worldwide, our research underscores the need for ongoing surveillance, adaptive strategies, and proactive measures to mitigate the evolving challenges posed by PD, ensuring the resilience and sustainability of the wine industry.

1 Introduction

Vector-borne diseases are caused by infectious agents transmitted by living organisms, called vectors, frequently insects. These diseases pose a significant threat to global human health, giving rise to illnesses such as malaria, yellow fever, and dengue, among others. Human diseases transmitted by vectors account for over 17% of all human infectious diseases, resulting in millions of cases and more than 700 000 deaths annually [1]. Besides humans, plants are also susceptible to bacterial and viral diseases transmitted by vectors, a scenario that is central to our study.

Pierce's Disease (PD) is a deadly vector-borne disease that originated in the Americas and was firstly reported in the 1880s [2]. This devastating malady primarily affects grapevines, Fig. 1(a), posing a significant threat to vineyards and wine production worldwide. It is caused by the bacterium Xylella fastidiosa (Xf), which infiltrates grapevines, gradually weakening them and often leading to their demise. The transmission of PD is facilitated by the Philaenus spumarius (Ps), a meadow spittlebug that feeds on xylem sap, unwittingly aiding in the disease's proliferation (Fig. 1(b)). Furthermore, PD is a thermal-sensitive disease, with temperature playing a pivotal role in its development and spread.

(a) **(b)**

Figure 1: (a) Visible effects of the Pierce's Disease in a grapevine leaf, caused by a strain of the bacteria Xylella fastidiosa [5]; (b) The meadow spittlebug, Philaenus spumarius, vector of PD [6].

The narrative of Pierce's Disease has transcended its American origins. In recent years, PD has crossed oceans and established its presence in new territories. Despite the rigorous quarantine protocols meticulously crafted to safeguard the wine industry, it has encroached upon vineyards nestled

on the Mediterranean isle of Majorca, Spain [2], especially noteworthy in our context. This finding, alongside the detection of PD in Taiwan among others, has raised concerns about its possible spread to continental Europe and other wine-producing regions worldwide. The consequences of such a scenario would be far-reaching. The global wine industry, with a farm value of production exceeding $3 billion annually, faces an existential threat. Moreover, the economic burden of PD is already substantial, with an estimated cost of approximately $100 million annually in California alone [5]. The stakes are high, and the need to unravel and predict the complex dynamics of PD, its interactions with Xf and Ps, and the role of temperature in its proliferation has never been more critical.

This work aims to shed light on these intricate factors. By employing dynamic epidemiological models and incorporating climate variables, we seek to uncover the geographic regions where PD could become endemic. Our goal is to assess the risk of its establishment, persistence, and subsequent epidemic development. The models take into account the limited knowledge about PD vectors in various wine-growing regions globally, offering a heuristic approach to understand the situation previously described and provide valuable insights in order to help safeguarding the future of the global wine industry.

2 Theoretical model

2.1 SIR compartmental model

The Susceptible-Infectious-Recovered (SIR) compartmental model is a classic, yet enduringly relevant framework in epidemiology, which has extensively been employed to describe the dynamics of infectious diseases and understand how they spread [4]. As the name suggests, this model divides our population into three distinct classes:

- Susceptibles (S): individuals inside this compartment are susceptible to the disease because they have not encountered the pathogen or have not yet been exposed to it; i.e., they are capable of becoming infected.
- Infectives (I): this compartment includes subjects who are currently infected with the disease and can potentially transmit it to the susceptibles.
- Removed class (R): individuals who have recovered from the disease and are assumed to have developed immunity, as well as those who are isolated until recovery or death, and thus are no longer susceptible to reinfection.

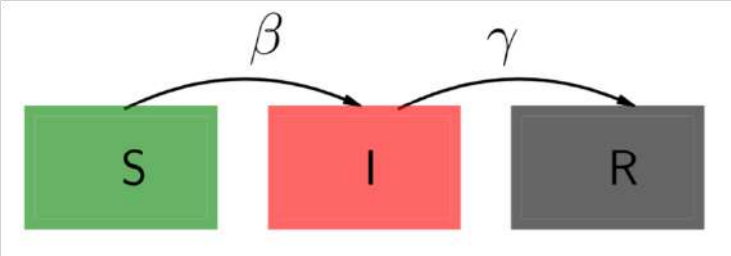


Figure 2: The progression of individuals among different compartments of the SIR model schematically represented [2].

The dynamics of the SIR model are governed by a set of ordinary differential equations (ODEs) that describe how individuals transition between these compartments over time (Fig. 2). Key parameters, such as the infection rate β and removal rate of infectives γ, influence the flow of the subjects through these sections:

$$\begin{aligned} \dot{S} &= -\beta \frac{SI}{N}, \\ \dot{I} &= \beta \frac{SI}{N} - \gamma I, \\ \dot{R} &= \gamma I, \end{aligned} \tag{1}$$

where $\frac{1}{\beta}$ and $\frac{1}{\gamma}$ represent the characteristic time inside the compartment (being $\beta, \gamma = cte > 0$).

The assumptions made about the transmission of the infection and incubation period are crucial in any model; these are reflected in the terms in the equations and the parameters [4]. Here, in the SIR model, we are assuming that: (i) The incubation period is short enough to be negligible; e.i., a susceptible who contracts the disease is infective right away. (ii) We consider the various classes as uniformly mixed; that is, each pair of individuals has equal probability of coming into contact with one another. This is an important assumption and in many situations does not hold as in most sexually transmitted diseases (STDs). (iii) We assume that the total size of the population, N, is taken to be constant. This fact can easily be seen by adding all terms in Eq. (1):

$$\dot{S} + \dot{I} + \dot{R} = 0 \quad \Rightarrow \quad S + I + R = N. \tag{2}$$

Nevertheless, the SIR compartmental model can be adapted to study a wide range of infectious illnesses, including vector-borne diseases, by incorporating additional compartments to represent hosts and vectors. This extension allows researchers to investigate the interplay between these populations and the disease's propagation. Moreover, this type of model is valuable for exploring various factors affecting disease dynamics, such as temperature, vector behaviour, and host susceptibility; as we will discuss in the following subsections.

On the other hand, a key question in any epidemic situation is, given β and γ, whether the infection will spread or not [4]. From the second expression in Eq. (1):

$$\dot{I}(0) = I(0)(\beta\tfrac{S(0)}{N} - \gamma) \begin{cases} > 0 \\ < 0 \end{cases} \quad if \quad \tfrac{S(0)}{N} \begin{cases} > \rho \\ < \rho \end{cases}, \quad \rho = \tfrac{\gamma}{\beta}. \tag{3}$$

And since, from the first expression in Eq. (1), $\dot{S} \leq 0$, $S \leq S(0)$ we have, if $\frac{S(0)}{N} \leq \rho$,

$$\dot{I} = I(\beta\tfrac{S}{N} - \gamma) \leq 0 \quad \forall t \geq 0, \tag{4}$$

in which case $I(0) > I(t) \to 0$ as $t \to \infty$ and so the infection dies out; that is, no epidemic can occur. Conversely, if $\frac{S(0)}{N} > \rho$ then $I(t)$ initially increases and we have an epidemic (meaning that $I(t) > I(0)$ for some $t > 0$). We thus have a threshold phenomenon, and taking into consideration that generally the population at the initial instant is made up mostly of susceptible individuals, $\frac{S(0)}{N} \approx 1$: if $1 > \rho$ there is an epidemic while if $1 < \rho$ there is not, being ρ the critical parameter. Henceforth, we will write:

$$R_0 = \frac{\beta}{\gamma} = \frac{1}{\rho}, \tag{5}$$

where R_0 is the basic reproduction rate of the infection, that is, the number of secondary infections produced by one primary infection in a wholly susceptible population. If more than one secondary infection is produced from one primary infection, that is, $R_0 > 1$, clearly an epidemic ensues. The whole question of thresholds in epidemics is obviously important. The definition and derivation or computation of the basic reproduction rate is crucial, but it can be quite complicated, as we can see exemplified in Eq. (12) of Sec. 2.3.

Finally, the maximum growth rate of the epidemic, relevant for an estimation of the risk of establishment [1], occurs when $S(0) \approx N$ (just as we have previously assumed), and is approximated by the (linearised) differential equation from (1):

$$\dot{I} \approx \beta I - \gamma I = \gamma I \left(\frac{\beta}{\gamma} - 1\right) = \gamma I(R_0 - 1), \tag{6}$$

where we have assumed the initial conditions: $S(0) \approx N$, $I(0) \approx 0$ and $R(0) = 0$. This linear differential equation can be integrated exactly:

$$I(t) = I(0)e^{\gamma(R_0-1)t}, \tag{7}$$

which will be of great help in a future when we try to introduce the climatic factor to our model.

2.2 Thermal-sensitive compartmental model

As presented in [2], this particular model addresses the intricate relationship between climatic conditions and the risk of PD establishment in vineyards. Although the biological reality involves vector-mediated transmission from host to host, for practical modelling purposes, we employ a SIR framework, Eq. (1), representing direct host-to-host communication. However, to capture the influence of temperature on disease dynamics, we introduce critical elements derived from direct measurements.

On the one hand, experimental data reveals that the probability of plant infection post-inoculation is temperature-dependent. Specifically, as temperature accumulates, there exists a threshold beyond which the likelihood of plant infection significantly increases (red curve, Fig. 3). On the other hand, a complementary temperature-related phenomenon comes into play – this is the so-called winter curing. In colder temperatures, the bacteria residing within infected plants are adversely affected, resulting in a probability of recovery. This healing likelihood is influenced by the cumulative cold exposure experienced by the plants during the winter months (blue curve, Fig. 3).

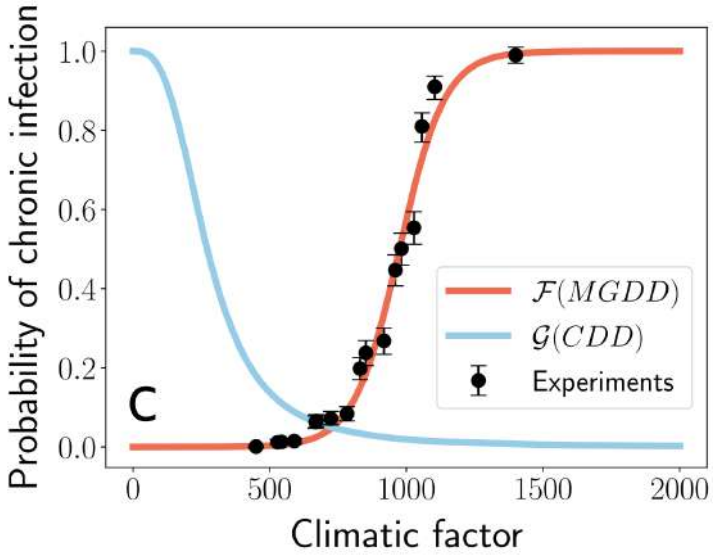


Figure 3: Nonlinear relationship between $\mathcal{F}(MGDD)$ (red line) and $\mathcal{G}(CDD)$ (blue line) and the likelihood of developing chronic infections [2].

The temperature-dependent infection probability and the winter curing effect are essential for understanding the seasonal dynamics of PD and play a pivotal role in our model. In fact, multiplied together, they give place to the cumulative probability of chronic infection, a crucial term represented as:

$$\Pi(t) = \mathcal{F}(MGDD)\mathcal{G}(CDD). \tag{8}$$

So, having built up an initial SIR framework using the exponential approximation, Eq. (7), and taking into account the thermal-sensitive dependence, we can depict the number of infected plants at time t as the product of two crucial terms, Eq. (9). Firstly, $I(t-1)$ encompasses the previous infected plant population, serving as a foundation, and we multiply it by an exponential factor, which represents the new inoculations. And secondly, we introduce the climatic factor Π, which is inherently temperature-dependent. Together, the transmission and climatic layers offer a comprehensive depiction of disease dynamics, combining historical infections, new introductions, and the profound influence of temperature on the system.

$$I(t) = I(t-1)e^{\gamma(R_0-1)}\Pi(t). \tag{9}$$

It is worth highlighting the term R_0 within the Eq. (9), i.e., the basic reproduction number, to which we all are familiar after COVID-19. As stated previously, the value of this element, Eq. (5), is fundamental in epidemiology, as it helps ascertain whether an epidemic is likely to occur. If R_0 is less than 1, an epidemic is improbable, while an R_0 greater than 1 indicates the potential for an

epidemic.

In addition, we also obtained the PD risk index, r, in order to compute the epidemic-risk maps, Eq. (10). In this equation, r implicitly delimits three different cases: (i) non-risk zones where $r \leq$ ≤ -0.1, and the number of infected plants decreases; (ii) transition areas where $-0.1 < r \leq 0.1$, and (iii) an epidemic risk-zone where $r > 0.1$ and PD can theoretically become established and produce an outbreak — the number of infected plants increases exponentially (see [2] for further details).

$$r = max\left\{\frac{log(I(\mathcal{T}))/I(t_0)}{\gamma(R_0 - 1)\mathcal{T}}, -1\right\}, \tag{10}$$

where $\mathcal{T}$ represents the period of time in which our simulation takes place, e.g., 5 years.

2.3 In-depth compartmental model

To address the issue of temperature dependency in disease dynamics, the previous model utilised some approximations in order to obtain Eq. (9). The aim of doing this is to simplify the model because, in the context of analyzing disease spread on a global scale, considering temperature-dependent parameters like β becomes exceedingly complex. Given that the values of these parameters can vary significantly with temperature, the approximation was employed to facilitate broader epidemiological investigations. It's important to note that when applying these approximations, the model offers an overview of disease potential on a global scale without requiring precise parameter values for specific locations.

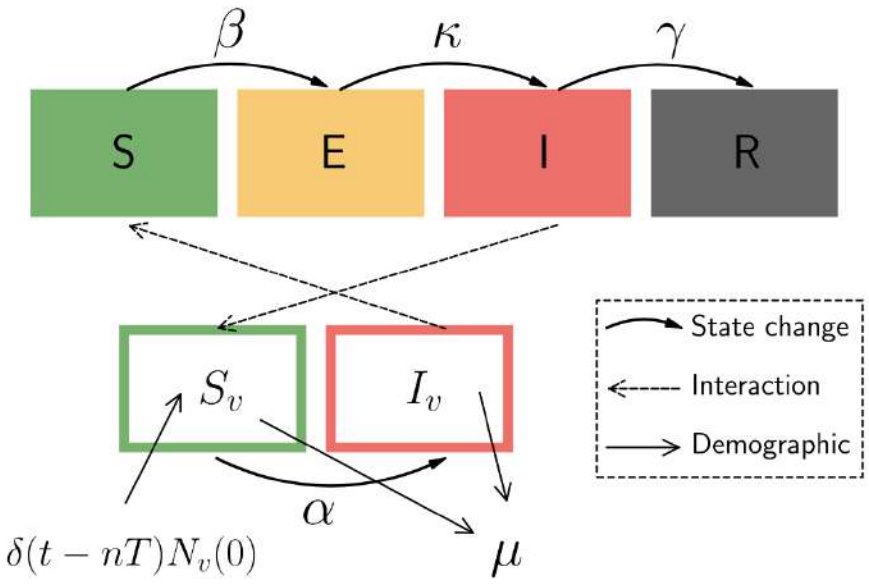


Figure 4: The progression of individuals among different compartments of the host-vector model schematically represented [3].

However, in the current model [3], our objective is to construct a highly precise representation of the disease dynamics at a local level. To achieve this, we incorporate various critical processes known to play a significant role in the disease propagation (Fig. 4). These include the introduction of vectors (in addition to an exposed compartment for the hosts, which represent individuals infected but not yet infectious), vector-to-host and host-to-vector transmission processes, as well as realistic vector dynamics that account for their birth and death rates. All these factors are taken into consideration by the following set of differential equations:

$$
\begin{aligned}
\dot{S}_H &= -\frac{\beta S_H I_v}{N_H}, \\
\dot{I}_H &= \frac{\beta S_H I_v}{N_H} - \kappa E_H, \\
\dot{E}_H &= \kappa E_H - \gamma I_H, \\
\dot{R}_H &= \gamma I_H, \\
\dot{S}_v &= N_v(0) \sum_{n=1}^{\infty} \delta(t - nT) - \frac{\alpha S_v I_H}{N_H} - \mu S_v, \\
\dot{I}_v &= \frac{\alpha S_v I_H}{N_H} - \mu I_v.
\end{aligned}
\tag{11}
$$

Regarding the seasonal dynamics of vectors, we assume that new adults emerge synchronously each year in fields being all susceptible. This is represented by the term $N_v(0)\sum_{n=1}^{\infty}\delta(t-nT)$, where $T = 1$ year is the period and $\delta(t-nT)$ is the Dirac delta function and basically implements a yearly pulse of new vectors at a certain moment in the year. Vectors are removed (die, move to herbaceous vegetation and other non-host trees, exit the field, etc.) at a given rate μ, which we consider identical for susceptible and infected vectors.

Finally, the basic reproduction number of this new model is given by [1]:

$$
R_0 = \frac{\beta\alpha}{\mu\gamma}\frac{S_H(0)}{N_H^2}\frac{N_v(0)}{\mu\tau}(1 - e^{-\mu\tau}). \tag{12}
$$

2.4 In-depth thermal-sensitive compartmental model

To reconcile the two modelling approaches, we aim to merge them effectively. While the previous temperature dependent model served as a useful approximation for global-level assessments, the in-depth epidemiological model offers a comprehensive framework for local-level analyses. By combining these approaches, we seek to incorporate the temperature-dependency of disease dynamics into our precise local model.

It is essential to highlight that the thermal-sensitive model primarily predicts the risk of disease establishment (r). This means it can indicate whether an outbreak is likely to occur but may not provide precise insights into the extent of the outbreak's impact ($R_\infty = R_H/N_H$). In contrast, the model described in Sec. 2.3 is highly precise and capable of quantifying impacts.

In this endeavour, we are merging these two methodologies to leverage the temperature-dependent aspects for global assessments while quantifying the disease's impact on a local scale. The differential equations used to accomplish this aim are the following ones:

$$
\begin{aligned}
\dot{S}_H &= -\frac{\beta S_H I_v}{N_H}, \\
\dot{I}_H &= \frac{\beta S_H I_v}{N_H}, \\
\dot{E}_H &= -\gamma I_H, \\
\dot{R}_H &= \gamma I_H, \\
\dot{S}_v &= -\frac{\alpha S_v I_H}{N_H} - \mu S_v, \\
\dot{I}_v &= \frac{\alpha S_v I_H}{N_H} - \mu I_v.
\end{aligned}
\tag{13}
$$

Where it is crucial to apply new initial conditions at the beginning of each year of our simulation in order to capture the seasonal dynamics of vectors and the thermal dependence of PD. This is the so called annual reset, given by the four expressions above:

$$\begin{aligned} S_H(t^*) &= S_H(t_f) + E_H(t_f)[1 - \Pi(t)], \\ I_H(t^*) &= I_H(t_f) + E_H(t_f)\Pi(t), \\ E_H(t^*) &= 0, \\ S_v(t^*) &= N_v(0), \end{aligned} \tag{14}$$

being t^* the initial condition of each year. In essence, we effectively restore the susceptible vector population to its initial numbers. This is because Ps have an annual generational life cycle, meaning they are born and die within a year. Additionally, we reassign exposed hosts to either the susceptible or infective categories based on the prevailing climatic conditions that either facilitate the proliferation or recovery of PD.

3 Results and discussion

During the Julia and Python simulations of our theoretical model (2.4), we have configured the parameter values, which are expressed as daily rates, as follows: $\alpha = 0.076$, $\beta = 0.013$ $\gamma = 0.0005$, $\mu = 0.022$. These parameter settings were used in conjunction with our specified initial conditions:

$$\begin{aligned} &\text{Populations} \quad \Rightarrow \quad N_H(0) = 1000, \;\; N_v(0) = fN_H(0); \\ &\text{Vectors} \quad \Rightarrow \quad S_v(0) = N_v(0), \;\; I_v(0) = 0; \\ &\text{Hosts} \quad \Rightarrow \quad E_H(0) = R_H(0) = 0, \;\; I_H(0) = 0.01N_H(0), \;\; S_H(0) = N_H(0) - I_H(0). \end{aligned}$$

Being f the Ps suitability in a certain location.

3.1 Phase diagram

Our initial outcome was the phase diagram depicted in Fig. 5, which served as a valuable tool for verifying the proper functioning of our simulation, but also for myself in order to become familiar with the programming language. In this image, we illustrate the impact of a PD outbreak, R_∞, concerning the climatic conditions that influence the infection probability, Π, and the Ps suitability, f, which in our case (due to the initial conditions), is proportional to its population.

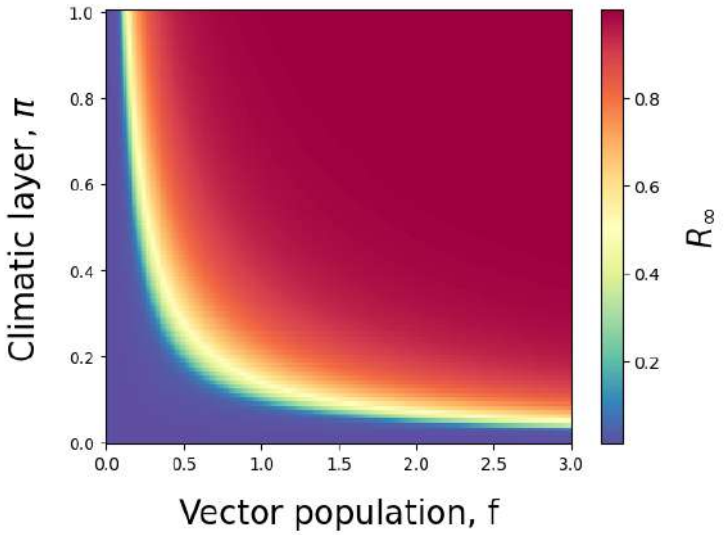


Figure 5: Phase diagram, showing the final size of the epidemic, R_∞, obtained from our simulation given 0 to 1 and 0 to 3 values of the Xf and Ps suitability, Π and f respectively, both with 0.01 steps.

The presented results follow an intuitive pattern: as Xf suitability increases, the deaths attributable to PD also rise, and likewise, with higher Ps suitabilities, the impact is more pronounced. However, what adds an intriguing dimension to this logical observation is the presence of a threshold (highlighted in blue), which suggests that, initially, we need not be overly concerned about a PD outbreak due to a very low R_∞.

3.2 PD impact and risk maps

Now that we have successfully tested our simulation, our main goal is to assess the real-world impact on a global scale and compare it with the PD risk as described in [2]. To accomplish this objective, we will utilise existing values of the Ps and Xf time-dependent suitability corresponding to different geographical coordinates, f and Π, with the latter one being a result of the climatic factor.

3.2.1 Historical data

In Fig. 6 we depict the outcomes derived from historical data. Notably, in both impact and PD risk assessments, coastal areas (such as Greece, Italy, and southern Spain) as well as Mediterranean islands (including Majorca, Sardinia, and Sicily) emerge as the most heavily affected regions.

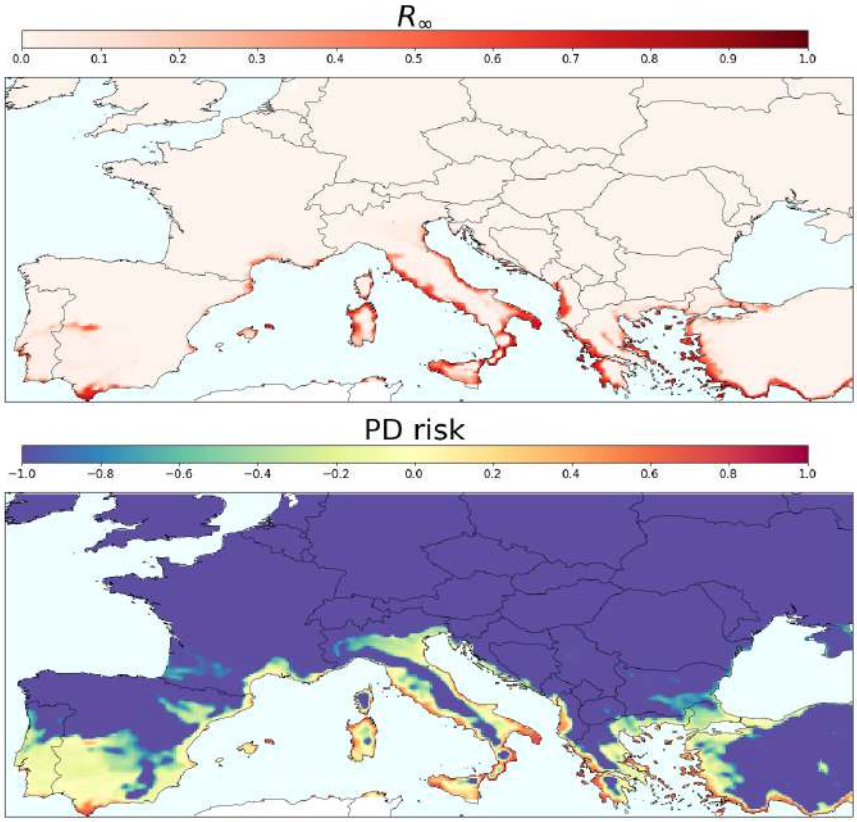


Figure 6: PD impact, R_∞, and risk, r, global maps acquired from historical data.

Moreover, given the striking resemblance between the behaviours of R_∞ and r, we conducted a comprehensive scatter plot analysis, as illustrated in Fig. 7, to delve into the potential correlation between these two variables. The results of this graphical representation unveil a distinct pattern: when the risk index assumes a negative value (indicating either an inconsequential risk or an exceedingly low likelihood of outbreaks), the resulting impact tends to be minimal. Conversely, as the risk index becomes positive, indicating an increasing probability of outbreaks, the impact escalates accordingly. Furthermore, it is worth noting that within the delineated "transition zone" demarcated by red dotted lines, encompassing risk values ranging from -0.1 to 0.1, the associated uncertainty is notably elevated, clearly evident through the presence of larger error bars, aligning with our expectations.

Last but not least, we can discern two distinct trends among the blue data points in the scatter plot, one positioned higher and the other lower. This intriguing observation may be linked to the geographical distribution of these points, potentially influenced by varying climatic conditions or other factors that warrant further investigation.

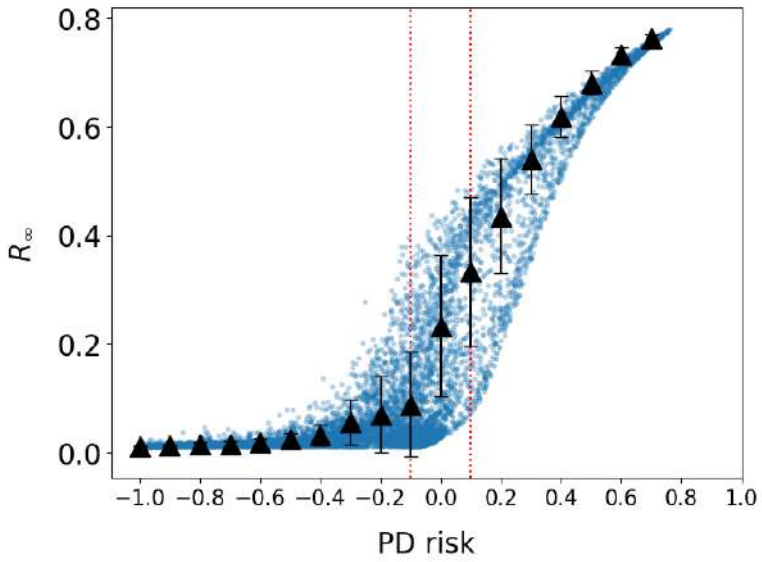


Figure 7: Average final size of the epidemic, R_∞, extracted from the scatter plot analysis (depicted in blue) is represented with black triangles and accompanied by its corresponding uncertainty, versus the risk of establishment of PD, r; the "transition zone" demarcated by red dotted lines.

3.2.2 Warming scenarios

Finally, we carried out a similar analysis to the one in the previous section, but this time under various Climate Change scenarios, in particular involving temperatures surpassing those observed during the pre-industrial era.

Remarkably, our findings mirror those derived from the historical data (Fig. 8 and 9), with coastal regions and Mediterranean islands consistently exhibiting the highest values for both R_∞ and r across all four scenarios. However, a distinct trend emerged in central regions, where the potential for a PD outbreak became increasingly evident, particularly in scenarios characterized by higher temperatures. This last observation is inherently intuitive, as it follows logical reasoning. In warming scenarios, the Xf suitability, represented by Π, naturally tends to increase, amplifying the likelihood of disease development within the vineyard after an inoculation event.

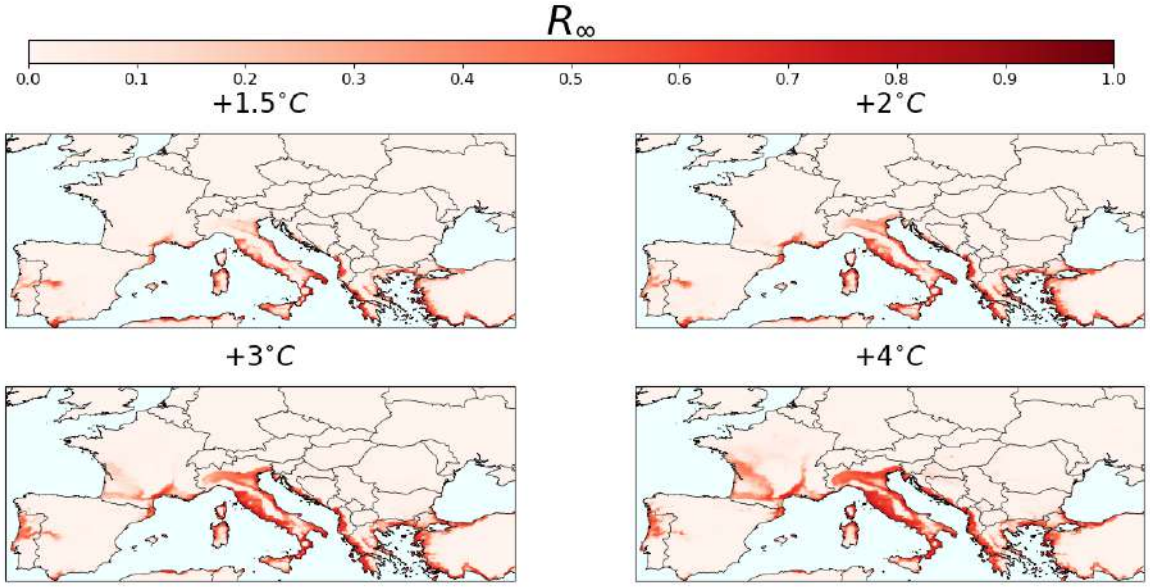


Figure 8: PD impact, R_∞, global maps acquired from different warming scenarios, where the starting temperature is the one corresponding to the pre-industrial period (+0.0℃).

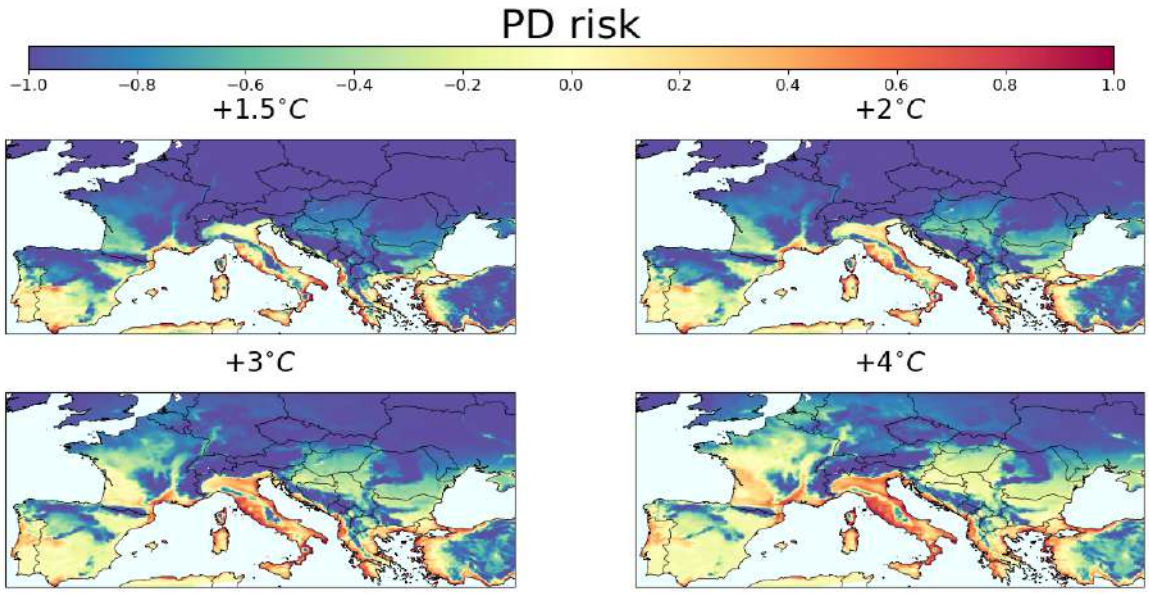


Figure 9: PD risk, r, global maps acquired from different warming scenarios, where the starting temperature is the one corresponding to the pre-industrial period (+0.0℃).

On the other hand, the results extracted from the scatter plot analysis between R_∞ and r across the four warming scenarios continue to exhibit the same pattern observed in the previous section. In essence, when the risk stays into negative territory (indicating a low likelihood of a PD outbreak), the corresponding impact remains consistently minimal, and vice versa. Furthermore, as anticipated, there is a discernible uptick in associated uncertainty within the "transition zone", in this case demarcated by black dotted lines. To streamline the presentation and prevent undue repetition, we have amalgamated all the findings into a unified graphic, depicting solely the average impact at 0.1 risk intervals (Fig. 10). This representation unmistakably underscores the previously mentioned trends and the striking similarity in the relationship between the final size of the epidemic, R_∞, versus the risk of espablishment of PD, r, among all four scenarios.

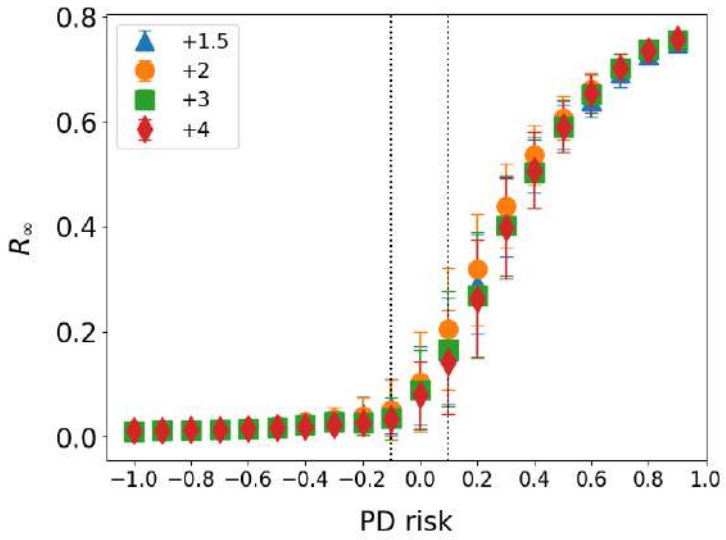


Figure 10: Average final size of the epidemic, R_∞, extracted from the scatter plot analysis is represented for the four climatic scenarios with their corresponding uncertainty versus the risk of establishment of PD, r; the "transition zone" is demarcated by black dotted lines.

4 Conclusions

In this study, we delved into the complex dynamics of Pierce's Disease (PD), shedding light on its potential implications for the global wine industry. Our research has unraveled critical insights:

PD stands as a formidable threat to viticulture on a global scale, with coastal and Mediterranean regions emerging as particularly vulnerable areas. The disease's propensity to thrive in these regions underscores the need for vigilant monitoring and proactive management strategies to safeguard vineyards.

Temperature, as our analysis has emphasized, plays a pivotal role in shaping the dynamics of PD. Under the spectre of climate change, regions previously considered low-risk may face an elevated threat. This emphasizes the urgency of implementing adaptive measures to curtail disease spread.

Furthermore, our findings have underscored the profound influence of vectors in disease transmission. Understanding vector behaviour is of paramount importance for crafting effective control strategies and mitigating the impact of PD outbreaks.

Last but not least, we have unveiled a significant direct connection between the risk of PD establishment, r, and the ultimate epidemic's size, R_∞, bridging the abstract and tangible realms of the disease dynamics.

In conclusion, our study underscores the vital necessity for ongoing research, vigilant surveillance, and climate-adaptive initiatives within the wine industry. Proactive measures are indispensable in fortifying vineyards against the evolving challenges posed by Pierce's Disease, thereby ensuring the resilience and sustainability of viticulture worldwide.

Future work

Taking the research forward, a logical extension would be a comprehensive modeling of Pierce's Disease to evaluate its potential economic repercussions at a global scale, leveraging our research as a cornerstone. This economic analysis would serve as a vital resource for stakeholders in the wine industry, enabling them to quantify the financial implications of PD's geographical prevalence and severity.

Furthermore, delving deeper into the different two behaviours identified in the previously shown scatter plots represents an exciting avenue for exploration. These two trends, which manifest as distinct clusters of data points, may hold valuable insights into the influences of specific geographical regions characterized by unique climates or other pertinent factors. Investigating these connections could unravel the intricate dynamics of PD across different locales, ultimately contributing to the development of region-specific strategies for disease management and prevention.

Acknowledgments

This work was supported by the SURF@IFISC fellowship, with the invaluable help and constant guidance provided by both my supervisors, Alex Giménez-Romero and Manuel A. Matías. Without their unwavering assistance this work could not have been possible. I would also like to express my sincere appreciation to the entire SURF team for giving me this incredible opportunity to take my first steps towards research in such a prestigious and reputed institution as IFISC.

References

[1] Giménez-Romero, A., Flaquer-Galmés, R. & Matías, M. A. *Vector-borne diseases with nonstationary vector populations: The case of growing and decaying populations*. Phys. Rev. E 106, 054402 (2022). https://doi.org/10.1103/PhysRevE.106.054402

[2] Giménez-Romero, A., Galván, J., Montesinos, M. et al. *Global predictions for the risk of establishment of Pierce's disease of grapevines*. Commun Biol 5, 1389 (2022). https://doi.org/10.1038/s42003-022-04358-w

[3] Giménez-Romero, A., Moralejo, E. & Matías, M. A. *A compartmental model for Xylella fastidiosa diseases with explicit vector seasonal dynamics*. The American Phytopathological Society (APS). https://doi.org/10.1094/PHYTO-11-22-0428-V

[4] Murray, J. D. Dynamics of Infectious Diseases: Epidemic Models and AIDS, *Mathematical Biology: I. An Introduction* (pages 315–394). Springer, New York, NY, 3 edition (2002) https://doi.org/10.1007/978-0-387-22437-4_10

[5] Tumber, K. B., Alston, J. M. & Fuller, K. B. *Pierce's disease costs California $104 million per year*. Calif. Agr. 68, 20 (2014). https://doi.org/10.3733/ca.v068n01p20

[6] https://www.flickr.com/photos/m-a-r-t-i-n/14446738521/in/photostream/

Optimizing Quantum Thermal Machines

Ignacio Megía Pérez, Rosa López Gonzalo
Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB)
Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain

Abstract

In the context of Stochastic Thermodynamics, a series of recently discovered inequalities known as *Thermodynamic Uncertainty Relations (TURs)*, have shown that high precision of non-equilibrium currents comes with an entropy cost. Demonstrations for these results are not valid for some important classes of dynamics. In this work, we analyze a simple example of those systems that can violate the original form of TUR: the electronic transport through a quantum dot connecting two metal leads. To do so, we apply *Landauer-Büttiker formalism* and make use of analytical calculations as well as numerical tools to find how TUR bound can be overcome. We then study how we could seize the properties of these system to build a quantum thermal machine.

1 Introduction

Thermodynamic Uncertainty Relations (TURs) have been generally proved under a variety of different assumptions, and checked in many processes for wich demonstration has not been done yet. The study of biomolecular processes, such as molecular motors, transcription and translation machinery and other enzymatic reactions motivated the analysis made in [2]. The authors propose a random walk model to describe a non-equilibrium chemical reaction catalyzed by an enzyme, and consider the number X of enzymatic cycles completed. Assuming an external environment of fixed temperature T, they demonstrate that after a time t, the relative uncertainty of X

$$\epsilon^2 := \frac{< X^2 > - < X >^2}{< X >^2}$$

and the entropy prodction rate, σ, which gives a total energy dissipation $T\sigma t$, verify

$$\mathfrak{Q} := T\sigma t\epsilon^2 \geq 2k_B T \tag{1.1}$$

After that, they show that the relation (1.1) holds for any network of states within linear response theory. For non-linear regime, they prove it analytically for unicyclic networks and give numerical examples for multicyclic cases. They finally conjecture this is a fundamental limit for nonequilibrium stationary states.

In [4], a stochastic model of a non-equilibrium system based on random jumps between discrete states is proposed. The system is supposed to interact with individual thermodynamic reservoirs with well-defined equilibrium propertes that mediate its transitions, which imposes the *local detailed balance* condition. The system relaxes into a non-equilibrium steady state characterized during a time window τ by an average entropy generation produced Σ_τ and irreversible currents, like a generic one $\mathfrak{J}_\tau$. These currents fluctuate when comparing different trajectories due to the inherent noise in the dynamics, and it can be shown that

$$\frac{\mathrm{Var}(\mathfrak{J}_\tau)}{< \mathfrak{J}_\tau >^2} \geq \frac{2k_B}{\Sigma_\tau} \tag{1.2}$$

This relation was proved within the framework of large deviation theory for jump processes. It can also be applied to more general situations (see [4] for details), but there are some interesting systems in which it is not valid. Here, we will study one of such processes, the electronic transport through a quantum dot connecting two metal leads linked to reservoirs.

Further investigations have generalised relations like the original form of TUR, (1.2), replacing the right-hand side of the equation by an arbitrary function of Σ_τ/k_B, but still, analysing how the first form can be violated is interesting due to the wide range of applicabitilty of this formula: In general, it is true for every *Markovian process* that verifies the *local detailed balance* condition.

2 Theoretical model

In this work, we study the flow of electrons between two metal leads connected by a quantum dot. Although it may seem a simple system, it presents many non-trivial problems, can work as a quantum thermal machine, and gives us the possibility to show how (1.2) can be violated in Quantum Mechanics.

2.1 Landauer-Büttiker formalism

We will make use of *Landauer-Büttiker formalism*, which is described in detail in [6]. We will take without demonstration some of the expressions that are derived there, especially in its first two chapters. This model of electronic transport is sometimes called *scattering approach* because it consists in studying the propagation of electrons as a quantum-mechanical scattering problem: A mesoscopic system, which we will call *scatterer* connects N_r macroscopic contacts that are considered to be equilibrium electrons reservoirs with fixed temperatures.

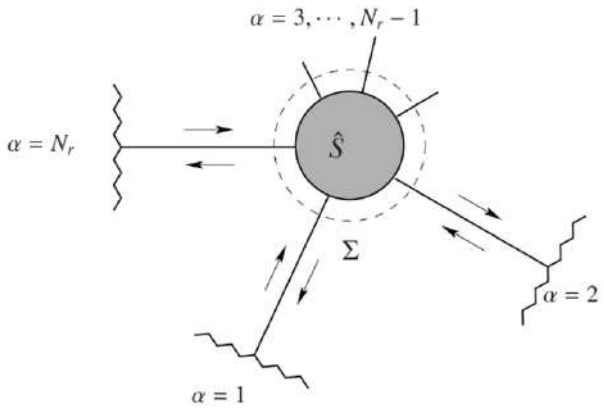


Figure 1: Scatterer and leads diagram. Figure taken from [6].

The main objective in Landauer-Büttiker formalism is to relate wave functions of scattered particles with those of incident ones. Instead of solving Schrödinger equation in general, it is enough to do it for the states of an orthonormal basis, $\psi_\alpha^{(\text{in})}$, $\psi_\alpha^{(\text{out})}$

$$\Psi^{(\text{in})} = \sum_\alpha a_\alpha \psi_\alpha^{(\text{in})} \quad \Psi^{(\text{out})} = \sum_\beta b_\beta \psi_\beta^{(\text{out})}$$

Solving the propagation problem for those states we get

$$\Psi_\gamma^{(\text{out})} = \sum_\beta S_{\beta\gamma} \psi_\beta^{(\text{out})}$$

where $\hat{S} = (S_{\mu\nu})_{\mu\nu}$ is the so-called *scattering matrix*, which allows us to write $b_\beta = \sum_\alpha S_{\beta\alpha} a_\alpha$. Scattering matrix has some important properties whose derivation can be found in [6]. Here we will make use of

$$\text{unitarity: } \hat{S}^\dagger \hat{S} = \hat{S}\hat{S}^\dagger = \hat{I} \implies \sum_\alpha |S_{\alpha,\beta}|^2 = \sum_\beta |S_{\alpha,\beta}|^2 = 1 \tag{2.1}$$

$$\text{symmetry: } \hat{S} = \hat{S}^T \iff S_{\alpha\beta} = S_{\beta\alpha} \tag{2.2}$$

where symmetry property changes in presence of a magnetic field, but we will not have to consider that situation. Instead of using a discrete basis, it is also possible apply this formalism to a continuous spectrum of states dependent of energy E, and so, from this point we will use the subscripts to mark the lead where the electron is.

In order to calculate the current flowing between the scatterer and the reservoirs, it is necessary to introduce the operators of *Second Quantization*, which create and annihilate particles $\hat{a}_\alpha^\dagger(E)/\hat{a}_\alpha(E)$ $\hat{b}_\alpha^\dagger(E)/\hat{b}_\alpha(E)$. With them, it is possible to define the *current operator*

$$\hat{I}_\alpha(t,x) = \frac{i\hbar e}{2m} \int \left\{ \frac{\partial \hat{\Psi}_\alpha^\dagger(t,\mathbf{r})}{\partial x} \hat{\Psi}_\alpha(t,\mathbf{r}) - \hat{\Psi}_\alpha^\dagger(t,\mathbf{r}) \frac{\partial \hat{\Psi}_\alpha(t,\mathbf{r})}{\partial x} \right\} dr_\perp \tag{2.3}$$

where $\hat{\Psi}^\dagger_\alpha$ and $\hat{\Psi}^\dagger_\alpha$ represent the *field operators*. It can be shown that creation and destruction operators for incoming and scattered electrons are also related via the scattering matrix:

$$\hat{b}_\alpha = \sum_{\beta=1}^{N_r} S_{\alpha\beta}\hat{a}_\beta \qquad \hat{b}^\dagger_\alpha = \sum_{\beta=1}^{N_r} S^*_{\alpha\beta}\hat{a}^\dagger_\beta$$

Making use of these relations, scattering matrix properties, and the assumption that incoming electrons have equilibrium properties, a quantum-statistical average calculation leads to the *measurable current*

$$I_\alpha =< \hat{I}_\alpha >= \frac{e}{h}\int_{-\infty}^{\infty} \left\{\phi^{(\text{out})}_\alpha(E) - \phi^{(\text{in})}_\alpha(E)\right\} dE \tag{2.4}$$

with

$$\phi^{(\text{in})}_\alpha(E) = f_\alpha(E) \equiv \frac{1}{1+e^{\frac{E-\mu_\alpha}{k_B T_\alpha}}} \quad \mu_\alpha = \mu_0 + eV_\alpha \quad \text{(Fermi function)} \tag{2.5}$$

$$\phi^{(\text{out})}_\alpha(E) = \sum_{\beta=1}^{N_r} |S_{\alpha\beta}(E)|^2 f_\beta(E) \tag{2.6}$$

where μ_0 is the (common) Fermi energy of all leads and V_α the electrical potential of lead α, so that $\mu_\alpha = \mu_0 + eV_\alpha$ is its electro-chemical potential.

Finally, substituting (2.5) and (2.6) in (2.4) and using (2.1) we can write

$$I_\alpha =< \hat{I}_\alpha >= \frac{e}{h}\int_{-\infty}^{\infty} \sum_{\beta=1}^{N_r} |S_{\alpha\beta}(E)|^2 \left\{f_\beta(E) - f_\alpha(E)\right\} dE \tag{2.7}$$

which, together with the properties of $\hat{S}$, (2.1) and (2.2), allows to prove a current conservation law

$$\sum_{\alpha}^{N_r} I_\alpha = 0 \tag{2.8}$$

In short, Landauer-Büttiker formalism allows us to calculate the currents between the scatterer and the leads once we know the scattering matrix. For the two-contact case, the scattering matrix can be written as

$$\hat{S} = e^{i\gamma}\begin{pmatrix} \sqrt{\mathcal{R}}e^{-i\theta} & i\sqrt{\mathcal{T}}e^{-i\phi} \\ i\sqrt{\mathcal{T}}e^{i\phi} & \sqrt{\mathcal{R}}e^{i\theta} \end{pmatrix} \tag{2.9}$$

where $\mathcal{R}+\mathcal{T}=1$ and ϕ is an odd function of the magnetic field H, and so vanishes if $H=0$, which will be our case. Now, using (2.9), (2.7) reads

$$I_\alpha =< \hat{I}_\alpha >= \frac{e}{h}\int_{-\infty}^{\infty} \mathcal{T}\left\{f_\beta(E) - f_\alpha(E)\right\} dE \quad (\beta \neq \alpha) \tag{2.10}$$

2.2 Currents through a quatum dot

The first step to calculate the scattering matrix is describing the system that connects the leads. In this work, we chose a quantum dot because these structures present interesting features: their transmission probability curve is highly dependent on the energy and it has been demonstrated that they have important thermoelectric properties, as well as quantum coherence.

A wide analysis of quantum dots can be found in [5]. For our purposes, see (2.10), we only need to know the transmission probability through the dot. Propagation of electrons can be described, in full analogy to a Fabry-Perot interferometer, as the interference of partial waves reflecting back and forth between the tunneling barriers of the dot. We encourage the reader to check in [5], section 18.3, the beautiful derivation of the formula

$$\mathcal{T}(E) = \frac{\Gamma_1\Gamma_2}{(E-\epsilon)^2 + (\Gamma/2)^2} \tag{2.11}$$

where Γ_1 and Γ_2 are positive energies whose sum is Γ and ϵ is the energy of the level that the dot has available for electronic transport. This is actually a simplification of the expression in [5],

where a sum over resonances is performed. Equation (2.11) gives us a lorentzian-form transmission probability that can also be written as

$$\mathcal{T}(E) = \frac{\gamma_1\gamma_2\Gamma^2}{(E-\epsilon)^2+\Gamma^2/4} = \frac{4\gamma_1\gamma_2}{1+\left(2\frac{E-\epsilon}{\Gamma}\right)^2} \quad \text{where} \quad \gamma_1 = \frac{\Gamma_1}{\Gamma},\ \gamma_2 = \frac{\Gamma_2}{\Gamma} \in [0,1] \tag{2.12}$$

Now we can finally write the complete expression to calculate, for example, the charge current through lead 1. We will change the sign of the integrand because from now on we will follow the same sign convention as in [1], instead of [6]. With this convention, we will consider positive those currents and energy flows towards the system, and negative those coming from the system. Then, we will write

$$\mathcal{I}_1 = \frac{e}{h}\int_{-\infty}^{\infty} \mathcal{T}(E)[f_1(E)-f_2(E)]\,dE = \frac{e}{h}\int_{-\infty}^{\infty} \frac{\Gamma_1\Gamma_2}{(E-\epsilon)^2+(\Gamma/2)^2}[f_1(E)-f_2(E)]\,dE \tag{2.13}$$

where we recall that f_α represents the Fermi function of contact α.

A similar discussion (see [3]) allows us to define and give an expression to calculate the *energy current* from lead α towards the scatterer

$$\mathcal{E}_\alpha = \frac{1}{h}\int_{-\infty}^{\infty} E\left\{\phi_\alpha^{(\text{out})}(E)-\phi_\alpha^{(\text{in})}(E)\right\}\,dE = \frac{1}{h}\int_{-\infty}^{\infty} E\sum_{\beta=1}^{N_r}|S_{\alpha\beta}(E)|^2\left\{f_\beta(E)-f_\alpha(E)\right\}\,dE \tag{2.14}$$

and the *heat current*,

$$\begin{aligned}\mathcal{J}_\alpha = \mathcal{E}_\alpha - \frac{\mu_\alpha}{e}\mathcal{I}_\alpha &= \frac{1}{h}\int_{-\infty}^{\infty}(E-\mu_\alpha)\left\{\phi_\alpha^{(\text{out})}(E)-\phi_\alpha^{(\text{in})}(E)\right\}\,dE = \\ &= \frac{1}{h}\int_{-\infty}^{\infty}(E-\mu_\alpha)\sum_{\beta=1}^{N_r}|S_{\alpha\beta}(E)|^2\left\{f_\beta(E)-f_\alpha(E)\right\}\,dE\end{aligned} \tag{2.15}$$

which for the two-terminal case and lead 1 reads

$$\mathcal{J}_1 = \frac{1}{h}\int_{-\infty}^{\infty}(E-\mu_1)\mathcal{T}(E)[f_1(E)-f_2(E)]\,dE = \frac{1}{h}\int_{-\infty}^{\infty}(E-\mu_1)\frac{\Gamma_1\Gamma_2}{(E-\epsilon)^2+(\Gamma/2)^2}[f_1(E)-f_2(E)]\,dE \tag{2.16}$$

Integrals (2.13) and (2.16) were solved analytically in [1] using digamma function, Ψ, to rewrite fermi functions and then use residue theorem. The results are

$$\mathcal{I}_1 = -\frac{2e}{h}\frac{\Gamma_1\Gamma_2}{\Gamma}\,\text{Im}\left[\Psi\left(z_1^+\right)-\Psi\left(z_2^+\right)\right] \tag{2.17}$$

$$\mathcal{J}_1 = -\frac{2\Gamma_1\Gamma_2}{h\Gamma}(\epsilon-\mu_1)\,\text{Im}\left[\Psi\left(z_1^+\right)-\Psi\left(z_2^+\right)\right] + \frac{\Gamma_1\Gamma_2}{h}\,\text{Re}\left[\Psi\left(z_1^+\right)-\Psi\left(z_2^+\right)\right] - \frac{\Gamma_1\Gamma_2}{h}\ln\left(\frac{T_2}{T_1}\right) \tag{2.18}$$

$$\text{where} \quad z_j^\pm = \left(\frac{1}{2}+\frac{\Gamma}{4\pi k_\text{B}T_j} \pm i\frac{\epsilon-\mu_j}{2\pi k_\text{B}T_j}\right)$$

This analytical expressions are hard to work with, due to the presence of digamma function. However, they will be very useful to save time in our calculations, and avoid precision problems, because *Mathematica* is able to quickly compute Ψ and with high accuracy.

TUR inequality also features entropy increase. For our experiment, we can use heat currents to easily calculate the entropy generation rate by means of *Clausius theorem*. Since the scatterer has fixed properties in time and entropy is a state function, its entropy is constant in time. Meanwhile, the leads do change their entropy, due to the heat flow between them and the scatterer. Consequently, and taking into account that according to our sign convention reservoir α receives a heat current $-\mathcal{J}_\alpha$, we may write that the entropy generation rate for the thermodynamic universe is

$$\mathcal{K} = -\frac{\mathcal{J}_1}{T_1}-\frac{\mathcal{J}_2}{T_2} \overset{(2.15)}{=} \frac{k_\text{B}}{h}\int_{-\infty}^{\infty}\mathcal{T}(E)(f_1(E)-f_2(E))\left[\frac{E-\mu_2}{k_\text{B}T_2}-\frac{E-\mu_1}{k_\text{B}T_1}\right]\,dE \tag{2.19}$$

The last ingredient we need to check the TUR is the magnitude of the fluctuations of $\mathcal{I}_1$, also known as *noise*. It is characterized by the mean square fluctuations of the current, and its calculation is explained in detail in [6]. A generalisation of the result derived there leads to

$$\mathcal{S}_1 = \frac{e^2}{h}\int_{-\infty}^{\infty}\left\{\mathcal{T}[f_1(1-f_1)+f_2(1-f_2)]+\mathcal{T}(1-\mathcal{T})(f_2-f_1)^2\right\}\,dE \tag{2.20}$$

where the energy dependency has been omitted for simplicity.

Finally, if we choose the current flowing from lead 1 to the scatterer, $\mathcal{I}_1$ as observable, TUR would be formulated as

$$\frac{\mathcal{S}_1}{\mathcal{I}_1^2} \geq \frac{2k_B}{\mathcal{K}} \tag{2.21}$$

3 Results and discussion

The main objective of this paper is to study the original form of Thermodynamic Uncertainty Relation for an experiment of electronic transport through a quantum dot. Let us summarize the situation. The inequality we are considering is (2.21)

$$\frac{\mathcal{S}_1}{\mathcal{I}_1^2} \geq \frac{2k_B}{\mathcal{K}}$$

where

- $\mathcal{I}_1$ is the current flowing from lead 1 towards the scatterer.
- $\mathcal{K}$ is the time derivative of the entropy of the thermodynamic universe.
- $\mathcal{S}_1$ is the noise of $\mathcal{I}_1$.

In order to be able to study whether and "how much" TUR inequality is verified or not, we can rewrite (2.21) by multiplying both sides by $\mathcal{K}/2k_B$, which is positive, so the inequality sign does not change. It must be true that $\mathcal{K}/2k_B > 0$, because otherwise this system would violate the Second Law of Thermodynamics. It will be proved later (see section 3.3.1) that indeed, this quantity is positive. Then, if we call the left-hand side of the resulting inequality as Θ, we have that TUR can be reformulated as

$$\Theta := \frac{\mathcal{K}}{2k_B}\frac{\mathcal{S}_1}{\mathcal{I}_1^2} \geq 1 \tag{3.1}$$

In conclusion, our main mission will be to determine whether it is possible to find a configuration in our experiment that yields $\Theta < 1$, how we can get it, and if we can take advantage of this situation to build more efficient thermal machines.

3.1 Analysis of parameters

There are several parameters in our electronic transport model. If we are able to understand, at least in a qualitative approach, how they affect the charge and entropy currents and noise, it will be easier to study the TUR and the "thermodynamic behaviour" of our system. Hopefully, this study will enable us to choose appropriate values for parameters to make a thermal machine out of it.

Since the mathematical expressions, both the integrals (2.13) and (2.16) and analytical solutions for currents, are hard to work with, we believe that the simplest way to deal with them is by a qualitative inspection. To do so, representations of the functions that appear in the integrands and those integrals themselves are extremely useful. That is why we made a *Manipulate* in Mathematica to plot together Fermi functions $f_1(E)$ and $f_2(E)$ and the transmission probability $\mathcal{T}(E)$. The instructions to implement it are provided in the appendix.

Let us set electrical potentials as $V_1 = V/2$, $V_2 = -V/2$, so that $V = V_1 - V_2$. Since $\gamma_1 + \gamma_2 = 1$, we can consider a single parameter $g = \gamma_1$ so that $\gamma_2 = 1 - g$. It will be convenient to work only with adimensional quantities and energies, so we will usually consider eV and $k_B T_\alpha$ instead of V and T_α. We have 7 different parameters to consider $(k_B T_1, k_B T_2, \mu_0, eV, \Gamma, g, \epsilon)$, which is one of the main difficulties of this model, because it is impossible to study them simultaneously in graphics and it makes analysis quite tedious. This fact, added to the difficulty of expressions we have for our magnitudes, led us to start approaching the problem with two different strategies: simplifying our set of variables and programming a flexible optimization function.

3.1.1 Simplifying the set of parameters

Some simple observations can help us to reduce the number of degrees of freedom in our experiment. At first sight, they may seem unimportant or even childish, but eliminating coordinates will make all the analysis much easier.

First of all, we can replace μ_0 and ϵ by a single parameter. This two quantities only appear subtracting from the energy E in the Fermi functions and the transmission coefficient, respectively.

Since E is the integration variable to calculate currents and noise, we can simply make a change of variable $E \to E + \mu_0$ to make the Fermi energy vanish in Fermi functions and transform the transmission coefficient into

$$\mathcal{T}(E) = \frac{\Gamma_1\Gamma_2}{[E-(\epsilon-\mu_0)]^2 + (\Gamma/2)^2}$$

so that now, the only quantity that matters is the difference $\epsilon - \mu_0$. Equivalently, we can simply assume $\mu_0 = 0$ from now on, which is, in fact, a way to choose the zero of our energies. That is why unless otherwise stated $\mu_0 = 0$.

Now we will show that the symmetric barrier, $\Gamma_1 = \Gamma_2 = \Gamma/2$ or equivalently $\gamma_1 = \gamma_2 = 1/2$ or $g = 1/2$, is always the optimal situation to make Θ smaller. Given that $\Gamma_1\Gamma_2 = g(1-g)\Gamma^2$ appears as a factor in $\mathcal{T}$, we can take $g(1-g)$ out of the integral in (2.13) and (2.16). Now if we rewrite the integrand of the noise expression (2.20)

$$\mathcal{T}[f_1(1-f_1)+f_2(1-f_2)]+\mathcal{T}(1-\mathcal{T})(f_2-f_1)^2 = \mathcal{T}\{f_1(1-f_1)+f_2(1-f_2)+(f_2-f_1)^2\underbrace{-\mathcal{T}(f_2-f_1)^2}_{<0}\} \tag{3.2}$$

we can see that another factor $g(1-g)$ can get out of the integral of $\mathcal{S}_1$, and join the one coming from $\mathcal{K}$ to cancel, in (2.21), the squared one from $\mathcal{I}_1$. Then, g disappears everywhere except for the negative term in (3.2). Since we want reduce $\mathcal{S}_1$, we can pick the value of $g \in [0,1]$ that maximizes $g(1-g)$, which is $g = 1/2$, i.e. the symmetric barrier case. Then, from this time on, g will be $1/2$ when omitted.

Last, but not least, we can also "get rid of" Γ by using it as our energy scale, dividing all energies by it. That is precisely what we did in (2.12) for $\mathcal{T}$, and we can also write

$$\frac{E-\mu_\alpha}{k_\mathrm{B}T_\alpha} = \frac{E/\Gamma - \mu_\alpha/\Gamma}{k_\mathrm{B}T_\alpha/\Gamma}$$

and make a change of variable $E = \Gamma w$, which gives a Γ factor that can be taken out of the integrals. In the end, we can replace our original parameters by their adimensional rates over Γ. This is especially interesting when computing the integrals with Mathematica.

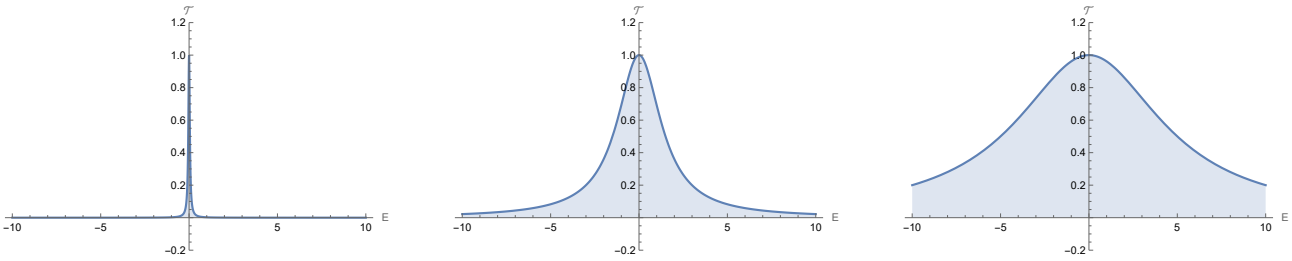
Although we could let Γ take negative values, we will assume $\Gamma \geq 0$, unless otherwise stated, because negative values are in general unnecessary and more difficult to interpret. Note that, appart from being used as our energy scale, Γ has a "geometrical" meaning as a mesure of the area under the transmission probability curve. This is because the integral of the lorentzian is

$$\int_{-\infty}^{\infty}\mathcal{T}(E)\,dE \overset{(2.12)}{=} 4\gamma_1\gamma_2\int_{-\infty}^{\infty}\frac{1}{1+\left(2\frac{E-\epsilon}{\Gamma}\right)^2}\,dE \overset{x=2\frac{E-\epsilon}{\Gamma}}{=} 4\gamma_1\gamma_2\frac{\Gamma}{2}\int_{-\infty}^{\infty}\frac{1}{1+x^2}\,dx = 4\gamma_1\gamma_2\frac{\pi\Gamma}{2} \tag{3.3}$$

and for every $E \neq 0$, $\mathcal{T}(E)$ grows with $|\Gamma|$, as can be demonstrated by observing the sign of

$$\frac{\partial T}{\partial\Gamma} = \frac{\partial}{\partial\Gamma}\frac{\gamma_1\gamma_2\Gamma^2}{(E-\epsilon)^2+\Gamma^2/4} = \frac{32\gamma_1\gamma_2(w-\epsilon)^2\Gamma}{[(E-\epsilon)^2+\Gamma^2/4]^2} \tag{3.4}$$

Figure 2 illustrates how transmission probability and the area under the curve depend on Γ. We can see that $\mathcal{T}(\epsilon) = 4\gamma_1\gamma_2$ for all $\Gamma \neq 0$, but for the rest of energies, the $\mathcal{T}(E) \to 0$ when $\Gamma \to 0$, so that the lorentzian gets sharper.



(a) $\Gamma = 0.1$ (b) $\Gamma = 3$ (c) $\Gamma = 10$

Figure 2: Geometrical meaning of Γ. To make this plots, we took $g = 1/2$.

3.1.2 minimACE

As we do not have simple analytical expressions for our magnitudes of interest, numerical calculations and graphics appear to be an interesting alternative to the standard analytical optimization

techniques. However, the high number of degrees of freedom makes it impossible to study how magnitudes depend on all of them together. Moreover, depending on the particular functions we wish to study, their number of variables will be different. This is what motivated us to build a tool in Mathematica to minimize functions with the most general conditions possible: *minimACE*.

This function takes as an argument a function F of n variables. The second argument is a matrix with n rows, each one of which contains the lower and higher values that a variable will take, and a positive integer to indicate how many evenly separated values, including the extremes, that variable must take. Next argument is a positive integer to choose the number of values of F that will be returned, and another positive integer to mark the maximum number of computations minimACE must make. This number is a manual limit for the loop minimACE uses and is also the reference for the progress messages it can print. The last two arguments are a couple of Boolean values to control whether the initial verifications of minimACE are to be done, and to ask minimACE to print progress messages when calculating. The output is a list of lists, each one formed by a list of values for the variables and the result of F applied to those values. Those values of F are the m lowest among the calculated ones. Figure 3 shows an example of how minimACE works, and the code of the function is provided in the appendix.

```
F[x_, y_] := x^2 + y^2;
A = {{-1, 1, 201}, {-2, 2, 401}};
minimACE[F, A, 10, 201 * 401 + 1, True, True]
```

```
Commencing minimization
10% of the maximum of operations completed
20% of the maximum of operations completed
30% of the maximum of operations completed
40% of the maximum of operations completed
50% of the maximum of operations completed
60% of the maximum of operations completed
70% of the maximum of operations completed
80% of the maximum of operations completed
90% of the maximum of operations completed
100% of the maximum of operations completed
```

$$\left\{\left\{0, -\frac{1}{50}\right\}, \frac{1}{2500}\right\}, \left\{\left\{\frac{1}{100}, \frac{1}{100}\right\}, \frac{1}{5000}\right\}, \left\{\left\{-\frac{1}{100}, \frac{1}{100}\right\}, \frac{1}{5000}\right\}, \left\{\left\{\frac{1}{100}, -\frac{1}{100}\right\}, \frac{1}{5000}\right\},$$
$$\left\{\left\{-\frac{1}{100}, -\frac{1}{100}\right\}, \frac{1}{5000}\right\}, \left\{\left\{0, \frac{1}{100}\right\}, \frac{1}{10\,000}\right\}, \left\{\left\{\frac{1}{100}, 0\right\}, \frac{1}{10\,000}\right\}, \left\{\left\{-\frac{1}{100}, 0\right\}, \frac{1}{10\,000}\right\}, \left\{\left\{0, -\frac{1}{100}\right\}, \frac{1}{10\,000}\right\}, \{\{0, 0\}, 0\}\right\}$$

Figure 3: Example of minimACE use.

The advantage of minimACE is its flexibility. It can minimize functions depending on an arbitrary number of variables and it lets the user easily choose how many values they want to see, the region of the $n-$dimensional parameter space they want to explore, and the number of points where the function is to be calculated, this way controlling the degree of detail of that exploration. Instead of using Mathematica's *Table* function, we programmed a system of counters to sequentially generate all the combinations of values for the variables according to the user's instructions, while comparing and saving only the number of lists {variables, function} requested, in order to save memory and computation time.

We found minimACE was very helpful to find appropriate values of the parameters in our experiment to reduce Θ. It allowed us to make wide range explorations to cover big regions of our space of parameters to find zones of interest. After that, we made more detailed analysis of those zones again with minimACE or applied other techniques, such as using Mathematica's optimization functions like *FindMinimum* or fixing some variables to make plots varying the rest of them. We would also like to comment that minimACE helped us to find patterns or make observations that guided us to some other results in our experiment. For example, even before realising that it was only the difference $\epsilon - \mu_0$, instead of separately ϵ and μ_0, what mattered in our experiment, we were able to see that by observing the results of minimACE.

3.1.3 Interesting cases

There is a couple of particular situations that we believe to be worth a brief independent analysis.

Equal temperatures case. This is the first situation we studied because in this case Θ gets a particularly simple form. However, we will later show that this may not be the only reason why

this is an interesting case. If we set $T_1 = T_2 \equiv T_0$, then from equation (2.19) we get

$$\mathcal{K} = \frac{k_B}{h}\int_{-\infty}^{\infty}\mathcal{T}(E)(f_1(E)-f_2(E))\left[\frac{\cancel{E}-\mu_2}{k_BT_0}-\frac{\cancel{E}-\mu_1}{k_BT_0}\right]dE = \\ = \frac{eV}{hT_0}\int_{-\infty}^{\infty}\mathcal{T}(E)(f_1(E)-f_2(E))\,dE \overset{(2.13)}{=} \frac{\mathcal{I}_1 V}{T_0} \tag{3.5}$$

where we have used $\mu_1 - \mu_2 = eV$. Substituting this into (3.1) yields

$$\Theta = \frac{eV}{2ek_BT_0}\frac{\mathcal{S}_1}{\mathcal{I}_1} =: \theta \tag{3.6}$$

where we do not cancel e in the fraction to keep energy dimensions in our parameters.

Small Γ limit. As we mentioned above, Γ determines the sharpness of the transmission probability curve. Let us consider the integral over $\mathbb{R}$ of a function $F(E)$ multiplied by $\mathcal{T}(E)$. If F is continuous at ϵ and has a global "good behaviour" (which all our functions have), provided Γ is small enough, we may approximate $F(E) \approx F(\epsilon)$ for $E \approx \epsilon$ and write

$$\int_{-\infty}^{\infty}\mathcal{T}(E)F(E)\,dE \approx \int_{-\infty}^{\infty}\mathcal{T}(E)F(\epsilon)\,dE \overset{(3.3)}{=} 4\gamma_1\gamma_2\frac{\pi\Gamma}{2}F(\epsilon) \overset{g=1/2}{=} \frac{\pi\Gamma}{2}F(\epsilon) \tag{3.7}$$

If we apply this approximation to estimate the integrals in (2.13) and (2.16), already for $g = 1/2$, we find

$$\mathcal{I}_1 \approx \frac{\pi\Gamma}{2}\frac{e}{h}(f_1(\epsilon)-f_2(\epsilon)) \tag{3.8}$$

$$\mathcal{K} \approx \frac{\pi\Gamma}{2}\frac{k_B}{h}(f_1(\epsilon)-f_2(\epsilon))\left[\frac{\epsilon-\mu_2}{k_BT_2}-\frac{\epsilon-\mu_1}{k_BT_1}\right] \tag{3.9}$$

The case of $\mathcal{S}_1$ is slightly different, because since the second term in the integrand of (2.20) the factor is not $\mathcal{T}$ but $\mathcal{T}(1-\mathcal{T})$ we need to split the integral in two before applying (3.7):

$$\mathcal{S}_1 = \frac{e^2}{h}\int_{-\infty}^{\infty}\mathcal{T}[f_2(1-f_2)+f_1(1-f_1)]\,dE + \frac{e^2}{h}\int_{-\infty}^{\infty}\mathcal{T}(1-\mathcal{T})(f_2-f_1)^2\,dE \approx \\ \approx \frac{e^2}{h}[f_2(\epsilon)(1-f_2(\epsilon))+f_1(\epsilon)(1-f_1(\epsilon))]\int_{-\infty}^{\infty}\mathcal{T}(E)\,dE + \frac{e^2}{h}(f_2(\epsilon)-f_1(\epsilon))^2\int_{-\infty}^{\infty}\mathcal{T}(E)(1-\mathcal{T}(E))\,dE = \\ = \frac{e^2}{h}[f_2(\epsilon)(1-f_2(\epsilon))+f_1(\epsilon)(1-f_1(\epsilon))]\frac{\pi\Gamma}{2} + \frac{e^2}{h}(f_2(\epsilon)-f_1(\epsilon))^2\frac{\pi\Gamma}{4} = \\ = \frac{\pi\Gamma}{2}\frac{e^2}{h}\left[f_2(\epsilon)(1-f_2(\epsilon))+f_1(\epsilon)(1-f_1(\epsilon))+\frac{(f_2(\epsilon)-f_1(\epsilon))^2}{2}\right] \tag{3.10}$$

Using (3.8), (3.9) and (3.10) in (3.1), recalling that $\mu_1 = \cancel{\mu_0}^{0} + eV/2$ and $\mu_2 = \cancel{\mu_0}^{0} - eV/2$ and operating, we find

$$\Theta \approx \frac{1}{2}\left[\frac{\epsilon+eV/2}{k_BT_2}-\frac{\epsilon-eV/2}{k_BT_1}\right]\frac{(f_1(\epsilon)+f_2(\epsilon))\left(1-\frac{f_1(\epsilon)+f_2(\epsilon)}{2}\right)}{(f_1(\epsilon)-f_2(\epsilon))} =: \xi \tag{3.11}$$

Apart from giving an analytical expression, the approximation in (3.11) has an important advantage. Despite still having 4 degrees of freedom (k_BT_1, k_BT_2, eV and ϵ), we can think ξ as a function with only two arguments. This is because since the fractions in the left bracket are precisely the arguments of Fermi functions, it is only the value of those fractions what determines ξ. Since those fractions can take any real value when we let our parameters free, we can consider them as our new coordinates, varying in all $\mathbb{R}$. In other words, we can replace (3.11) by

$$\xi(x_1,x_2) = \frac{1}{2}[x_2-x_1]\frac{(f(x_1)+f(x_2))\left(1-\frac{f(x_1)+f(x_2)}{2}\right)}{(f(x_1)-f(x_2))} \quad \text{where} \quad (x_1,x_2)\in\mathbb{R}^2,\ f(x)=\frac{1}{1+e^x} \tag{3.12}$$

and even more, taking into account that f is a diffeomorphism between $\mathbb{R}$ and the interval $(0,1)$, we can use it to make change of coordinates

$$\mathbb{R}^2 \xrightarrow{\quad\sim\quad} (0,1)\times(0,1) \\ (x_1,x_2) \longmapsto (y_1 = f(x_1), y_2 = f(x_2)) \tag{3.13}$$

which transforms (3.12) into

$$\xi(y_1,y_2)=\frac{\ln\left(\frac{1}{y_2}-1\right)-\ln\left(\frac{1}{y_1}-1\right)}{2}\frac{(y_1+y_2)\left(1-\frac{y_1+y_2}{2}\right)}{y_1-y_2}=$$
$$=\frac{1}{4}\ln\left(\frac{\frac{1}{y_2}-1}{\frac{1}{y_1}-1}\right)\frac{(y_1+y_2)(2-y_1-y_2)}{y_1-y_2}\quad\text{where}\quad(y_1,y_2)\in(0,1)\times(0,1)\qquad(3.14)$$

In spite of being analytical, the expression in (3.14) is not easy to study at all, due to the transcendental equations that arise from almost any manipulation we can try with it. It is not that difficult to prove that when y_1 and y_2 tend to the same value, ξ tends to 1. However, this approximation has an obvious advantage, which is that it allows us to visualize all possible configurations with a simple *Plot3D* in Mathematica, which is shown in 4.

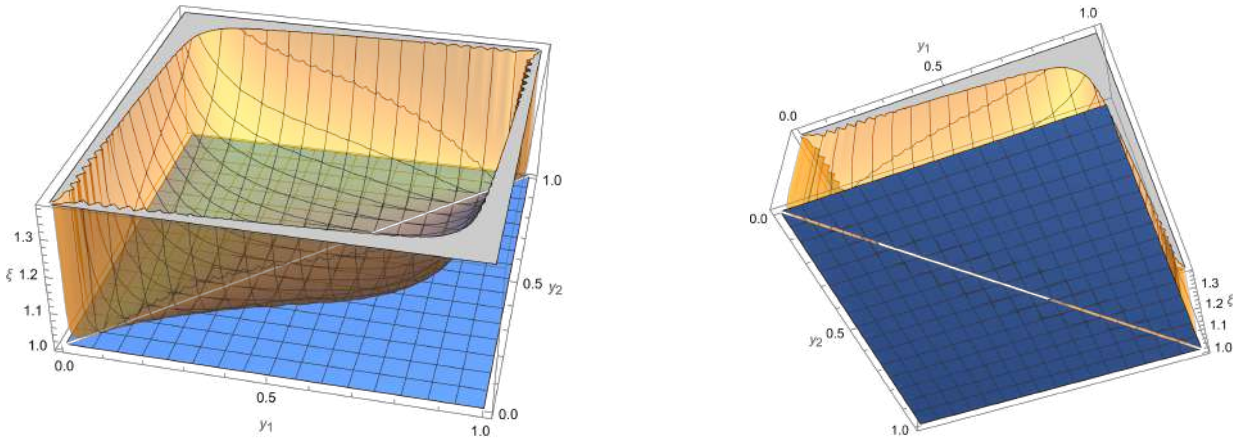


Figure 4: Plot3D of $\xi(y_1,y_2)$ from two different angles. The blue plane is the graph of constant 1 function.

In view of the representation of $\xi(y_1,y_2)$ we conclude that in the small Γ limit, our TUR cannot be violated.

The last result has another important consequence due to a certain equivalence between small Γ and high temperatures. After making a change of variable $E=\Gamma w$, it is easy to check that for Θ, multiplying T_1 and T_2 by the same positive quantity λ gives the same result as dividing Γ, eV and ϵ by λ. This means that increasing both temperatures takes us to a small Γ configuration, in which we have already checked that violating TUR is not possible. In conclusion, high temperatures regime also prevent us from escaping TUR.

3.2 Escaping TUR

After the previous discussion, we will now seize our results and tools to explore the possibility to violate TUR in its original form. Unless otherwise mentioned, we will use Γ as our reference for energies. Since Γ factors comming out of the integrals cancel in Θ, we do not need to worry about them. Consequently, from now on, any energy quantity without units must be understood as its rate over Γ.

Investigating the low temperatures region of our space of parameters, and guided by miniMACE, we found a TUR violation. The minimum of Θ was found, with the help of FindMinimum, close to the point $(k_BT_1=0.0866667, k_BT_2=0.0866667, eV=0.23, \epsilon=0)$, for which $\Theta(0.0866667, 0.0866667, 0.23, 0)=0.997092$. To get an idea of what happens there with Θ, we make a Plot3D fixing the temperatures and find what we show in figure 5. To get it, we used a function to calculate θ, (see equation (3.6)) instead of the general Θ, to simplify calculations.

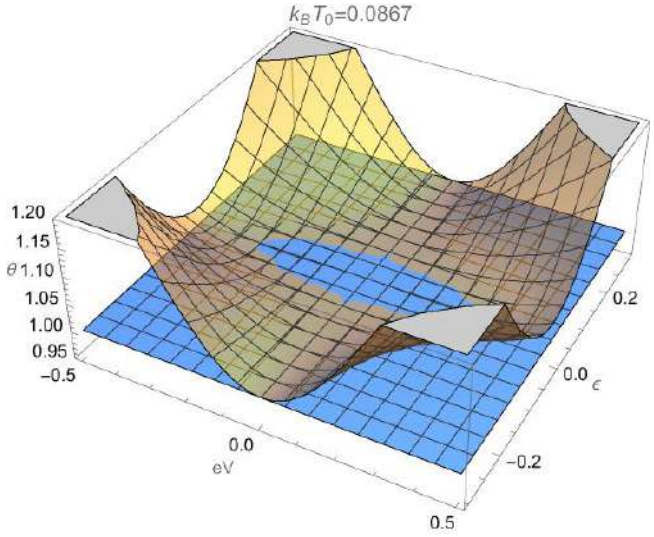
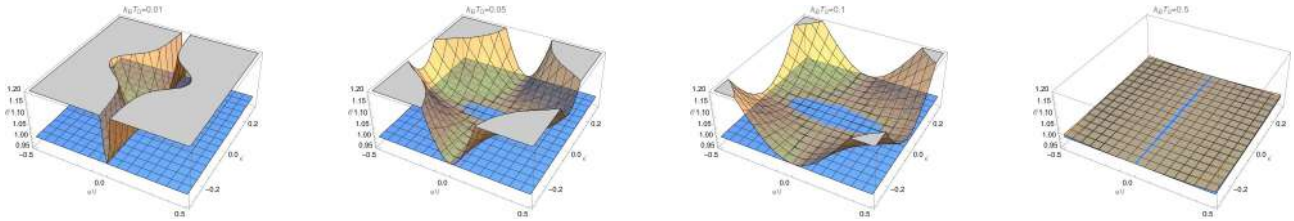


Figure 5: Plot3D of Θ close to the minimum.

In order to explore the influence of the common temperature, we repeat this Plot3D for different values of $k_{\mathrm{B}}T_0$ and get 6.



(a) $k_{\mathrm{B}}T_0 = 0.01$ (b) $k_{\mathrm{B}}T_0 = 0.05$ (c) $k_{\mathrm{B}}T_0 = 0.1$ (d) $k_{\mathrm{B}}T_0 = 0.5$

Figure 6: Plot3D of $\theta(k_{\mathrm{B}}T_0, eV, \epsilon)$ for different temperatures.

In order to understand why in this region we find a minimum in Θ, let us go back to the $T_1 = T_2 = T_0$ case. Here we had a simpler expression for out TUR indicator:

$$\theta = \frac{eV}{2ek_{\mathrm{B}}T_0}\frac{\mathcal{S}_1}{\mathcal{I}_1}$$

and if we recall the integral expression for noise, (2.20), we find that it is composed by two terms, known respectively as *thermal noise* and *shot noise*

$$\mathcal{S}_1 = \frac{e^2}{h}\int_{-\infty}^{\infty} \mathcal{T}\left\{[f_1(1-f_1) + f_2(1-f_2)]\right\}\, dE + \frac{e^2}{h}\int_{-\infty}^{\infty} \mathcal{T}(1-\mathcal{T})(f_2-f_1)^2\, dE \tag{3.15}$$

We have found a combination of our parameters in which:

- Charge current is high because the maximum of transmission probability is in the middle of the interval of energies in which $f_1 > f_2$, and the fast fall of Fermi functions places almost the whole area under $\mathcal{T}$ inside of the region between them.
- Thermal noise is low because low temperatures make $f_1(1-f_1)$ and $f_2(1-f_2)$ be near 0 almost everywhere due to the fast decay of Fermi functions.
- Shot noise is low too thanks to the fact that $\mathcal{T}(1-\mathcal{T})$ is small in most of the interval in which $(f_2-f_1)^2$ is high.

This observations can be easily ilustrated by means of plots like those in figure 7.

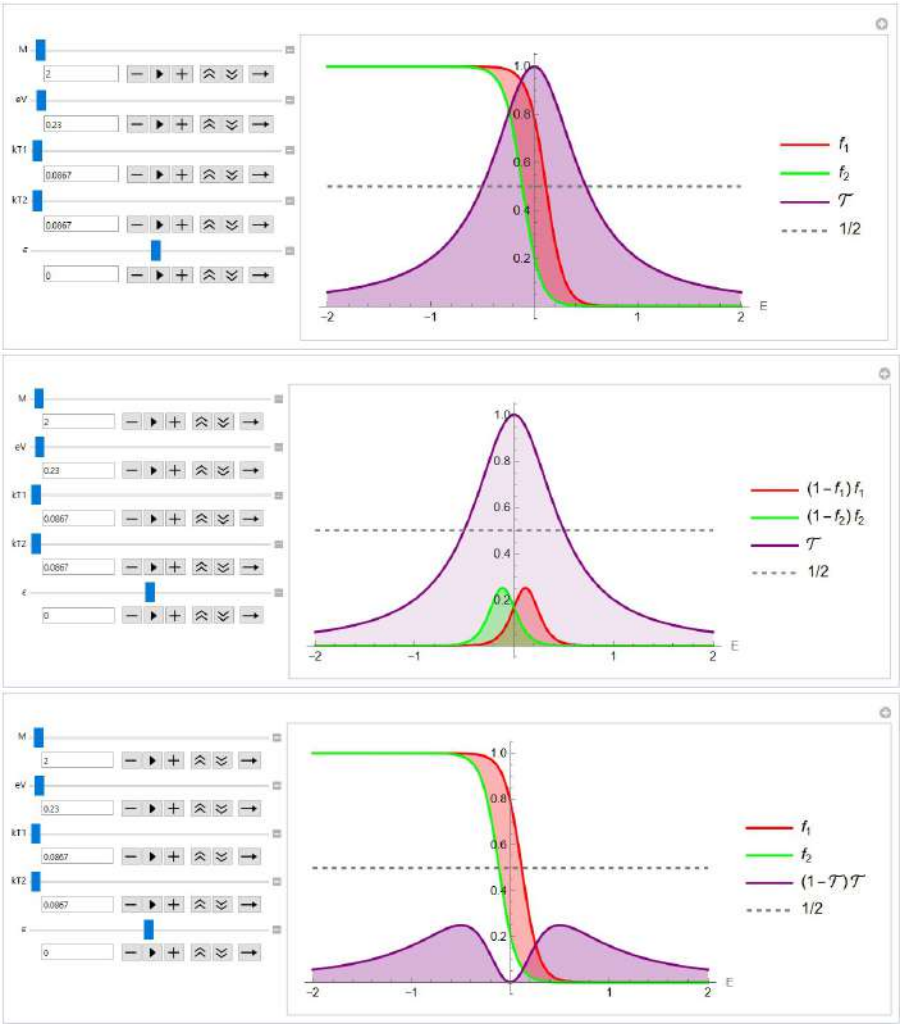


Figure 7: Behaviour of functions in the integrands for the parameters that let us escape the TUR.

This special conditions quickly disappear when we introduce a temperature gradient, and as a result, it is difficult to escape the TUR when $T_1 \neq T_2$. This is what we show in figure 8, in which we fix $k_B T_1$ to plot Θ for varying eV and $k_B T_2$.

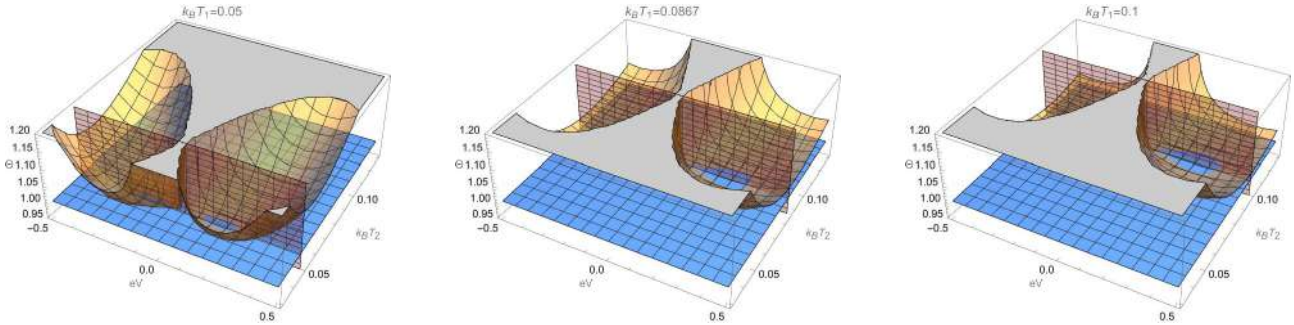


Figure 8: Temperature gradient makes TUR hold.

3.3 Thermodynamic description of the experiment

After the quantum description of our two-terminal scattering experiment, we can now investigate its "thermodynamic behaviour" and its possible applications. In order to avoid confusion, let us review our sign convention:

- Currents towards the system (scatterer) are positive, and currents towards the leads are negative.

- The voltage V between leads 1 and 2 is $V = V_1 - V_2$, and so, it will be positive when reservoir 1 has a higher potential than 2.

3.3.1 Verification of First and Second Laws of Thermodynamics

We can start our thermodynamic analysis of our experiment by checking that it obeys the First Law of Thermodynamics. Following our criteria, the total heat rate it takes from the reservoirs must be the sum of the heat currents flowing from both terminals, which can be calculated as

$$\begin{aligned}\mathcal{J}_1 + \mathcal{J}_2 &\overset{(2.16)}{=} \frac{1}{h}\int_{-\infty}^{\infty}(E-\mu_1)\mathcal{T}(f_1-f_2)\,dE + \frac{1}{h}\int_{-\infty}^{\infty}(E-\mu_2)\mathcal{T}(f_2-f_1)\,dE = \\ &= \frac{1}{h}\int_{-\infty}^{\infty}(\mu_2-\mu_1)\mathcal{T}(f_1-f_2)\,dE = \frac{-eV}{h}\int_{-\infty}^{\infty}(f_1-f_2)\mathcal{T}\,dE \overset{(2.13)}{=} -\mathcal{I}_1 V \end{aligned} \tag{3.16}$$

or equivalently

$$\mathcal{J}_1 + \mathcal{J}_2 + \mathcal{I}_1 V = 0 \tag{3.17}$$

which is simply the mathematical expression of the First Law of Thermodynamics. Indeed, in a time lapse Δt, a charge $\mathcal{I}_1\Delta t$ flows from lead 1 to the scartterer, and thanks to the DC current conservation (2.8), $\mathcal{I}_1 + \mathcal{I}_2 = 0$, the same charge passes from the scatterer to lead 2, which results in a electrical potential energy loss of $\mathcal{I}_1\Delta t V$. Then, $\mathcal{I}_1 V$ is the electrical power dissipated into the system. Thus, equation (3.17) may be obtained by derivating $Q + W = \Delta U$ with respect time, because since the scatterer has fixed properties, its internal energy is constant. This result is an important verification, and will also be useful when considering this experiment as a thermal machine.

We can as well check that our results are consistent with the Second Law of Thermodynamics, by proving that the entropy generation rate for our thermodynamic universe, $\mathcal{K}$, is always positive. To do this, let us rewrite its integral expression (2.19):

$$\begin{aligned}\mathcal{K} = -\frac{\mathcal{J}_1}{T_1} - \frac{\mathcal{J}_2}{T_2} &\overset{(2.15)}{=} \frac{k_B}{h}\int_{-\infty}^{\infty}\mathcal{T}(E)(f_1(E)-f_2(E))\left[\frac{E-\mu_2}{k_B T_2} - \frac{E-\mu_1}{k_B T_1}\right]dE = \\ &= \frac{k_B}{h}\int_{-\infty}^{\infty}\mathcal{T}(E)\left(f\left(\frac{E-\mu_1}{k_B T_1}\right) - f\left(\frac{E-\mu_2}{k_B T_2}\right)\right)\left[\frac{E-\mu_2}{k_B T_2} - \frac{E-\mu_1}{k_B T_1}\right]dE \end{aligned} \tag{3.18}$$

where again $f(x) = (1+\exp x)^{-1}$. Since f is a monotone decreasing function, we have

$$\frac{E-\mu_2}{k_B T_2} > \frac{E-\mu_1}{k_B T_1} \Leftrightarrow f\left(\frac{E-\mu_2}{k_B T_2}\right) < f\left(\frac{E-\mu_1}{k_B T_1}\right)$$

and as $\mathcal{T}(E) > 0$ for all E, the integrand in (3.18) is always positive, and therefore $\mathcal{K}$.

3.3.2 Engine

We now would like to explore the possibility to make a quantum thermal machine out of this experiment. In particular, we wish to extract some work of it. This can be achieved by making electrical current flow from the low-potential contact to the high-potential one, as it is a way of generating profitable electrical energy.

Let us assume for example (the opposite situation is equivalent) that $V > 0$, so contact 1 has a higher voltage than contact 2 and we want to make current flow from 2 to 1. Then, according to our sign convention, we wish $\mathcal{I}_1$ to be negative (and automatically $\mathcal{I}_2 > 0$). This will make $\mathcal{I}_1 V < 0$ and, owing to equation (3.17), $\mathcal{J}_1 + \mathcal{J}_2 > 0$, which means that the system is absorbing heat and using that energy to do work, this is, acting as an engine.

However, getting that behaviour is not easy. If we calculate the current $\mathcal{I}_1$ for the parameters that let us violate the TUR we find

$$\mathcal{I}_1(k_B T_1 = 0.0867, k_B T_2 = 0.0867, eV = 0.23, \epsilon = 0) \approx 0.21\frac{e\Gamma}{h} > 0 \tag{3.19}$$

so it is not working as an engine, but as an electric warmer. If we go back to the first plot in figure 7, it is obvious that the integral (2.13) will be positive, since the integrand is positive for all E. The only way to get an negative current is by introducing a gradient of temperatures $T_1 > T_2$ and adjusting ϵ to place and adequate region between Fermi functions under the central part of the curve of $\mathcal{T}(E)$, as we show in 9.

On the other hand, the mentioned temperature gradient prevents us from escaping TUR. Figure shows that the engine condition does not occur where TUR is violated. Until now, we have not found a configuration to meet both objectives at the same time.

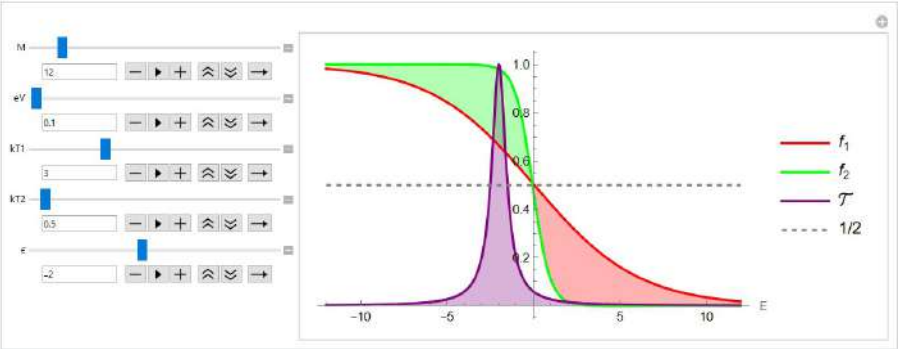


Figure 9: Fermi functions and transmission probability in an engine configuration.

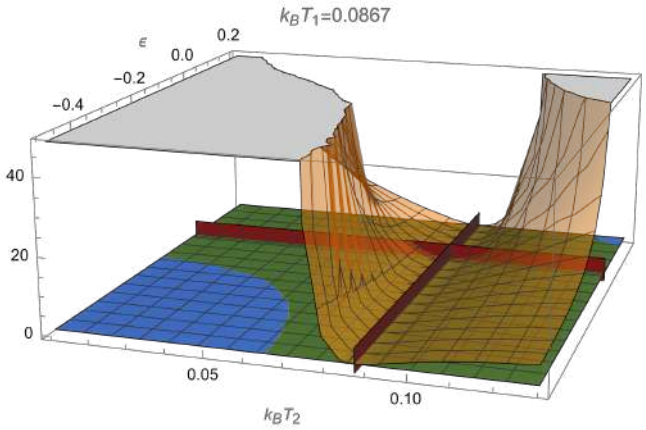


Figure 10: Engine condition and violation of TUR do not occur at the same points in our parameter space. In this plot, blue plane marks 1, the yellow graph is Θ and the green graph is $\mathcal{I}_1 + 1$, so that the objective would be finding a point in which both graphs are under the blue plane. However, we see this does not happen, at least in this region of our parameter space. The red cross marks the point where TUR was previously violated.

4 Conclusions

In this work, we have applied the so-called Landauer-Büttiker formalism to study the electronic transport through a quantum dot. Its quantum approach, after manipulations and statistics, lead to expressions for charge current and thermodynamic magnitudes that offer a model of the problem that, while based in Quantum Mechanics, involve measurable macroscopic properties of electrons reservoirs. The reason why this phenomenon is interesting is that, by virtue of its quantum nature, it makes it possible to violate some limitations that appear in classical systems. This kind of phenomena offer us an opportunity to, by seizing this sort of "quantum advantages", devise new thermal machines that may be able in the future to surpass current ones.

The main difficulties of our problem are the amount of degrees of freedom we have to consider and that the expressions we had to work with are problematic when applying analytical techniques of calculation and optimization. That is the reason why we had to search for alternative techniques to study our system, specially through computing tools and graphical methods.

Although we have tried to exploit the possibilities of our system to the most, it still has many aspects to study and open questions, some of which we are still investigating:

- The possibility to try more complex and effective calculation, exploration and optimization methods (either analytical, numerical, graphical...) may give us a deeper understanding of the behaviour of our system, which we could use to search for more possibilities to break the TUR.

- Even without violating TUR, searching for ways to build an engine that approaches the $\Theta = 1$ may yield important theoretical or practical results. Analysing the efficiency of this engines, may give interesting implications too.
- Here we have explored the way to turn our system into an engine, but we could try to build different types of thermal machines, such as refrigerators and heat pumps.
- In this work, we have only considered Clausius definition of entropy. However, Statistical Physics and Information Theory provide alternative definitions for entropy. We have already investigated the relations between those definitions in this experiment, but we still have not completed that analysis. The importance of this question lies in the information it may give us about the reasons why TUR can be violated by this system.

Moreover, this is just a possible approach to the problem of violating TUR. Changing parts of our system can yield better results when trying to escape from it. For example, metal contacts can be replaced by semiconductors, or the quantum dot may be substituted by some kinds of quantum potentials or different objects. If more complicated systems allow to reach lower values for Θ or to violate TUR in more general conditions, our difficulties to build a TUR-breaking engine might be overcome.

Acknowledgments

Acknowledge support and useful discussions:
This work was supported by the SURF@IFISC fellowship.

References

[1] Javier Argüello-Luengo, David Sánchez, and Rosa López. "Heat asymmetries in nanoscale conductors: The role of decoherence and inelasticity". In: *Physical Review B* 91.16 (2015), p. 165431.

[2] Andre C Barato and Udo Seifert. "Thermodynamic uncertainty relation for biomolecular processes". In: *Physical review letters* 114.15 (2015), p. 158101.

[3] Anton Bruch, Caio Lewenkopf, and Felix von Oppen. "Landauer-Büttiker approach to strongly coupled quantum thermodynamics: Inside-outside duality of entropy evolution". In: *Physical review letters* 120.10 (2018), p. 107701.

[4] Jordan M Horowitz and Todd R Gingrich. "Thermodynamic uncertainty relations constrain non-equilibrium fluctuations". In: *Nature Physics* 16.1 (2020), pp. 15–20.

[5] Thomas Ihn. *Semiconductor Nanostructures: Quantum states and electronic transport*. OUP Oxford, 2009.

[6] Michael V Moskalets. *Scattering matrix approach to non-stationary quantum transport*. World Scientific, 2011.

5 Mathematica code

5.1 Manipulate

This code creates a *Manipulate* that plots together the Fermi functions and the transmission probability. It is possible to introduce an electrical energy gradient, with contact 1 with a higher voltage. The maximum of the transmission probability can be changed with ϵ and M controls the range of energies to represent. It is advisable to reduce M when any of the temperatures is very low, in order to avoid precision problems.

```
(*Definitions*)
f1[kT1_, \[Mu]0_, eV_, w_] := 1/(1 + Exp[(w - \[Mu]0 - eV/2)/kT1]);
f2[kT2_, \[Mu]0_, eV_, w_] := 1/(1 + Exp[(w - \[Mu]0 + eV/2)/kT2]);
Ts[\[Epsilon]_, w_] := 1/(4 (w - \[Epsilon])^2 + 1);
(*Plot*)
Manipulate[
 Plot[{f1[kT1, 0, eV, w], f2[kT2, 0, eV, w], Ts[\[Epsilon], w],
   1/2}, {w, -M, M},
  PlotStyle -> {Red, Green, Purple, Directive[Gray, Dashed]},
  Filling -> {{1 -> {{2}, {Directive[Green, Opacity[0.3]],
       Directive[Red, Opacity[0.3]]}}, {3 -> {Axis,
      Directive[Purple, Opacity[0.3]]}}},
  PlotLegends -> {\!\(TraditionalForm\`
SubscriptBox[
StyleBox["f", "TI"], "1"]\), \!\(TraditionalForm\`
SubscriptBox[
StyleBox["f", "TI"], "2"]\), \[ScriptCapitalT], "1/2"},
  AxesLabel -> {"E", None}],
 {{M, 50}, 0.1, 100}, {eV, 0, 10}, {{kT1, 2}, 0.01, 10}, {{kT2, 2},
  0.01, 10}, {\[Epsilon], -20, 20}
 ]
```

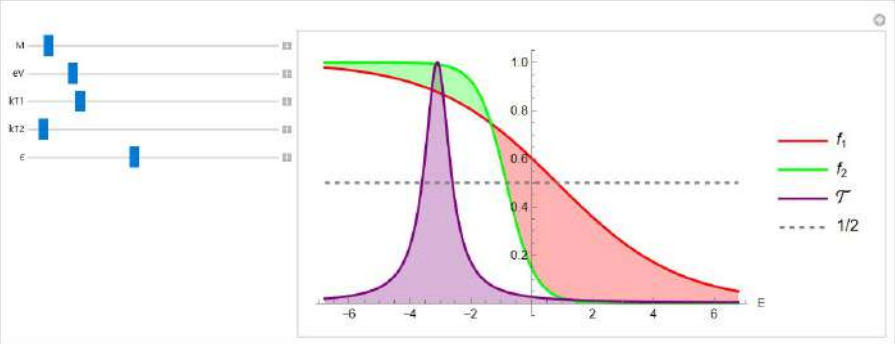


Figure 11: Example of the Manipulate execution.

5.2 minimACE code

```
    minimACE[F_, AA_, m_, tope_, comprob_ : True, marcas_ : False] :=
 Module[
  (*** LOCAL VARIABLES ***)
  {A = AA, At, actucont, actulista, continuarActualizacion = True, n,
   X, X0, Xm, J, j, k, S, R, V, w, RV, iter = 0, parc, r},
  (*** FUNCTION BODY ***)
  If[comprob,(** Some verifications **)
   (*Checking AA is a real matrix*)
   If[TrueQ[MatrixQ[AA]], At = Transpose[AA],
    Print["Second argument of minimACE must be a Matrix"];
    Return["minimACE::MatrixError"]];
   If[TrueQ[Dimensions[AA][[2]] == 3], ,
    Print["The matrix must have 3 columns"];
    Return["minimACE::MatrixError"]];
   If[TrueQ[MatrixQ[AA, RealValuedNumberQ]], ,
    Print["Second argument of minimACE must be a matrix of real \
numbers"]; Return["minimACE::errorMatriz"]];
   (*We need last column of A to have positive integers*)
   At[[3]] = Floor[At[[3]]];
   A = Transpose[At];(*We need those integers to be 1 or greater*)
   j = At[[3]]; n = Length[A];
   For[k = 1, k <= n, k++,
    If[TrueQ[j[[k]] >= 1], , Print["Error con número de puntos"];
     Return["minimACE::errorNumPuntos"]]];
   (*Lets check that applying F to the arguments yields a real number*)
   If[TrueQ[RealValuedNumberQ[F @@ At[[1]]]] &&
     TrueQ[RealValuedNumberQ[F @@ At[[2]]]], ,
    Print[
     "The function applied to initial or final values of variables \
did't yield a real number"];
    Return["minimACE::errorArgsFuncion"]];
   (** Verifications complete **),
   At = Transpose[AA]; A = Transpose[At]; j = At[[3]]; n = Length[A];];
  (** Auxiliary functions definitions **)
  (*Function for updating counters*)
  actucont[L_, l_] := Module[{X1 = L, i = 1, m1 = Length[l]},
    While[i <= m1 && X1[[i]] == l[[i]], X1[[i]] = 0; i++;];
    If[i <= m1, X1[[i]]++, continuarActualizacion = False];
    Return[X1]];
  (*Function for updating the list with lower values of F found*)
  actulista[L_, V1_, vbles_, f_, n1_] := Module[{i = 1, L2, V2},
    While[i <= n1 && f < V1[[i]], i++];
    If[i == 1, Return[{L, V1}], V2 = Drop[Insert[V1, f, i], 1];
     L2 = Drop[Insert[L, {vbles, f}, i], 1]; Return[{L2, V2}]]];
  (** End of auxiliary functions definitions **)
  (*n=number of variables, X= variables values vector,J=
  counter vector actualised until j, S=jump vector*)
  X0 = At[[1]]; X = At[[1]]; Xm = At[[2]]; j = At[[3]] - 1;
  J = Table[0, n];(*Now we'll build the jump vector*)S = Table[0, n];
  For[k = 1, k <= n, k++,
   If[j[[k]] == 0, , S[[k]] = (Xm[[k]] - X[[k]])/j[[k]];];];
  (*Initializing the vector of F values and the results list*)
  V = Table[Infinity, m]; R = Table[{0, Infinity}, m];
  (* COMMENCING CALCULATIONS*)
  Print["Commencing minimization"];
  If[\[Not] TrueQ[marcas],
   While[iter < tope && continuarActualizacion, w = F @@ X;
     RV = actulista[R, V, X, w, m]; R = RV[[1]]; V = RV[[2]];
     J = actucont[J, j]; X = X0 + J*S;
     iter++;];,(*Alternative with progress marks*)
   For[r = 1, r <= 10, r++, parc = tope*r/10;
     While[iter < parc && continuarActualizacion, w = F @@ X;
      RV = actulista[R, V, X, w, m]; R = RV[[1]]; V = RV[[2]];
      J = actucont[J, j]; X = X0 + J*S; iter++;];
     Print[10*r, "% of the maximum of operations completed"];];
   ]; If[iter == tope, Print["¡¡Maximum of operations reached!!"]];
  Return[R]]

 minimACE0::usage = "minimACE[F_,AA_,m_,tope_] returns the m lower values of F from those calculated. AA is a real matrix.
 In each row, it stores, for a variable of F, its starting value, its ending value and the number of values that
 variable must take (counting initial an ending ones). The output is a list of lists, each one formed by a list of
 values of the variables and the value of F calculated with them, so that those values of F are the m lowest ones,
 ordered from the highest to the lowest.
 tope is the maximum number of iterations for the loop used by minimACE. It is given by the user to serve as a safety
 mechanism and to enable to show the progress of the calculations.  minimACE admits two optional logical
 arguments. The first one is True by default, and decides whether or not the initial verifications of the function will
 be done. The second one is False by default, and if True, the loop for calculation is broken into 10 parts, showing a
 progress message after completing each part."
```

Power grid stability in scenarios of large renewable penetration

Marina Camps, María Martínez-Barbeito, Pere Colet
Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB)
Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain

Abstract

With the continuously increasing presence of renewable energy sources, it is necessary to study how they affect power grids. Since the generated power must be balanced with the demand in real-time, the intermittent and variable nature of these new sources can compromise the stability of the network. We present a simple model applicable to grids with enough inertia and use the Balearic power grid as a test. We explore the time evolution of the grid in case one line fails and present four representative cases of individual line failures. It will be shown that the model can handle sudden, significant perturbations and reproduces a time evolution of the system which is consistent with reality.

1 Introduction

In the ineludible scenario of climate change, renewable sources have increased in importance due to the necessity of reducing dependence on conventional energy sources. In particular, the energy transition accentuates in islands, given their subservience to external sources (importation, links to the mainland, etc.).

On the other hand, the stability of the power grid can be compromised in a framework in which renewable sources predominate over conventional ones. This is because this stability relies on the requisite of the grid to operate around a given reference frequency, which has to be maintained in real-time given the lack of large-scale energy storage capacity. This is achieved by adjusting generation to the power demand. Using traditional power plants, this is relatively easy owing to their controllable nature. When the demand increases (decreases), the frequency of the grid slows (runs above the nominal frequency), so plants with rotating generators rise (reduce) their output.

In contrast, VRES (Variable Renewable Energy Sources) are far more unpredictable, intermittent and uncontrollable, which makes more difficult the challenge of balancing the power demand with the available generation in real-time.

Another key aspect that has to be taken into account when talking about the stability of the network is inertia. Conventional power plants have a considerable amount of inertia, which helps to significantly reduce the amplitude of frequency fluctuations. On the contrary, VRES (mainly solar and wind power) have very low or even null inertia, a problem that introduces frequency fluctuations at different timescales. This added to the large fraction of VRES, their limited interconnectivity and their reduced scale make islands' grids more vulnerable to high demand variations and, therefore, more frequent failures. In a situation, thus, in which VRES progressively increase their presence, especially in *per se* more fragile networks, the global flexibility of the grid decreases.

2 Theoretical model

In the numerical model that has been used for the simulations shown in this report, the power grid is constructed as a set of nodes linked by transmission lines [1]. These nodes are treated as interacting, non-linear oscillators that should be synchronized at the same frequency. We will choose a reference frame rotating to the reference frequency of the grid, which will be, thus, normalized to zero. We will also work within the lossless line approximation.

Furthermore, we distinguish between two classes of nodes: conventional generators and consumers. Consumers correspond to substations, which include their attached load. Conventional generators obey the swing equation (1), which considers their inertial response.

$$\begin{aligned} \dot{\theta}_i &= \omega_i \\ \dot{\omega}_i &= \frac{\omega_R^2}{2\hat{H}_i\left(\omega_i+\omega_R\right)}\left(P_i^m - P_i^e\right) \end{aligned} \tag{1}$$

Power plants are thus modeled as generators with a rotor spinning at the reference frequency of the grid ω_R plus a small deviation ω_i. In (1), θ_i is the voltage phase of the node i, ω_i is the angular frequency, $\hat{H}_i = H_i P_i^G$ with H_i and P_i^G the inertia per unit power and nominal power generation. P_i^m is the supplied mechanical power, and P_i^e is the electric power at the node.

Power plants also have two frequency fluctuation control mechanisms that act at different timescales. The primary control (2) performs a change of the frequency drift in a few seconds, which stabilises its value progressively to the reference frequency. The secondary control (3) then reduces the deviation ω_i by activating the spinning reserve power. It acts in a timescale of minutes.

$$\dot{P}_i^m = \frac{1}{\tau_i}\left(P_i^s - P_i^m - \frac{P_i^c}{R_i}\frac{\omega_i}{\omega_R}\right) \tag{2}$$

$$\dot{P}_i^s = -\kappa_i \frac{\omega_i}{\omega_R} - \frac{1}{\tau_i^{ref}}\left(P_i^s - P_i^{ref}\right), \tag{3}$$

Here, τ_i is the turbine time constant, P_i^s the spinning reserve power, R_i the governor speed regulation parameter, P_i^c the primary control power, and κ_i the secondary control gain parameter.

The second term in (3) is added in order to break a degeneracy. Without this addend, grids with a single power plant would be capable of adjusting the generation to cover the demand using secondary control. On the contrary, if there were several generators involved, there would be a degeneracy and it would not be possible to define what fraction of generation corresponds to each plant. In this way, plants are forced to operate near a set point P_i^{ref}, and generation dispatch data is allowed into the model. τ_i^{ref} is then the timescale of the forcing. Based on demand data, grid operators are capable of establishing a reference power for each plant in the network.

Consumers, on their side, do not have generation nor frequency fluctuation control mechanisms, but they do present inertia (very low compared to the nodes with power plants). As a consequence, we set in this case $P_i^m = 0$ in (2) and (3), and $\hat{H}_i^{cons} << \hat{H}_i^{gen}$. Lastly, VRES are described as negative, frequency-independent demand given their essentially null inertia. The connection to the mainland is modeled as an extra power in the corresponding node and it does not depend on the frequency of the grid.

$$P_i^e = \left(1 + D_i\frac{\omega_i}{\omega_r}\right)P_i^l + \sum_j B_{ij}\sin\left(\theta_i - \theta_j\right) - P_i^{VR} \tag{4}$$

In any node, the electric power is given by (4). The first term accounts for the load in the node, and the fraction $D_i\omega_i/\omega_r$ introduces a frequency-dependent damping. The second addend represents the power transmitted from node i to node(s) j. $B_{ij} = V_{ij}^2/Z_{ij}$ is called the susceptance parameter: V_{ij} is the voltage of the line that connects node i to node j, while Z_{ij} is its impedance, estimated by the longitude of the link. P_i^{VR} is the energy generated by renewable sources if any.

As mentioned in previous sections, the main objective of the project is to study possible line failures and the potentially resulting cascade collapses of other links due to the first malfunction. This situation can be described by setting $B_{ij} = 0$, being i and j the nodes connected by the nonfunctional line. A new part has been added to the original code in order to make a line fail with a given probability rate if it runs over its voltage capacity. This probability rate is introduced in order to delay individual line failures in time and observe how their stress increases along cascade events. The possibility of reducing the capacity of the lines is also included in this new version.

This model can be applied to grids of any size. As a test, we will use the Balearic power grid, which is a paradigmatic case of a network with both an insularity condition and renewable energy sources constituting 3.5% of the total power of the system in 2015 [2].

3 Results and discussion

To begin with, parameters appearing in equations (1), (2), (3) and (4), the power supplied by the connection to the mainland, dispatch and demand data have been provided by Pere Colet and María Martínez.

In this section, we will explore four distinct cases of line failures and the consequent time evolution of the grid. First, it is necessary to clarify that we have reduced the global capacity of the lines to 80%. This is because links are built to have standard voltage capacities and power grids in islands are constructed with lines with much larger voltage limits than the maximum predicted demand. As a consequence, it is very unlikely that a line will run over its capacity. To illustrate, thus, cascade events similar to the ones that would occur on the mainland, these capacities have been lowered.

Also, we set the probability rate to 0.01; which means that, in each step of integration, a line carrying a larger voltage than it is capable of has a probability of 0.01 of failing. The integration step is 0.1s, so the overstressed lines will have faild in approximately 10s.

3.1 Cas Tresorer - Son Orlandís

Here, we present the resulting time evolution of the grid when the line connecting Cas Tresorer to Son Orlandís fails at $t = 10$s.

As shown in Fig. 1, four other lines fail as a consequence of the initial collapse. At $t = 10$s, when Cas Tresorer - Son Orlandís fails, all other lines (that will later collapse too) get more stressed and, after a transient, stabilise to a stationary value of voltage. The link between Colisseu and Son Molines (which is a particularly fragile line, since it has a large tendency to fail when other links collapse) runs largely over its capacity and it is the second line to break down. After this second collapse, the other lines stress again and consecutevly fail.

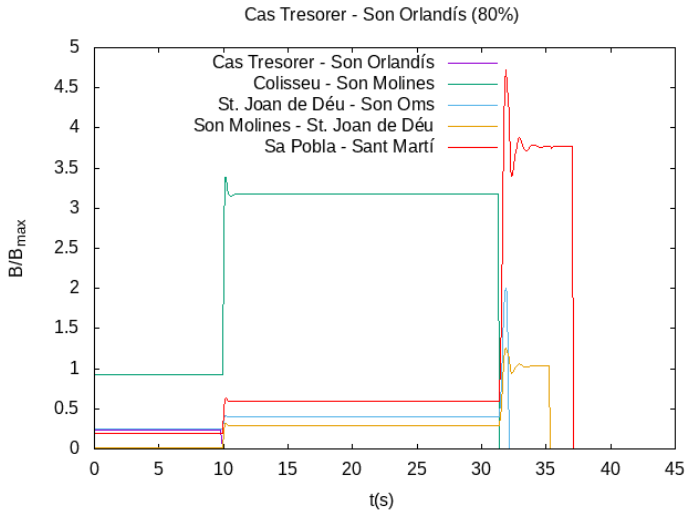


Figure 1: Relative carried capacity of the lines as a function of time. Initial failure: Cas Tresorer - Son Orlandís.

In Fig. 2, we plot the frequency of the grid over time. We can observe three transients as consequences of the first three failures at $t = 10$s (Fig. 3a), $t \approx 32$s (Fig. 3b) and $t \approx 34$s (Fig. 3c). At these times, other links fail (cascade event), so the grid experiments frequency fluctuations that increase in amplitude as more lines collapse, but it is still capable of recovering the nominal value of the frequency within seconds. When the line connecting Sa Pobla and Sant Martí fails, the frequency drops to almost -1.4Hz, indicating a large decrease of generation in the nodes for which we have represented the frequency. In this case, it will take minutes to drive the frequency back to the reference value.

Finally, in Fig. 4, we show the lines that have failed in this cascade event. We can observe that

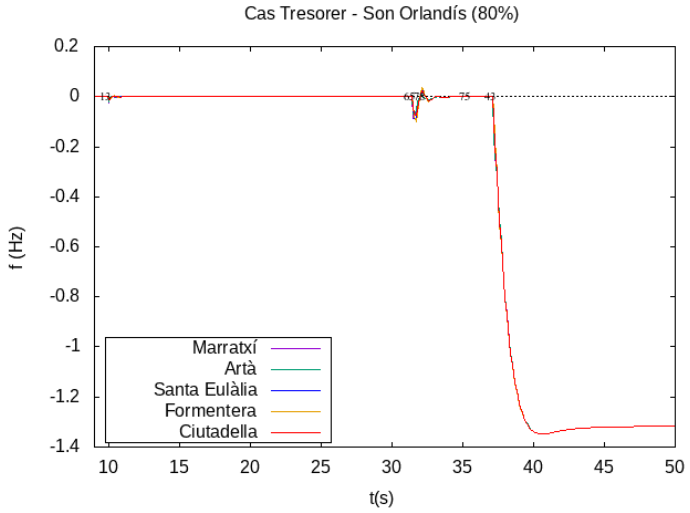
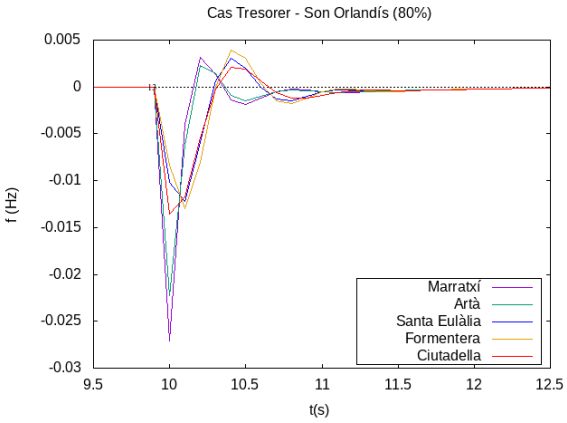
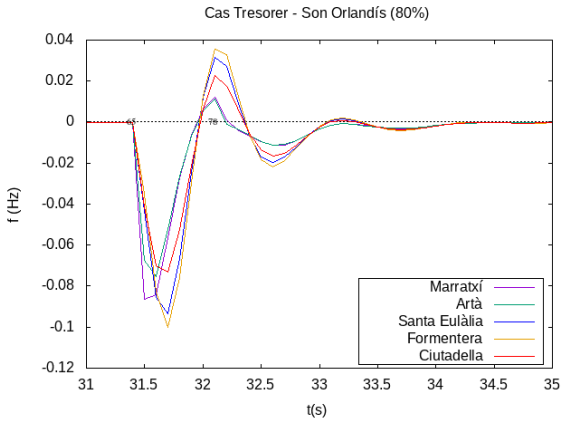


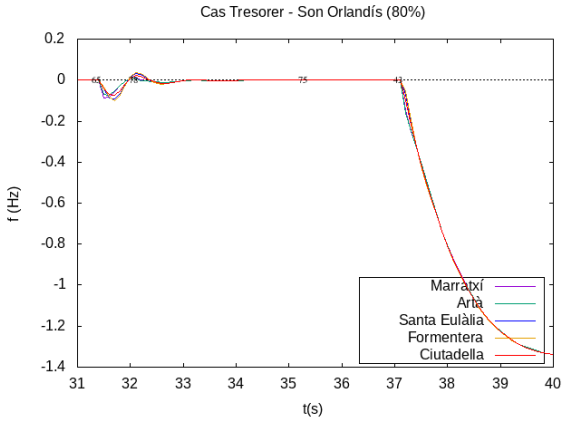
Figure 2: Frequency of different nodes as a function of time. Initial failure: Cas Tresorer - Son Orlandís.



(a) Initial frequency transient.



(b) Second frequency transient.



(c) Final frequency transient and posterior frequency drop.

Figure 3: Frequency transients over time for the initial malfunction of Cas Tresorer - Son Orlandís.

the collapsed links are not necessarily consecutive, which means that a failure in a particular line affects not only its surroundings but the whole grid.

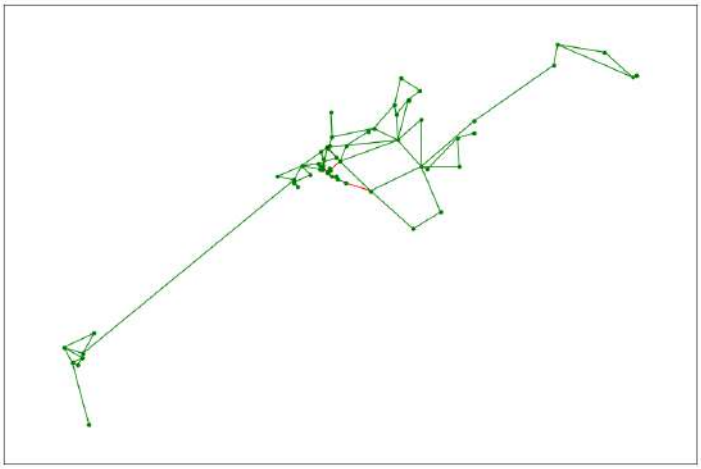


Figure 4: Balearic grid. The collapsed lines are drawn in red. Initial failure: Cas Tresorer - Son Orlandís.

3.2 Son Reus CCGT - Son Reus main node

In this case, we will explore the reaction of the grid and its time evolution when the line that connects the main node of Son Reus (power plant) with the combined cycle gas turbine (CCGT) collapses.

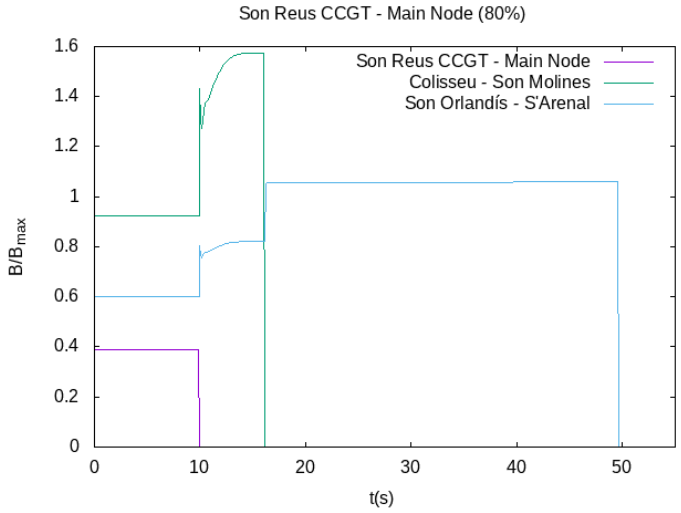


Figure 5: Relative carried capacity of the lines as a function of time. Initial failure: Son Reus CCGT - main node.

In Fig. 5, we can see that, at the breakdown time ($t = 10$s), the stress of the other two lines that will later collapse increases. The link between Colisseu and Son Molines (which also failed in the previously shown cascade event) runs over its capacity and collapses after a transient and a stabilisation. At that instant, Son Orlandís - S'Arenal also exceeds its limit and fails after approximately 35s.

On the other hand, we can see in Fig. 6 that, when the initial malfunction occurs, the frequency of the grid drops to ~ -2Hz. This is because the starting nonfunctional line connects one of the turbines of the power plant of Son Reus to its main node. With the failure, its generation is lost, leading to the observed frequency decrease.

With the third line breakdown, some nodes experiment a lower frequency drop ($t \approx 50$s), but, after that, the whole grid is synchronized to the same frequency and slowly returns to its nominal value, which will occur in a timescale of minutes as a result of the action of secondary controls.

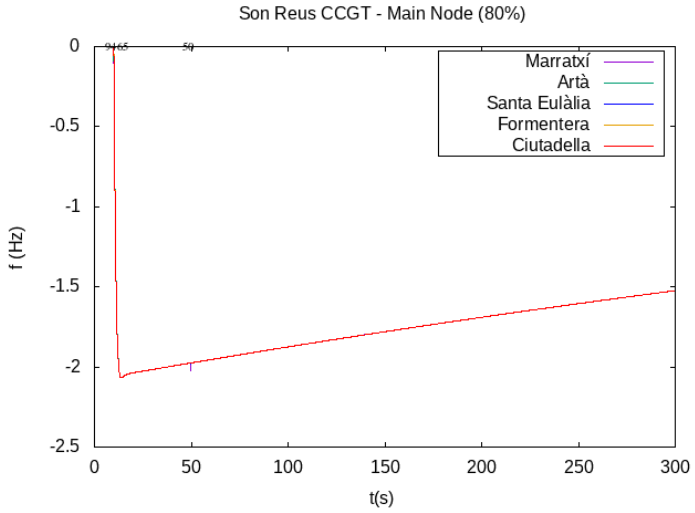


Figure 6: Frequency of different nodes as a function of time. Initial failure: Son Reus CCGT - main node.

3.3 Grid splitting

In this scenario, the initial nonfunctional line connects Mallorca with Menorca, leaving two separated networks evolving independently from one another (see Fig. 7). This is a prticularly interesting case since the frequency in Menorca will be different from the rest of the network as a consequence of this splitting.

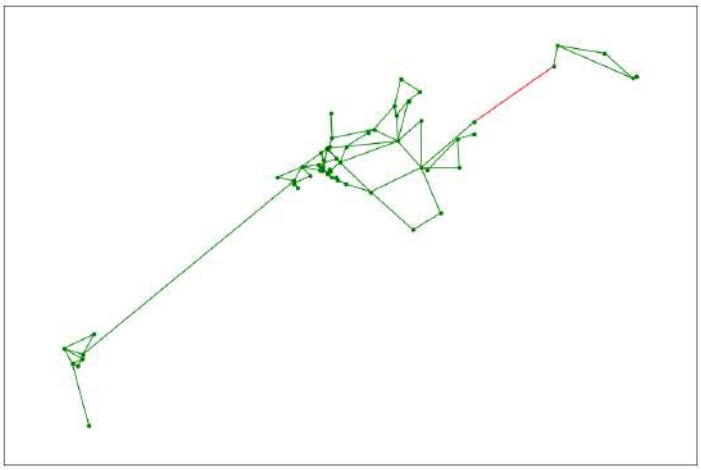


Figure 7: Grid splitting

As seen in the introduction, a frequency decrease is a synonym for a disbalance between the generation and the demand, the last one larger than the first. Consistently, the simulation shows a frequency drop in Ciutadella (to approximately -1.1 Hz) when the line failure occurs ($t = 10$ s). The reason is that a significant fraction of the generation in Ciutadella comes from Mallorca. When the line connecting the two islands fails, this generation is lost in Menorca, producing this frequency decrease.

On the other hand, the frequency in the rest of the grid has a less pronounced positive peak at the breakdown time. This is a sign of the demand reduction caused by the detachment of Menorca, which introduces, as a whole, more consumption than generation into the system.

As shown in Fig. 8, the primary control acts within the first seconds of the failure and stops the frequency drift, while, in a timescale of minutes, the secondary control (in both now independent systems) redirects the frequency to its nominal value. Letting the system evolve in time, no other

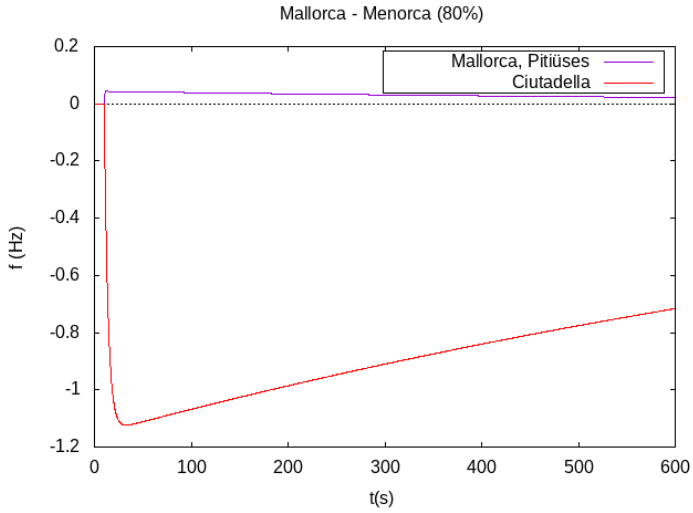


Figure 8: Grid splitting

line failures are observed.

3.4 Sta. Ponça - EC Sta. Ponça

In this last example, the line that initially fails links Santa Ponça with a nearby node, EC Sta. Ponça, which is connected, in turn, to the continent. EC Santa Ponça has a second connection with Eivissa. Therefore, the node that receives the power from the mainland has only two links: one that goes to Mallorca and another one that goes to Eivissa. In this case, the first line collapses, so all the electricity provided by the Peninsula goes to Eivissa and Formentera.

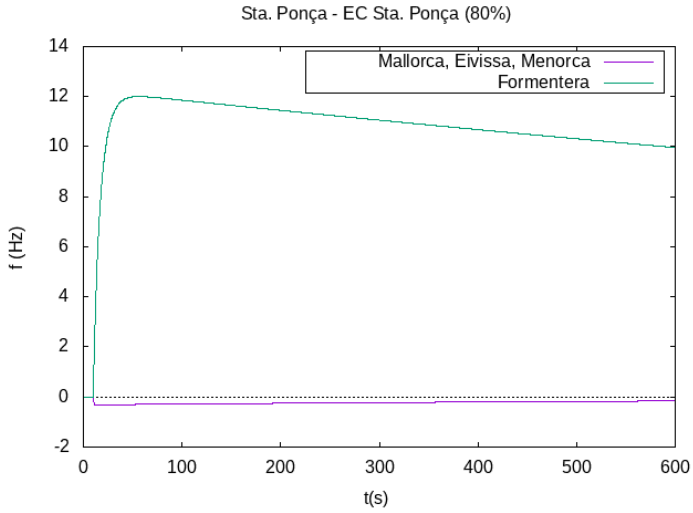


Figure 9: Grid splitting

In Fig. 9, we can see that, at the breakdown time ($t = 10$s), the frequency in Eivissa and Formentera has a high positive peak as a consequence of all the power supplied by the continent beign driven to these islands. In contrast, Mallorca and Menorca experiment a frequency decrease (much less in absolute value than in Eivissa and Formentera) because of the lack of generation. Also it stabilises back to its nominal value significantly earlier than the Pitiüses.

Letting the system evolve in time, just one other line fais due to the initial malfunction (Fig. 10). The link between Colisseu and Son Molines, as in previous cases, runs over its capacity at the breakdown time and fails after approximately 2s.

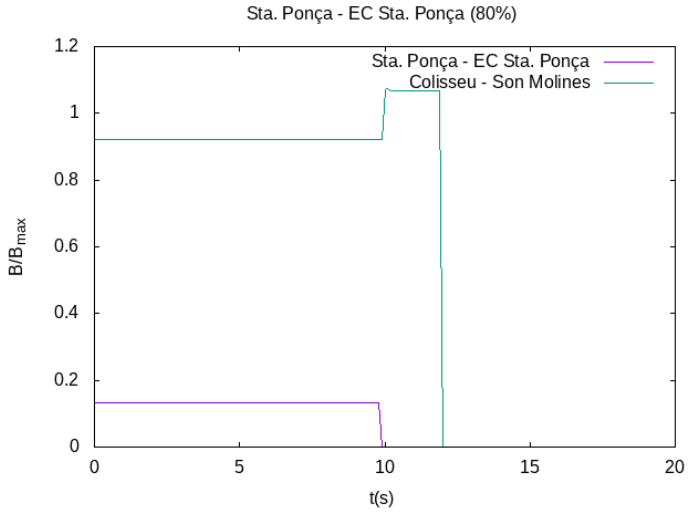


Figure 10: Grid splitting

4 Conclusions

We have presented a simple model for power grids with large renewable penetration in order to simulate the time evolution of the network under line failures. We have used the Balearic power grid as a test and it has been shown that this model allows high, sudden perturbations. It gives results consistent with reality such as cascade failures due to one inital line collapse.

We have seen that the model can correctly handle individual line failures and reproduce the resulting frequency fluctuations in any node of the grid. The time evolution of the carried power of the lines shows how, as other links fail, the system as a whole adjusts and redistributes the generated power to recover the nominal value of the frequency.

A further extension of this model could include a low probability of failure even though a line does not run over its capacity in order to account for external factors that could make a link collapse. Also, if two nodes are connected by two or more lines, the model uses just one link with an effective voltage capacity and an effective impedance. It would be useful to include multiple links to obtain more realistic results.

Acknowledgments

I want to thank to Pere Colet and María Martínez. Without their guidance and advice, none of this work would have been possible. This work was supported by the SURF@IFISC fellowship.

References

[1] Martínez-Barbeito, M., Gomila, D., & Colet, P. (2023). Dynamical model for power grid frequency fluctuations: application to islands with high penetration of wind generation. *IEEE Transactions on Sustainable Energy*.

[2] Govern de les Illes Balears (2015). Energías renovables y eficiencia energética en las Islas Baleares: Estrategias y líneas de actuación.

Quantifying higher-order interactions in social systems and their temporal evolution

Leyla Gómez, Beatriz Arregui & Sandro Meloni

Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB)
Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain

Today there is a large number of data that can be modeled or reconstructed through networks. Such networks can represent complex systems that until recently were intensively studied as connections between pairs of nodes or entities. Recent research shows that a more advanced representation is needed to understand group interactions and better encode the information that is transmitted in this type of network, giving way to what have been defined as Higher-Order-Interactions (HOI). In this work, the results obtained for a data set will be countered when considering only classical pairwise interactions, also called pairwise, and a second method based on the analysis of motifs, defined as small connected subgraphs. Additionally, the quantitative formation of a special structure of subgraphs, called cliques, will be studied throughout different aggregation times, thus evidencing how the characteristic scale of the data under study varies.

It is important to note, in an introductory way, that the science of networks is composed essentially of two parts; the structural part that seeks to understand how networks are formed, and the dynamics, corresponding to the part in charge of processing the information, controlling with it, what goes in and out of the nodes, that is, everything that is fundamentally wanted to know. Following the line of the previous description, graphs are presented as the first concept. Graphs are a mathematical tool and are defined as a set of nodes and links that usually encode interactions between pairs of units, also known as dyadic or pairwise interactions. This concept is a tool commonly used since its inception, as a way of describing complex systems.

Although it is true that network science has been developing for many years with the traditional or pairwise approach, the recent empirical evidence that has emerged since the turn of the millennium, indicates that interactions also occur with larger groups, that is, in interactions with more than two nodes, which are known as higher order interactions [1-2]. There are several examples of this class of interactions, such as face-to-face human relationships [3], collaboration networks [4], structural and functional brain networks [5-6], etc.

Understanding the effect of this class of higher-order interactions, beyond a pairwise approach, is essential when it comes to understanding how information is processed in systems with high-order interactions and also their consequences, since they can contribute significantly more information that is extracted from the traditional method. In addition, an incorrect reading of that information, by not considering the simultaneous interaction that can occur between more than two nodes, can considerably limit the ability to characterize the local structure of systems with group interactions [7-8].

To properly encode this new type of higher order interactions [1-2], new mathematical tools are needed that can describe these more complex structures. One of them is *hypergraphs* [9]. The hypergraphs correspond to a generalization of a graph, whose edges are called hyperedges, hyperlinks or hyperborders, which relate, unlike the graph, more than two nodes or vertices. These hyperlinks are responsible for encoding or describing higher order interactions between nodes.

Considering a new analysis using hypergraphs, will allow the construction of models based on more accurate data, which despite being a relatively recent approach, today has already led to the discovery of new collective phenomena and dynamic behaviors, such as social contagions, human cooperation, diffusion models and synchronization [10-11].

But there is a big problem with higher order interactions, which refers to how difficult it is usually to empirically model these hypergraphs because of the eminent lack of key data, such as the context [12]. Until today, many databases used for this analysis have been built with only pairwise interactions in mind, ignoring that reality is often much more complex, since the local density of these networks can well be explained because the entities involved usually have a shared context, that is, a higher order interaction that would allow generating this class of dense connections [13]. There is in many cases a lack of data such as data obtained from the brain [14], ecological competition data [15] and data from typical social interactions [16].

A simple example is conducting a survey of individuals to learn about their network of friends as part of the AddHealth study, carried out by the US National Longitudinal Study of Adolescent to Adult Health [17], although it is known that people usually interact naturally between groups of friends, the survey cannot reveal them since it only inquiries into relationships, without being able to represent in this way, the social dynamics between people.

The following image (figure 1) shows how friendships are measured between pairs of participants (nodes and links), even though the fundamental unit should be the groups of friends (colored areas). This reveals that there are many different representations of a higher order interaction compatible with the same data network, making it very difficult to choose its best representation [18].

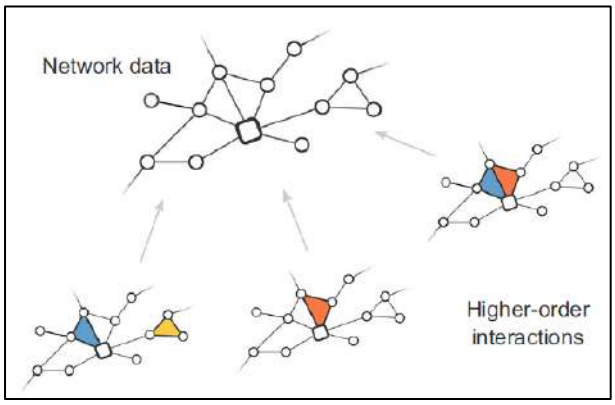


Figure 1. Diagram showing the extended network of a group of friends of the respondent (square node) of the AddHealth study [17]. The set of nodes and links shows the friendship between pairs of participants, even though the fundamental unit are friends [16].

This work seeks to obtain the higher order interactions between different groups of data with different existing methods. The first method will get and list all possible cliques that can be generated from the dataset to be studied. The second method will dig into motif analysis to find statistically validated cliques. In this way, it will be possible to counteract both results, as well as to find out how much a correct statistical validation can influence the total number of cliques of different orders. Likewise, an analysis will be made showing the variation of the results over time, thus evidencing how the different sizes of cliques evolve according to the aggregation time, with which it will be possible to appreciate the behavior of the groups. Likewise, the characteristic scale of the size of the groups will be found.

Hypergraphs

A hypergraph is a generalization of a graph, where the edges can connect more than two vertices. In graph theory, hypergraphs are represented by sets of vertices (nodes) and sets of edges (hyperlinks), where each edge is a subset of vertices. For example, a hypergraph with three nodes {a,b,c} and two hyperlinks {a,b} and {b,c} can be represented as (V,E), where V = {a,b,c} and E = {{a,b},{b,c}}.

Although hypergraphs are useful in modeling relationships between elements, as in set theory or in database theory, its use remains very open to detect hyperlinks since the network contains encoded information from which it is difficult to reconstruct. A practical example is the case of a scandal that compromises a government and its main political party. These politicians formed foundations very quickly with which millionaire projects were awarded instantly to help the community. But, even if the network of relationships between them could be correctly drawn or traced, there would still be questions that cannot be answered, that is, it would not be possible to arrive at the necessary and transcendental information, such as: Who exactly are the participants behind a foundation?, How many foundations will a specific group of individuals have created?.

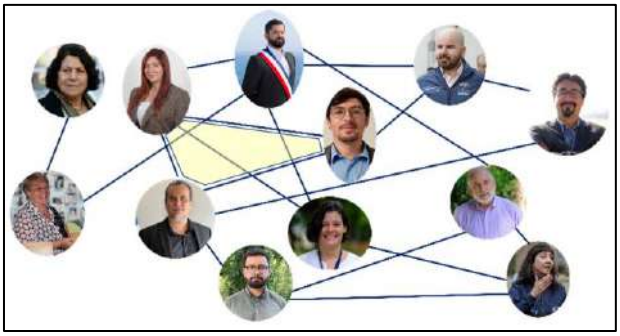


Figure 2. Network showing an alleged corruption scheme within Chilean politics. Despite knowing in advance the relationship between the participants (yellow figure), there are still certain questions to be answered that cannot be seen with a simple view.

To address the ill-posed nature of the problem and to understand how these networks are formed, various approaches have been used to reconstruct the information.

- Cliques Analysis: Finds and enumerates the minimum set of cliques that cover the network.
- Motifs analysis: they appeal to notions of randomness and generative modeling to regularize the problem.

Both methods work explicitly with the data, inferring them and reconstructing the networks, in this way, a probability can be assigned to possible higher-order data representations, which allows selecting the most appropriate one to understand their real community structure.

There are also other methods that do not fall into the previous category, that is, within the explicit analysis

of the data, such as the Bayesian Approach. This method works with an input network, which then identifies the parts of the network best explained by higher-order latent interactions. The method makes use of a Bayesian generative model to deduce the hypergraph.

Analysis of cliques

For the discovery and identification of higher order interactions, previous work experiences have often employed clique identification and enumeration [19-20]. Cliques are fully connected subgraphs, that is, sets of nodes that are fully interconnected. Although it is easy to count the cliques, to know their possible real representation within a hypergraph is necessary its decomposition, which does not turn out to be a satisfactory solution for the problem of information recovery, since a clique admits many possible decompositions, thus, difficult to know which one to choose. For example, from figure 1, the triangle admits a single 2-clique and three 1-clique [18]. In summary, the multiplicity of possible solutions implies that the recovery of higher order interactions is a poorly stated inverse problem [18].

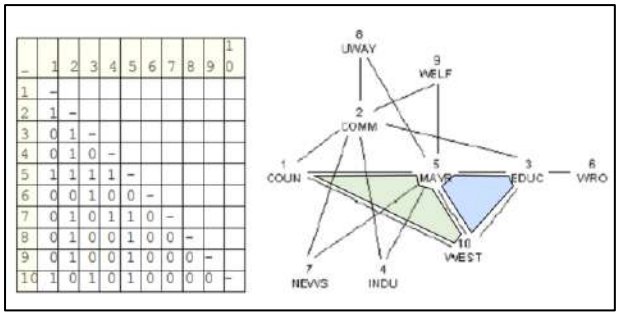


Figure 3. Data matrix and hypergraph from a database of 10 organizations concerned about social well-being in a city in the United States. Here the participating nodes in the blue shaded figure make up a clique of order 2 (3 nodes involved), while it is not possible to get a clique of order 4 (blue shaded figure plus green shaded figure, with 4 nodes involved), since a connection between node 1 and 3 is missing.

Analysis of network motifs

The second method under study mentions the motif. Motifs are the basic components of networks, also known as building blocks. They are defined as connected subgraphs of a given number of nodes. If motifs are studied, the general dynamic behavior is understood, because it provides the fingerprints of local network structures [21-22].

The objective of this method is to develop an efficient algorithm to evaluate the statistical weight of each motif in higher order interactions, in this way, the micro-scale fingerprints of the networks will be kept, enabling the identification of their structural and functional components by focusing on the characteristic patterns of higher-order interactions between small groups of nodes [23].

For this analysis, the heterogeneity of the system is taken into account, so it is interesting to evaluate if the weight of a hyperlink is compatible with a null model in which all the nodes randomly select their partners. In this way, it will be possible to see if they are statistically over-expressed or not with respect to a null model, that is, if within an observed network the motifs appear at a higher or lower frequency in relation to a random model [24]. In complex networks the null model is a random version of an original graph (G) with edges that are randomly wired, its only restriction is that the degree of each vertex coincides with the degree of the original vertex.

In summary, the statistically validated hypergraph is obtained by putting together all the hyperlinks of different sizes that are validated against a null hypothesis.

The steps required to analyze in data-base motifs with higher order interactions are:

1. Count the frequency of each higher-order interaction motif in the network (isomorphism problem).
2. Compare said frequency with that of a null model network.
3. Statistically evaluate its over or under-expression in a hypergraph, according to abundance (Δ) with respect to a null model.

$$\Delta_i = \frac{Nreal_i - (Nreal_i)}{Nreal_i + (Nreal_i) + \epsilon}$$

where ϵ ensures that |Δ| is not misleadingly large when the subgraph appears very few times in both the real and random networks (here ϵ = 4) [31].

4. The above is concatenated/computed into a Significance Profile (SP), which is similar to a fingerprint of the local network structure.

$$SP = \frac{\Delta_i}{\sqrt{\sum {\Delta_i}^2}}$$

One of the problems presented is that it can be difficult to observe higher order interactions in some systems or only data are recorded in pairs, moreover, algorithms for traditionally counting motifs fail to capture information in group interactions, as they cannot consider hyperarist patterns [23]. This newly developed solution uses a Bayesian framework to reconstruct connections in higher order interactions following the principle of parsimony.

When comparing the frequency of each motif with a null model, also called *reshuffling*, in which a network containing the same number of nodes is randomized several times, but with different links, a probability can be calculated to infer whether the cliques are statistically relevant or not, that is, whether they were formed by mere chance. In this way, motif analysis provides much more accurate information than the traditional peer-to-peer method.

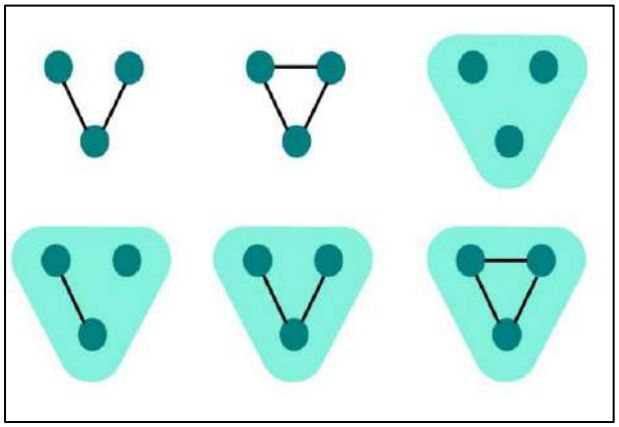


Figure 4. Representation of a 3-node motif with its possible decompositions. Those shaded represent higher order interactions, while those with black lines represent pairwise interactions [23].

The next graph (figure 5) shows another major problem presented when studying higher order interactions. When 3 nodes connected under the pairwise approach are studied in non-directed interactions, there are only two possible patterns (figure 4), however, when non-directed higher-order interactions are considered, the possible combinations increase to 6. This shows that there is an over-exponential growth of the motifs depending on the order of the motifs, that is, depending on the number of nodes that composes a motif. This combinatorial explosion is evident when going from a motif of order 3 (with 6 possible combinations) to a motif of order 4, which reaches up to 171 combinations [0,2]. All these calculations must necessarily be stored in memory, which is why computationally it becomes very demanding, thus normally reaching an analysis that does not exceed order 4 or 5 of the motifs.

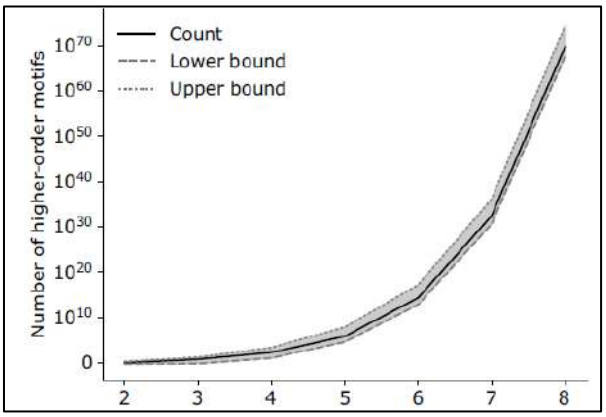


Figure 5. Number of possible decompositions of the motifs according to the order of the motifs. The upper and lower bounds of k nodes should be computed analytically to easily encode the dependence of the number of higher-order motifs according to the order of the motif [23].

Likewise, the motifs are especially helpful in differentiating networks. A network has preferential connectivity patterns at the microscale. This causes a characteristic fingerprint to be encoded, which is relevant according to the functions of the system (eg transport network, whose function is to simplify the flow of traffic; neural networks, which have a function related to information processing) [23]. These patterns are possible to quantify with the network motifs. Thanks to their studies, it is possible to detect early structural changes that, for example, could cause a crisis in the financial network [25], or study direct or indirect interaction networks between species in ecology [26-27].

Differentiation of hypergraphs with different domains according to their SP

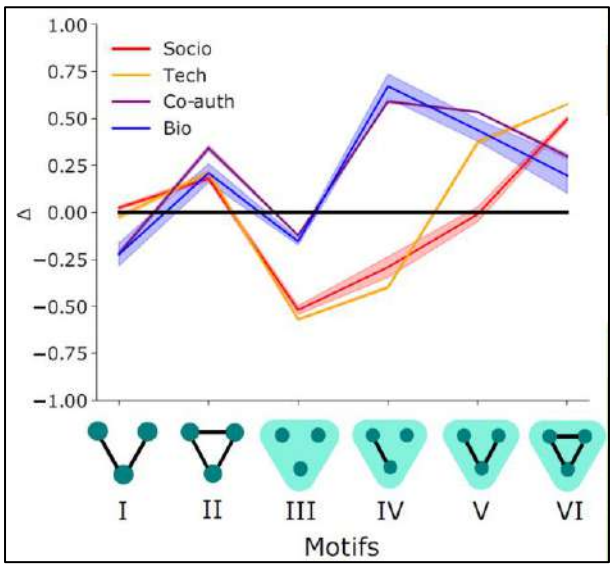


Figure 6. Significance profiles (SP) of different domains for different motif decompositions up to order 3, according to the abundance of each motif in a random network. The average for each motif profile is represented with the solid line, and the standard error of the average is represented with the shaded area [23].

The graph in figure 6 shows the significance profiles (SP) for different higher order motifs of up to 3 nodes in different networks of the same domain, which is calculated by concatenating the over or under-expression of each motif decomposition, leading to the formation of local footprints of the network structure [23]. The result shows that it is possible to identify different families of hypergraphs, characterized by different patterns of higher order connections at a local scale. This analysis in the different domains highlights the relative structural importance of these patterns.

The over-expressed reasons are associated with different functionalities of the system, with which certain behaviors can be generalized for each specific domain, such as:

• Image 2 shows a highly overexpressed motif in all domains.
• In motifs with a hyperedge of 3 and at least one dyadic link, the greatest differences between domains occur.
• In the social and technological domain, the last motif is highly overexpressed, which suggests that entities that interact in groups tend to also interact strongly in pairs.
• In co-authorship networks, the most over-expressed reason corresponds to images 4 and 5, which indicates that there is surely a hierarchical structure, which prevents all nodes from interacting equally between peers, (eg research leader who is co-author of articles with students and postdocs, the latter are not co-authors of articles without the former).
• One can also analyze anti-motifs, that is, under-expressed motifs. Image 3 shows that it is very likely that in the social and technological domains within a group interaction, there are no interactions in pairs either before or after.

Results

For the analysis of both methods in higher-order real-world interactions, freely available network databases were taken. Both databases studied here come from the same sociological domain, that is, from the close contact between the participants. The first database to be analyzed corresponds to the SFHH conference with face-to-face interactions of 405 participants during two days in Nice, France, in 2009. RFID sensors (radio frequency sensors that identified and tracked people located close enough) were used to collect the data. The second database used corresponds to contacts and friendly relations between students at a secondary school in Marseille, France, in December 2013 for 5 consecutive days. The measurement was done with portable sensors. Both data groups correspond to a non-directed network.

Characteristic scale

A program was created to analyze the number of cliques of different order. In the program, the database was grouped into specific time periods. Figure 7 shows the evolution of the number of cliques of order 3 from one-minute intervals to one-hour periods throughout the entire database at the SFHH conference. It can be evidenced a clear grouping that is increasing in number of cliques as the aggregation time passes, with three peaks that remain in time.

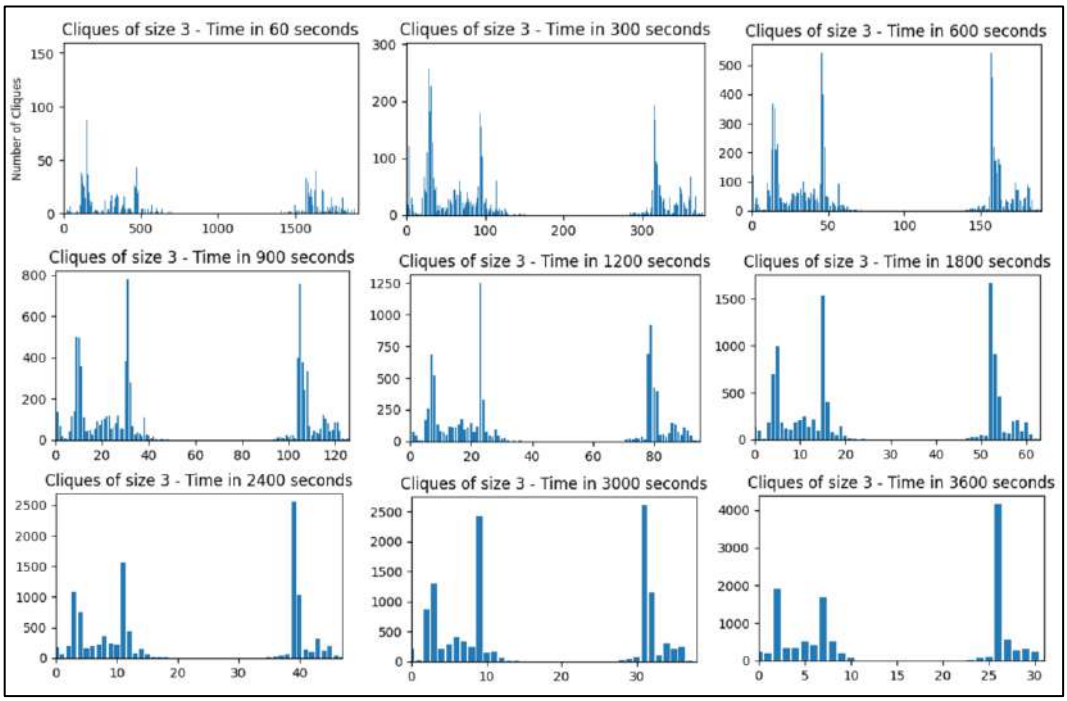


Figure 7. Order 3 clique numbers found in the database at the SFHH conference, for different aggregation times.

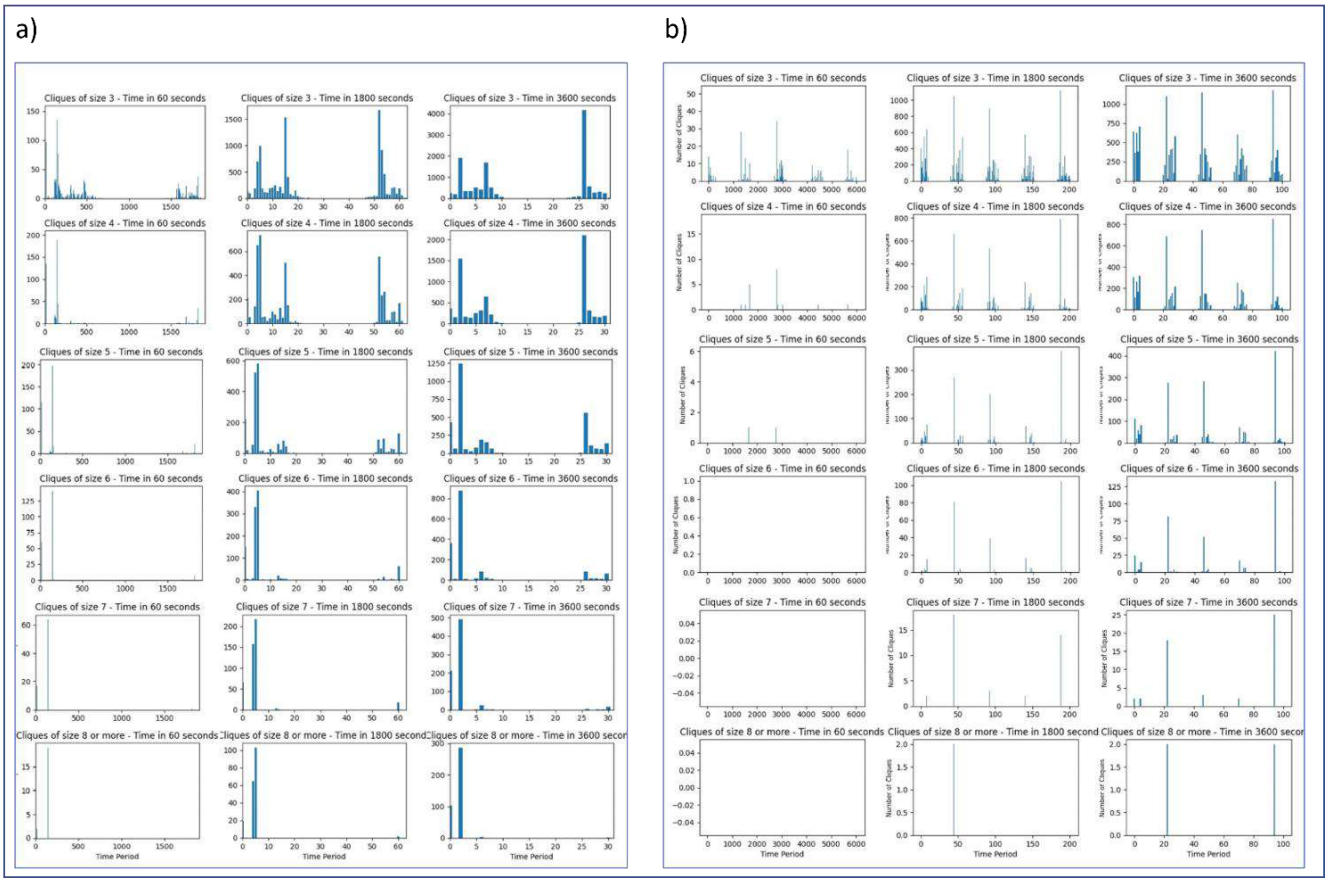


Figure 8. Analysis for different aggregation times (columns) and different clique sizes (rows) for the SFHH conference database (a) and for the secondary school in Marseille (b). In this way, the data are grouped according to the period of time under study, which correspond to 1 minute, 30 minutes and 1 hour.

For the data obtained from a conference (figure 8a), in which the different rows of graphs in the image show the increasing orders of the cliques, and each column shows the increased aggregation times, it is evident that by increasing the interaction, that is, by going down the graphs through the different orders of cliques, it shows that the number of groups with more members is gradually decreasing (higher order cliques), which suggests that there is a temporal evolution in which for an infinite time, will not necessarily occur more close proximity interactions.

This is attributed to the fact that there is a characteristic scale of the size of the groups in each stage under study. The conference shows that the number of cliques above size 5 is no longer significantly in the number of periods, so it could be inferred that it is no longer relevant.

A similar case can be seen with the other database in figure 8b, which corresponds to the interactions obtained in 5 consecutive days. As the analysis of clique's size progresses, there is a certain limit, also close to the order of 5, in which if exceeded, number of interactions ceases to be relevant if correlated with the number of periods existing between row of cliques of order 5 and row of cliques of order 6. The behavior described above remains stable across the different time intervals.

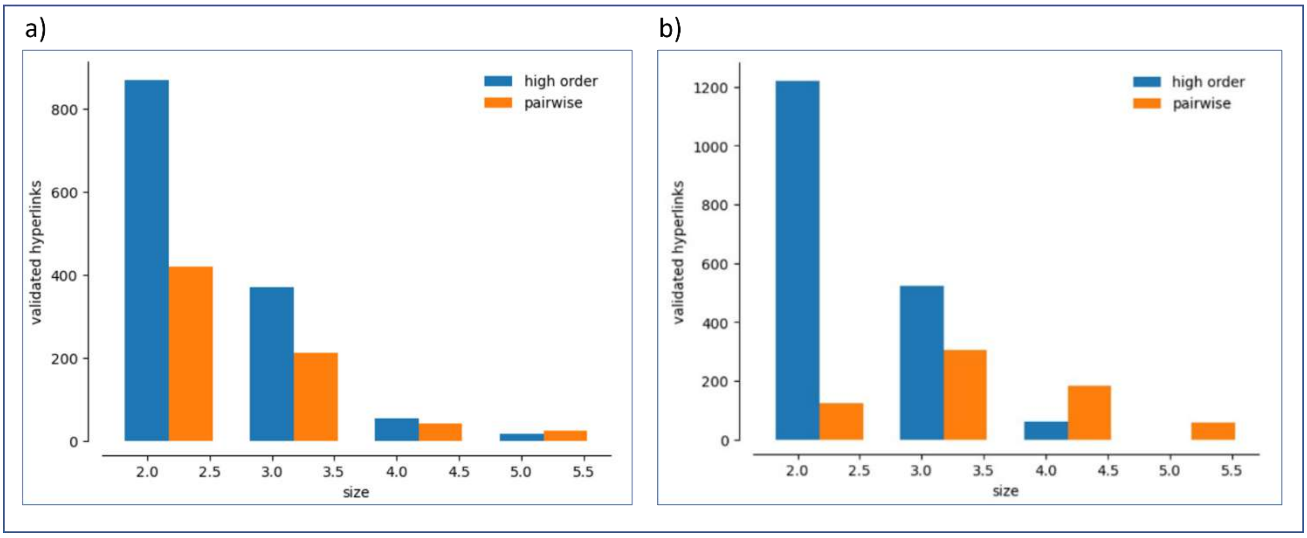


Figure 9. Contrast of the database analysis with the traditional pairwise method and the motif analysis method in higher order interactions.

Figure 9 shows that using a higher order interaction motif analysis approach, more statistically validated hyperlinks can be found than using the pairwise method for lower order cliques. As the order of cliques grows, the pairwise method slightly outnumbers the motifs approach for statistically validated hyperlinks against a random network.

If we only count the number of cliques with the traditional pairwise approach, without doing the subsequent step that consists of the statistical validation process, it can be seen in figure 8 that despite the fact that the longest period of time considered is one hour (whose time is less than that considered in figure 9), it can be seen that the analysis of cliques without statistical validation finds a much larger number of cliques than the one that considers statistical validation, thus accounting for the relevance it has the process also called *reshuffling* to filter out cliques that are not real.

Conclusion

If a network is studied using a traditional pairwise approach, there is a risk that even if a species is known in detail in its individual scope, it is very likely that a failure may arise when trying to understand the population dynamics of said species. The failure is given because this type of approach does not consider the rich pattern of non-linear interactions between the components of the system, a

fundamental ingredient to meaningfully understand a complex system [28].

After years of reductionism, science has left behind the idea that a complex system can be understood by considering only the units as an isolated system [29], the introduction of higher order interactions has begun to shape most of the relational data sets, even if they are not explicitly coded. With the new studies it is possible to recover these interactions from a database, even when that information is delivered following a pair structure.

On the other hand, the analysis of the variation in the size of the cliques as time passes, revealed that although the data can be grouped in larger periods, in both cases studied, people will not necessarily tend to interact in groups of more members. Therefore, a limit is reached, also called characteristic scale, in which cliques of higher order are no longer formed. To find statistically validated cliques against a random network, the motif approach turned out to be much more advantageous than a pairwise approach. Likewise, carrying out the statistical validation process, also called reshuffling, eliminates an important portion of the noise in the networks, that is, those cliques that are not real and therefore do not provide valuable information about the group's dynamic behavior.

Additionally, the study and analysis of the hypergraphs show that there are superfamilies of networks, that is, groups of networks with similar local structure that tend to behave in the same way. This allows us to build models based on more accurate data using network motifs.

With the effort of the scientific community to develop and formalize tools that allow not only the study of higher order interactions, but also, include the temporal variable, giving way to the analysis of non-static interactions [30] has succeeded in making this multidisciplinary field a flourishing discipline, with applications covering a wide range of sciences from fundamental physics to the social sciences [28].

References

1) Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., ... & Petri, G. (2020). Networks beyond pairwise interactions: Structure and dynamics. Physics Reports, 874, 1-92. (2000).
2) Battiston, F. et al. The physics of higher-order interactions in complex systems. Nat. Phys. 17, 1093–1098 (2021).
3) Cencetti, G., Battiston, F., Lepri, B. & Karsai, M. Temporal properties of higher-order interactions in social networks. Sci. Rep. 11, 7028 (2021).
4) Patania, A., Petri, G. & Vaccarino, F. The shape of collaborations. EPJ Data Sci. 6, 18 (2017).
5) Petri, G. et al. Homological scaffolds of brain functional networks. J. R. Soc. Interface 11, 20140873 (2014).
6) Giusti, C., Ghrist, R. & Bassett, D. S. Two's company, three (or more) is a simplex. J. Comput. Neurosci. 41, 1–14 (2016).
7) Carletti, T., Fanelli, D. & Lambiotte, R. Random walks and community detection in hypergraphs. J. Phys.: Complex. 2, 015011 (2021).
8) Chodrow, P. S., Veldt, N. & Benson, A. R. Generative hypergraph clustering: from blockmodels to modularity. Sci. Adv. 7, eabh1303 (2021).
9) Berge, C. Graphs and Hypergraphs (North-Holland Pub. Co., 1973).
10) Bick, C., Ashwin, P. & Rodrigues, A. Chaos in generically coupled phase oscillator networks with nonpairwise interactions. Chaos 26, 094814 (2016).
11) Gambuzza, L. et al. Stability of synchronization in simplicial complexes. Nat. Commun. 12, 1–13 (2021).
12) Battiston, F. et al. Networks beyond pairwise interactions: structure and dynamics. Phys. Rep. 874, 1 (2020).
13) Latapy, M., Magnien, C. & Del Vecchio, N. Basic notions for the analysis of large two-mode networks. Soc. Netw. 30, 31–48 (2008).
14) White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of thenervous

system of the nematode Caenorhabditis elegans. Philos. Trans. R. Soc. Ser. B 314, 1–340 (1986).

15) Grilli, J., Barabás, G., Michalska-Smith, M. J. & Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. Nature 548, 210–213 (2017).
16) Atkin, R. Mathematical Structure in Human Affairs (Heinemann, 1974).
17) Resnick, M. D. et al. Protecting adolescents from harm: findings from the national longitudinal study on adolescent health. J. Am. Med. Assoc. 278, 823–832 (1997).
18) Young, J. G., Petri, G., & Peixoto, T. P. Hypergraph reconstruction from network data. Communications Physics, 4(1), 135. (2021).
19) Patania, A., Vaccarino, F. & Petri, G. Topological analysis of data. EPJ Data Sci. 6, 7 (2017).
20) Petri, G., Scolamiero, M., Donato, I. & Vaccarino, F. Topological strata of weighted complex networks. PLoS ONE 8, e66506 (2013).
21) Hong-lin, X., Han-bing, Y., Cui-fang, G. & Ping, Z., Social network analysis based on network motifs. J. Appl. Math. 2014, 874708 (2014).
22) Juszczyszyn, K., Kazienko, P. & Musiał, K. Local topology of social network based on motif analysis. In International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, 97–105 (Springer, 2008).
23) Lotito, Q. F., Musciotto, F., Montresor, A., & Battiston, F. Higher-order motif analysis in hypergraphs. Communications Physics, 5(1), 79 (2022).
24) Milo, R. et al. Network motifs: simple building blocks of complex networks. Science 298, 824–827 (2002).
25) Saracco, F., Di Clemente, R., Gabrielli, A. & Squartini, T. Detecting early signs of the 2007–2008 crisis in the world trade. Sci. Rep. 6, 1–11 (2016).
26) Bascompte, J. & Stouffer, D. B. The assembly and disassembly of ecological networks. Philos. Trans. R. Soc. B: Biol. Sci. 364, 1781–1787 (2009).
27) Simmons, B. I. et al. Motifs in bipartite ecological networks: uncovering indirect interactions. Oikos 128, 154–170 (2019).
28) Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., ... & Petri, G. Networks beyond pairwise interactions: Structure and dynamics. Physics Reports, 874, 1-92 (2020).
29) P.W. Anderson, More is different, Science 177 (4047). 393–396 (1972).
30) P. Holme, J. Saramäki, Temporal networks, Phys. Rep. 519 (3) (2012) 97–125.
31) Milo, R., Itzkovitz, S., Kashtan, N., Levitt, R., Shen-Orr, S., Ayzenshtat, I., ... & Alon, U. Superfamilies of evolved and designed networks. Science, 303(5663), 1538-1542 (2004).