POLITECNICO DI TORINO

Corso di Laurea Magistrale
in Ingegneria Matematica

Tesi di Laurea Magistrale

Dynamics of popularity in social media:
The Roles of Network Effects and Recommendation Systems

Relatori
prof. Chiara Ravazzi
prof. Paolo Frasca
prof. Fabrizio Dabbene
firma dei relatori

. . . . . .
. . . . . .
. . . . . .

Candidato
Gaya Cocca

firma del candidato

. . . . . .

Anno Accademico 2024-2025

Summary

This thesis aims to present a new model for studying the evolution of popularity on social networks, taking an innovative route compared to what is currently in the literature.

Content appreciation is described through a stochastic nonlinear model, from which insights can be derived to study the popularity of influencers. The mathematical formulation emphasizes three terms: user interactions, the role of the content recommendation platform, and the intrinsic characteristics of the influencer himself. These are called external factors because they are external to the dynamics of the network.

The goal is to understand what kind of characteristics this model has, its steady-state behavior, and whether it admits equilibrium points. The process begins with the study of a deterministic model, where all the special cases are analyzed, obtained by canceling out some contributions rather than others. Each case shows clear equilibrium points dependent in most cases on the value of the external factors associated with the influencers. This deterministic model approximates the expected value of the stochastic model so it gives a clearer view of the latter. Numerical simulations are implemented for each case to prove the results.

The stochastic model is analyzed by repeating the same approach as the deterministic one, that is, by canceling one or more contributions and studying all special cases. The definition of a Markov chain and the above approximation help to characterize this model as well, understanding its behavior over the long run. Once again, convergence conditions are obtained and numerical simulations are exploited to prove the results.

In parallel with the theoretical efforts, a dataset containing information on a set of YouTube channels is explored. From this data, several measures of popularity are derived and potential external factors associated with the channels are explored. This approach not only provides insight into real-world dynamics but also lays the foundation for future analyses aimed at validating the theoretical model with larger and more representative datasets.

This work was carried out partly at the CNR (Centro Nazionale delle Ricerche) Department of Electronics and Telecommunications, within the IEIIT institute in Turin — a National Research Council of Italy reference center for advanced scientific and technological research in information engineering. At the CNR, I had the opportunity to interact with researchers in a collaborative environment, benefiting from discussions and insights that enriched the development of my research. The second part of the work was conducted at GIPSA-lab in Grenoble, a joint research laboratory of the CNRS (Centre National de la Recherche Scientifique), Grenoble-INP, and Grenoble Alpes University. At GIPSA-lab, I pursued my research in collaboration with the CNRS, within a multidisciplinary setting where theoretical and applied studies are developed in the fields of control, signal processing, and artificial intelligence.

Contents

List of Tables

List of Figures

It may seem difficult at first, but all things are difficult at first.
[Miyamoto Musashi, The Book of Five Rings]

Abstract

The goal of this thesis is to analyze how social networks, social imitation, combined with recommendation systems, and content quality can play a crucial role in popularity dynamics in social media. In particular, we formulate a new model of popularity dynamics in social networks and study its mathematical properties. The proposed model is nonlinear and includes randomness components to capture the unpredictable fluctuations and stochastic dynamics typical of social media. Using data collected on YouTube videos over the period 2020-2023, we explore real-world trends in user engagement and video virality to gain insights into the mechanisms driving popularity dynamics.

Chapter 1

Introduction

We live in a digital world where we spend hours every day scrolling through smartphones and enjoying content of all kinds. Actively participating in the reality of social media such as YouTube, TikTok, Instagram, and many others, means constantly interfacing with the popularity of *creators*. Creators (or *influencers*) are individuals who produce content that shapes the interests, preferences, and behaviors of their audience. Through their work, they introduce lifestyles, products, and ideas, encouraging users to engage with and adopt them.

Although we are not directly affected and aware of it, such social networks day by day are characterized by upward or downward trends in the popularity of influencers, which lead to the prevalence of the contents of one, rather than others. An increase in their popularity has huge benefits for the creators themselves, who automatically become more visible on the platform, and have greater monetary compensation and offers of partnerships with top companies. Understanding the dynamics of popularity, in fact, could have large implications for marketing strategies, platform policies, and even the spread of fake news. Companies invest heavily in influencer partnerships, while social networks fine-tune recommendation algorithms to balance engagement and fairness. A rigorous model of these dynamics can provide insights for businesses and policymakers.

Given the vastness of the topic, several techniques have been implemented over time. Most existing research relies on epidemiological and diffusion models to describe the evolution of popularity, adapting frameworks originally developed for different contexts, such as disease spread. While they are concerned with monitoring popularity from an information perspective, that is, analyzing how a certain piece of information spreads in a population, they struggle to capture the interplay between algorithmic recommendations, peer influence, and content quality, which jointly determine an influencer's success. Our work fills this gap by explicitly incorporating these factors into a stochastic framework.

Our contribution is both empirical and theoretical: we explore and analyze real-world data to extract meaningful information, and we formulate and study a theoretical mathematical model. Although at this time the analysis of the dataset is purely exploratory, it is interesting to understand what kind of information can be gleaned from it, with a view to subsequent work that might deal with validating the model on larger and more representative databases.

Building on this, our empirical contribution primarily focuses on exploring a dataset associated with YouTube, deriving measures of user engagement (and consequently popularity), and analyzing their trends. The observation of the evolution and a hypothetical convergence, allow us to get a first glimpse of what is happening on the platform and to understand whether there are outcomes that we could predict and reproduce in some way, or not. At the same time, because influencer characteristics are central to our dynamic, it turns out to be interesting to understand in the context of YouTube what channels' features might lead a generic user to view more of its videos rather than others. Again, the dataset is exploited to obtain information in this sense.

From the theoretical point of view, we propose a nonlinear stochastic model to formalize user appreciation and derive insights into the popularity of influencers. Based on the theory of opinion dynamics, it represents an innovative formulation compared to the current literature.

The model emphasizes three key terms: users' interactions, the recommendation platform of the social network itself, and aggregated intrinsic features of the influencer (external to the network). Its structure, obtained by dividing and weighing each contribution, has two strengths: it is a holistic approach, which allows us to shape different terms and assess how their presence or absence results in substantial changes in the dynamics, and it is a formulation that is simple to understand, without excessive modeling and abstractions.

The three key terms governing the dynamics appear natural to figure out. Since we want to model the network's appreciation of content to gain insights into the popularity of influencers, we must first understand how users interact with each other and exchange information. We then need to recognize how impactful are recommendation platforms, which by suggesting one piece of content rather than another, contributes to increasing the popularity of certain creators while penalizing others. Anything left out of these contributions is included under the umbrella of *intrinsic influencer characteristics* (or *external factors*), which could be the expertise, the quality of content, or the creator's experience.

Our theoretical contribution is therefore aimed at understanding the characteristics and behavior of the model in a rigorous way, studying all its special cases, obtained by canceling one or more terms. Under each condition we analyze the steady-state behavior, getting assumptions and theorems that guarantee convergence. In addition, where possible, we obtain and formalize the equilibrium points of the dynamics. Each theoretical analysis and conclusion is supported by simulations in Matlab that show any peculiarities found numerically.

It is worth noting that we start by analyzing a simplified deterministic model, which has the same basic setting. It is an approximation of the expected value of the stochastic model and serves as a starting point for us to understand and simplify the treatment of the stochastic part. In this way, we can get a more complete picture of all possible cases.

This work lays the groundwork for future research, providing a baseline model that can be expanded with additional factors and complexities. Future studies may refine its assumptions, incorporate more detailed user behavior, or explore alternative modeling approaches.

To make the various concepts used understandable and to frame our work in the current research, Chapter 2 will be aimed at introducing the state-of-the-art of popularity

evolution models and recommendation systems on YouTube, as well as recalling the basics of opinion models. In this way we can place our work in a broader framework, reflecting on what mathematical techniques are being exploited to date to address problems of this type. Chapter 3, on the other hand, will focus on the empirical part of the thesis, describing the dataset, the data cleaning and pre-processing, and the various elaborations performed on it. It is relevant to note that the empirical analysis supports both the interest in a stochastic model and a deterministic approximation of it, as not only measures of user engagement will be derived, but their time-averaged evolution will also be analyzed. At the same time, there will be a focus on the topic of characteristics that determine success, consistent with what will be analyzed in theory, in which quality/competence factors will be discussed. Chapters 4 and 5 will be aimed at describing the entire theoretical formulation, starting with the deterministic model and then presenting the characteristics of the general stochastic model. Each case, associated with different parameters, will be studied, obtaining rigorous results in turn supported by numerical simulations. For all scenarios, conditions of convergence will be derived. Finally, Chapter 6 will show several noteworthy special cases associated with interesting behavior in simulation. It will be further understood how important the term modeling the influencer's quality/competence is. The Appendix 8 will be used to recall theorems and definitions important for understanding all the formulations and demonstrations. An alternative version of a proof that we will be running in the following chapters will also be included in this section.

Chapter 2

State of the Art

In this chapter, we will analyze the state of the art of popularity evolution models and recommendation platforms, as well as recall several concepts useful for understanding our work in its entirety. First, we will do a brief literature review of what are the most widely used approaches to modeling popularity on social networks nowadays.We will then introduce the context of opinion dynamics on which our model relies. Finally, in the last two sections of this chapter, we will make an introduction to the world of YouTube and recall the state of the art of its recommendation platforms, which represent an important contribution to formalizing our model.

2.1 Modeling Popularity: A Literature Review

Nowadays, since it is a central research theme, numerous models of different types have been proposed to explain how popularity evolves and what kind of characteristics it has. They can be divided into four main categories, related to a different nature of the model: *epidemic models*, *self-exciting processes*, *Bass-like diffusion models*, and *Neural Network-based models*.

As their name suggests, epidemic models arise to study the evolution of diseases within a population. They were later applied to a variety of contexts, including the evolution of popularity. One of the most famous epidemic models is the SIR model [3], in which the population is divided into three categories: Susceptible, Infected, and Recovered individuals. Once parameters such as death rate and infection rate are defined, the evolution of the disease is formulated through ordinary differential equations, where each explains how one of the subsections of the population varies. Such methodologies can be easily translated to social dynamics and in particular to the transmission of content (hence the popularity), replacing the concept of disease with that of information (knowledge of content): from the latter, the infected group represents the informed population while the removed one represents those who forget the information. Daley and Kendall in [4] were the first to apply the epidemiological model to this new context, trying to analyze the spread of rumors. They were later followed by many others, including Bettencourt et al. [5] who proposed a variation of the SI model (susceptible and infected only) to understand

the widespread use of Feynman diagrams across the U.S. physics-theoretical communities. To do so, they included two new classes: Exposed, meaning those exposed to the new idea, and Skeptics, meaning those aware of the idea but skeptical of applying it. Subsequently, Jin et al. [6] applied the same model to online information dissemination, showing an interesting applicability, further proven by Xiong et al. [7] who showed how including a delay term between receiving the idea and adopting it could refine the dynamics even more. Richier et al. in [8], on the other hand, used a differently constructed epidemic model and managed to fit the 90% popularity of a dataset containing YouTube videos. The use of this type of models has also proved successful in defining the evolution of YouTube video virality, as demonstrated by [9].

The second category of models applied to the evolution of popularity is self-exciting processes, of which the work of Crane and Sornette [10] is a landmark. The model aims to study the evolution of videos on YouTube by giving relevance to three aspects: user interactions in the exchange of views, the existence of influences external to social media, and non-homogeneous patterns over time. The model is defined by conditional Poisson processes with a view rate of

$$\lambda(t) = \lambda_0(t) + \sum_{i,t_i<t} \mu_i \phi(t - t_i),$$

where $\lambda_0(t)$ is the exogenous source of views while the summation models the imitation process; μ_i is the number of viewers influenced by i and $\phi(t)$ represents the rate at which individuals consume information.

Turning now to analyze the third group of models, Bass diffusion models were first introduced in [11] to monitor the dynamics of the introduction of new products in the market. This is a very simple formulation that divides the population into only two categories: innovators and imitators. Over time this approach has been complicated to include macroscopic agent characteristics, such as user habits and content characteristics [12], or user features [13]. Other studies have used Bass's model by varying diffusion networks, starting with variations in network degrees [14], and arriving at the assumption of the existence of different communities to simulate various populations on the Web [15].

The most recent category is represented by models that exploit neural networks to predict the evolution of popularity. Neural networks are well-suited for capturing complex, nonlinear relationships in social media data. Many different approaches have been implemented in recent years to simulate various effects. In 2019, Cao et al. [16] proposed a model that uses Graph Neural Networks (GNN), to predict online content popularity while considering network effects—particularly the cascading effect of information spread. The cascading effect is related to the interplay between user activation (engagement) and content diffusion. The model consists of two interdependent GNNs, each focusing on a different aspect of content virality: the interpersonal influence of users in the network, and the user engagement with the content. The results validated that the model outperforms traditional methods in predicting content popularity. The same tool was used by Jin et al. [17] to model information diffusion and temporal dynamics within social media networks. Their primary goal was to predict emerging trends by learning effective entity representations from social data. Finally, unlike prior works that considered only social influence (the power of early adopters to spread content) or homophily (the tendency of similar

users to engage with the same content), Shang et al. in 2022 [18], proposed a data-driven framework to combine both. Also, in this case, they used two Graph Neural Networks to simulate social homophily and social influence. Additionally, instead of treating early adopters as independent, they grouped them into social clusters based on user interest. After testing on two real-world datasets, the results showed that their model significantly outperforms existing approaches, confirming that combining social influence, homophily, and social groups improves accuracy.

2.2 Opinion Dynamics: Context and Foundations

The use of models typical of the opinion dynamics branch represents an innovative step in the study of the evolution of popularity, as there is not strong literature and research in place. To clarify the context and tools that will be used in the future, we show in this section the main concepts associated with these models.

Opinion dynamics is a field of study in mathematics that aims to understand how individuals update their beliefs, attitudes, or preferences through interactions with others. It is particularly relevant in understanding how consensus, polarization, or fragmentation emerge in societies and online communities. Central to these models are agents, who interact with each other and exchange views. They are the entities of social networks, which is modeled in various ways depending on how the interactions are represented. For this reason, the main components of an opinion dynamics model are:

- Agents: individuals holding an opinion;
- Opinions: they represent the point of view (or the ideas) of an agent;
- Interaction Rule: it defines how opinions change when there are interactions between agents;
- Network Structure: it describes how agents interact and influence each other. It is represented by a graph.

As these elements vary, different patterns are obviously found, among them:

- *Consensus Models*: These models explore how opinions converge over time. A fundamental example is the *DeGroot Model* [19], where agents take a weighted average of their neighbors' opinions. A key extension is the Friedkin-Johnsen Model [20], which introduces stubbornness by allowing agents to retain some weight on their initial opinions rather than fully conforming to their neighbors.
- *Bounded Confidence Models*: In models like the Hegselmann-Krause [21] and Deffuant-Weisbuch [22] models, agents only interact with others whose opinions are sufficiently close to their own, leading to opinion clusters.
- *Discrete Opinion Models*: The Voter Model [23] and the Sznajd Model [24] describe opinion evolution in binary or categorical settings, often used to study political shifts or brand competition.

In analyzing the special cases of our theoretical model, we will find in some cases known results and models already studied: in that case we will refer to the literature associated with that particular model. Instead, our contribution is related to the theoretical understanding and study of all the part not known in the literature, which we will formalize from scratch.

2.3 Understanding YouTube as a platform

YouTube was founded in 2005 by Chad Hurley, Jawed Karim and Steve Chen. Although it was initially conceived as a dating platform where users could introduce themselves via video, the developers quickly decided to make it a space for publishing videos of all kinds.

The publication of videos is done through channels, which represent the entities that populate the social network. Each user can react to each channel's videos with likes and comments. He can also observe how many times a certain video has been viewed, as YouTube associates each piece of content with the number of views since it was uploaded.

The platform has grown enormously in popularity over time. Being bought by Google a year after its launch, it became the second most visited website in the world in 2020. Nowadays, the number of active users is 2.7 billion monthly [25]; in other words, 1 in 4 people in the world accesses YouTube at least once a month. This implies a strong influence of the social network in shaping users' opinions.

Given its powerful influence, the platform has been at the center of numerous controversies over time, from fighting fake news circulating on it, which led to the phenomenon known as "The Purge", resulted in many videos being deleted and demonetized, to copyright infringement. At the same time, however, it expanded increasingly, attracting more and more users who wished to earn income from uploading a wide variety of videos. YouTube, in fact, ensures that videos are monetized, mainly through the presence of commercial product sponsorships or paid subscriptions that some users implement for the channels they like most. Therefore, a popular channel that manages to upload videos receiving a large numbers of views, has substantial revenue in economic terms.

Although content of all kinds is present, we will focus only on channels associated with the dissemination of news, analyzing which of these manage to attract the most audience, and thus be most popular, and trying to find reasons for this phenomenon. In fact, it is well known that in the context of breaking news, there is a strong prevalence of the platform: a recent study [26] reveals that 26% of Americans get their news on YouTube, and 72% of news consumers consider it an important way to get news.

2.4 Recommendation Systems on YouTube

In analyzing the state of the art of recommendation systems, we will focus exclusively on YouTube, as it is our main topic of study, given the data we have available.

As evidenced by the fact that it is the second most widely used social network after Facebook, YouTube is characterized by a strong dynamic of publications: 3.7 million videos uploaded every day, 2.500 new videos per minute, amounting to about 500 hours of video [27], represent a huge set of content that every user is faced with. It is therefore

natural that such a platform, on par with the major social platforms around today, should have a recommendation system that can present users with videos that are more similar to their tastes and that might capture their attention more. It is clear, therefore, that the choice of videos to play is not only associated with the user's search (mainly based on several factors that we will discuss later) but also to this inherent operation of the platform itself.

The main recommendation systems today are based on artificial intelligence techniques. Information about them is not numerous, as they are proprietary; all that is known is mainly based on papers published by the developers themselves, in which the main structure is explained in general. To this purpose, an interesting analysis was carried out by Maria Castaldo in her doctoral thesis [28], from which some insights are taken. To date, *Deep Neural Networks for YouTube Recommendations* [1], although it was published in 2016, is still an important reference point for understanding the logic of the recommendation mechanism. Subsequent papers [29] and [30], confirmed the approach on neural networks presenting changes in terms of layers and activation functions, for which [1] will be used as a reference, being very clear about how the recommendation process takes place.

An effective recommendation platform must first and foremost take into account three aspects typical of data on social networks: *scalability*, given the number of new videos uploaded every day; *novelty* so that the showcase of proposed products improves each time; and *noise* since predicting the generic user's viewing history entirely is very complex due to external factors. The model described by Covington et al. in [1], developers at Google, is based on deep learning techniques. It is divided into two basic parts, as shown in the figure 2.1, a *candidate generation* part and a *ranking* part. The first phase, taking the user's YouTube activity history as input, returns a subset of videos (about 100), making a first big screening. In contrast, the ranking phase associates a score with each video, using a large set of features for both the user and the video itself. The user will end up with the videos in the order they were assigned by ranking.

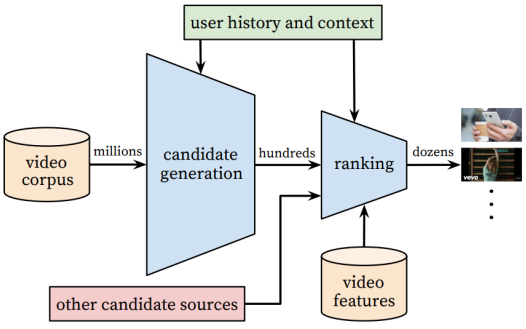


Figure 2.1: Recommendation system architecture from [1]

Recommendation is defined as a multiclass classification where one must classify a video to watch at time t, i.e. w_t, among millions of videos i that correspond to classes, defined in a set V based on user U and context C. N corresponds to the number of videos

that we want to obtain from the screening. The probability to watch the specific video i is defined by

$$P(w_t = i|U, C) = \frac{e^{v_i u}}{\sum_{j \in V} e^{v_j u}}$$

where $u \in \mathbb{R}^N$ represents the *embedding* of the user-context pair, while $v_j \in \mathbb{R}^N$ represents embeddings of each candidate video. An *embedding* is a map of sparse entities in a dense vector. The main idea is to learn user embeddings as a function of the user's history and context, which are useful for discriminating among videos. The generic architecture of the network is shown in the figure 2.2, taken from [1].

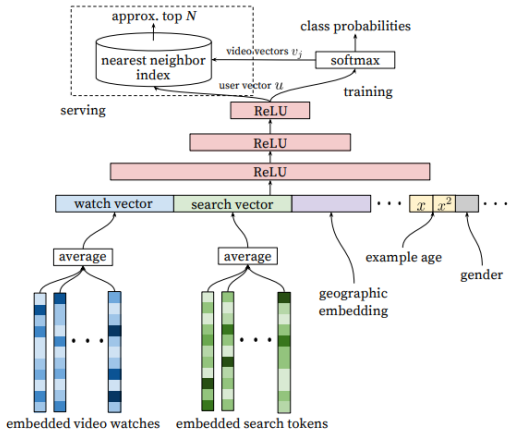


Figure 2.2: Deep candidate generation model from [1]

The various embeddings related to the videos watched and other more or less personal characteristics of the user are averaged and then fed to the neural network. Once the embeddings u and v_j $\forall j \in V$ are obtained, a typical Machine Learning technique, *Nearest Neighbor*, is applied to generate the suggested N videos. During the ranking phase, it is possible to access more features describing the video and the user's relationship to the video, as the number of items is no longer too large. A neural network is used also in this case, with a structure similar to the previous phase. The specific score for each video is obtained using a logistic regression. Figure 2.3 shows that more embeddings are used as input in the neural network for ranking.

Beyond the mechanism inherent in the recommendation platform, we most want to fix the factors taken as input by the neural networks. This will allow us to reason about how to account for and model the recommendation in our theoretical model. The authors of the paper [1] insist on three basic concepts: *user similarity* (homophily), *previous popularity*, and *freshness*. The first is met with the definition of embeddings mentioned earlier, in which users and content are mapped together in N-dimensional space (they become points in this space). The distance in that space represents the similarity and affinity between users or videos. The second is previous popularity, which is related to the propagation of viral content, known in the literature as the dynamics of the *rich getting richer*. Finally,

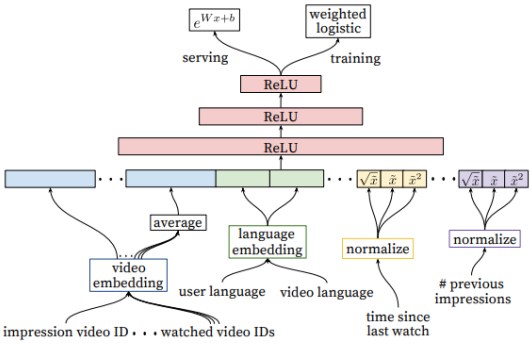


Figure 2.3: Deep ranking network architecture from [1]

the third aspect involved in content selection is content freshness, modeled by inserting an age to the elements of the training set in the training phase, since users prefer newer videos.

These concepts are further confirmed by later articles, primarily [31] where Massimo and Ricci once again point out how the themes of novelty and prior popularity are central in recommendations. Popularity has become a key focus of the study of such mechanisms, as confirmed by [32] and also evidenced by the numerous articles (of which [33] is an example) that question the extent to which such recommendation platforms amplify biased popularity in input. In our case, as will be illustrated more precisely in later chapters, we will take into account only popularity to model the impact of recommendation, given its great importance in the literature and its key role as an input to neural networks.

Chapter 3

Data Collection and Empirical Analysis

This chapter aims to define the applied part of our study. Our purpose is to analyze the evolution of popularity of different YouTube channels using empirical data. To analyze channels' popularity, we will need to:

- define a metric of *user engagement* with which to associate the concept of popularity,
- understand what *external factors* influence popularity, trying to answer why some channels appear more "competent"/"interesting" than others.

A careful analysis of many user engagement metrics and possible external indicators that could be defined will clarify what are the major phenomena that we can reproduce with the theoretical model.

3.1 YouTube Data Collection

The dataset exploited for empirical analysis in this study includes detailed information about the videos of a subclass of YouTube channels. This section outlines the key characteristics of the dataset as well as motivating all the choices and directions we made over time.

The dataset was collected in 2023 by Yahya Bajdadi at the Gipsa Lab in Grenoble, under the supervision of Prof. Paolo Frasca, [34]. It is composed of 316908 videos, each characterized by: channel of publication, identifying name, video title, date and hour of publication, number of views, number of likes, number of comments, and duration. The range of publication channels consists of 47 major news channels in France, which can be divided into 5 categories due to their different nature:

- newspapers,
- radio,
- TV,

- YouTubers,
- Web.

The videos span a time range from 2006 to 2023.

channelName	video_id	title	publishedAt	viewCount	likeCount	mmentCou	duration
01net	1FQC73IObsI	Ne faites jamais confiance à l'IA !	2023-08-19 12:00:01	1319	84	2	0,983333
01net	UnNTalcff4A	Android peut maintenant détecter les AirTag indésirables !	2023-08-18 11:06:56	4173	154	4	0,883333
01net	bNJijypzyWY	Comment pourrait être l'Apple Watch Series X ?!	2023-08-17 11:55:21	2056	111	1	0,733333
01net	_s9P3sGMmlE	Le prix de Disney+ augmente et la publicité débarque	2023-08-16 16:14:54	1045	45	3	0,916667
01net	MPoJfNcdxIg	Xiaomi se prépare à défier Samsung !	2023-08-16 13:09:33	2919	141	3	0,783333

Figure 3.1: Example of rows in the dataset

The collection was mainly done using the YouTube API, with support from third-party sites.

This dataset was then further analyzed by Titouan Vial in June 2024 [2], under the supervision of Prof. Paolo Frasca, in order to identify interesting or common publication practices. Several statistical tests were performed to find possible correlations between the variables, publication trends, and variation in duration over time. Titouan's study served as a starting point for a better understanding of the dataset and to obtain information more quickly, particularly on the distribution of videos over time. The latter test had shown, in fact, how the largest publications were concentrated in the years 2019 to 2023, Figure 3.2.

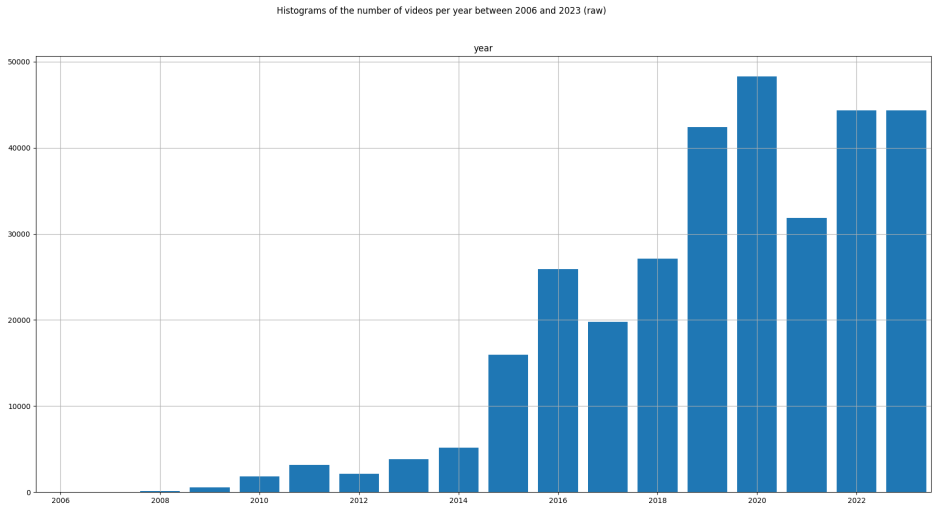


Figure 3.2: Distribution of the number of videos published in the years 2006-2023 in [2]

Taking into account, moreover, that the arrival of COVID-19 had greatly impacted the dynamics of publication, showing a more fluctuating trend due to the forced inactivity of many channels, it was decided to focus our study on the period of June 2020 - July 2023, thus about 3 years.

Closer analysis showed that many channels had sporadic publishing practices, alternating between short moments of activity and long moments of inactivity. This is likely associated with the fact that for many of them YouTube is not the primary showcase, but rather a complement to be used only in certain situations. For this reason, the dataset was cleaned in terms of channels considered. For a more accurate analysis in these terms, we calculated how many months of total inactivity there were for each channel, if any. After several trials, we decided to eliminate from our study those channels with a number of full months of inactivity greater than 8. This reduced our list to a total of 21 channels (Figure 3.3).

- 20 Minutes France
- AJ+ français
- BFMTV
- Brut
- France Culture
- France Inter
- l'express
- La Provence
- Le Figaro
- Le Monde
- Le Parisien
- Le Point
- LeHuffPost
- Les Echos
- Mediapart
- Mouv'
- Ouest-France
- RFI
- Slate.fr
- Sud Ouest
- Valeurs Actuelles

Figure 3.3: Channels of the final dataset

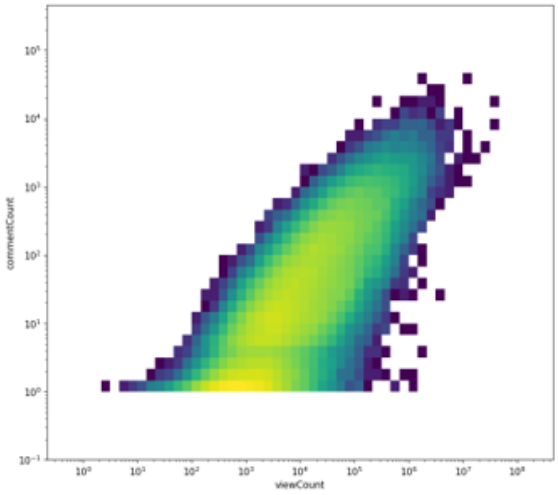
As anticipated earlier, all videos posted by these channels, are characterized in terms of duration, number of views, comments and likes. It is interesting to note for the purpose of subsequent analyses that there is correlation between number of views and number of likes (Figure 3.4a), and between number of views and number of comments (Figure 3.4b), as already confirmed by many earlier studies ([35],[36]). While the former is more reminiscent of a logarithmic relationship between the two variables comments and views, the latter shows that the number of likes can be regarded as a power of the number of views.

3.2 User Engagement Metrics

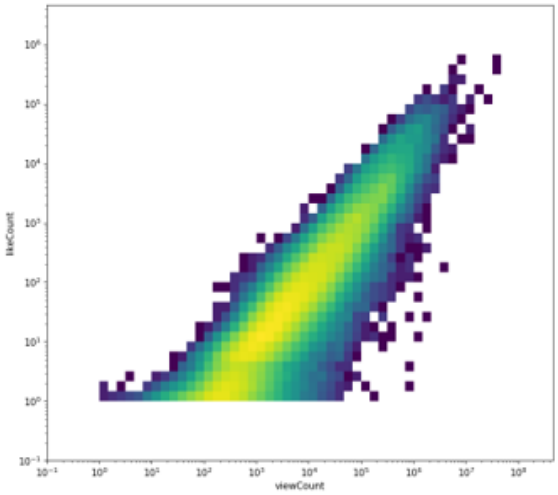
In the context of social media, it is very significant to define *user engagement metrics*, which can give an idea of how much a user is attracted to the platform.

User engagement is defined as how users interact with the product, how often they use it, and how long they stay on it. This is a very important indicator because it provides insight into how much a certain piece of content is effective on users, or in other words, how much that content is popular.

Of course, each platform allows its creators to monitor their user appreciation with internal and private statistics. From outside, however, it becomes a bit more complex to understand how users are attracted to one piece of content rather than another, as there are little data available.



(a) Correlation between num. of views and num. of comments



(b) Correlation between num. of views and num. of likes

Figure 3.4: Correlations in [2]

In the specific context of YouTube, any user from outside can access to several indicators of appreciation of a video, which are the number of views, the number of likes and dislikes, and the number of comments. As noted in the previous section, our dataset makes us aware of all these indicators except the number of dislikes. Obviously, a more complete investigation would require additional elements, such as the average viewing time of a certain video, the number of shares of the video itself, and the rate of clicks on videos linked to the same channel. Since this is not just a problem we encounter, but is related to the very structure of YouTube and what the platform wants and does not want

to show, many studies have addressed it by proposing several methodologies for modeling user engagement. In many cases [35], [36], [37], the metrics used to model popularity through user engagement referred simply to the use of classic indicators separately, so number of views, likes, and dislikes.

Although it is somewhat dated work now, Liikkanen et al. in [38] used YouTube as a case study to discuss a range of new user engagement measures. However, the metrics used are combinations of the same data, hence, the number of views, likes, dislikes, and comments. They discussed three metrics: *Comments per thousand impressions* (CpkI), *Votes per thousand impressions* (VpkI), and *Dislike proportion* (DisP), demonstrating through statistical tests a good prediction of user engagement of various YouTube channels. Since the dataset at our disposal does not include the number of dislikes, the only one of these metrics on which we will focus our attention is the first one, defined as

$$CpkI^{(i)}(t) = \frac{N_c^{(i)}(t) \cdot 1000}{N_v^{(i)}(t)},$$

where $N_c^{(i)}(t)$ is the number of comments and $N_v^{(i)}(t)$ is the number of views of the channel i at time t.

We therefore decided to proceed by working with a dual measure of engagement with which to associate the concept of popularity: on the one hand simply using the number of views of the videos of the channel i ($N_v^{(i)}(t)$), and on the other hand using the measure just presented ($CpkI^{(i)}(t)$). In this way we tried to understand what kind of trends the channels have over time in terms of different measures of popularity. Obviously, because they are different, each puts the focus on different aspects, so we expect specific trends.

Before showing the results obtained, it is important to point out that in both cases the user engagement is normalized, so as to model the competition between the channels. It is in our interest to understand which channel turns out to be more popular, which less so, and how this popularity varies over time, so normalizing the user engagement values makes it easier to answer this question. For example, if $N_v^{(i)}(t)$ represents the number of visualizations for the channel i at time t, the normalized measure of number of visualizations will be

$$\pi_1^{(i)}(t) = \frac{N_v^{(i)}(t)}{\sum_j N_v^{(j)}(t)}, \quad \forall i, \tag{3.1}$$

and similarly for $CpkI^{(i)}(t)$

$$\pi_2^{(i)}(t) = \frac{CpkI^{(i)}(t)}{\sum_j CpkI^{(j)}(t)}, \quad \forall i. \tag{3.2}$$

Another important point to note, is that the time step considered to evaluate the trend of the popularity measure is the week. It appears to us to be a natural subdivision of time for the publishing practices of channels, which have higher rates during the workweek, while are less active on the weekend.

We will show the trends associated to each user engagement measure separately.

Number of visualizations (N_v)

On a video playback platform such as YouTube, the number of views is the first measure we are faced with and the most natural one with which to associate the concept of popularity. The number of views is mostly in the thousands, reaching, in some cases, even over a million.

Figure 3.5 shows the evolution of normalized views $\pi_1^{(i)}(t)$ for each channel i as defined in (3.1), week by week.

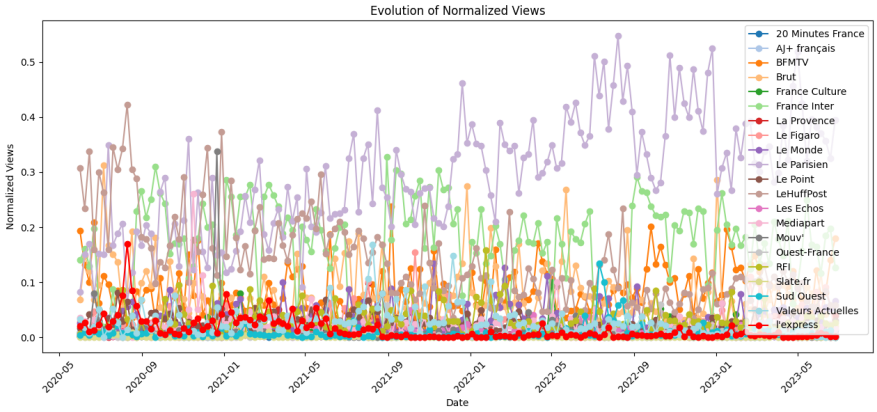


Figure 3.5: Evolution of the popularity measure: normalized number of visualizations

It can already be seen that there are a few channels that prevail in terms of popularity, while many have low and mostly constant values over time.

To better analyze what kind of trends there are, we calculated the time average of normalized views

$$\hat{\pi}_1^{(i)}(T) = \sum_{t=1}^{T} \pi_1^{(i)}(t), \quad \forall i. \tag{3.3}$$

.

Figure 3.6 shows the trends for each channel over time. It can be seen that the most significant growth characterizes *Le Parisien* channel, while *Le HuffPost* channel has a decreasing trend although the initial values are higher. *France Inter*, *Brut* and *BFMTV* are the other most significant channels in terms of popularity, but they maintain that value mostly constant over time.

Comments per thousand impressions (CpkI)

The measure, discussed by Likkanen et al. in [38], beyond the statistically significant results obtained in the paper, appears interesting because it takes into account the number of comments that users leave, but still keeping a reference to the number of views. In fact, in this case, when views are equal, channels with more comments are rewarded. Leaving a comment under a YouTube video, could be interpreted as an even stronger act of engagement, since the user is so taken by the content of the video itself that he decides to react with a comment.

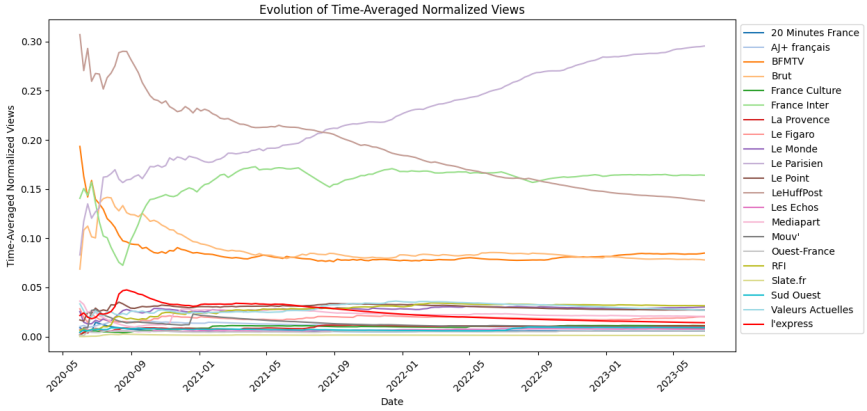


Figure 3.6: Evolution of the time-averaged popularity measure: normalized number of visualizations

The figure 3.7 shows the evolution of normalized comments per thousand impressions $\pi_2^{(i)}(t)$ defined in (3.2), for each channel i week by week.

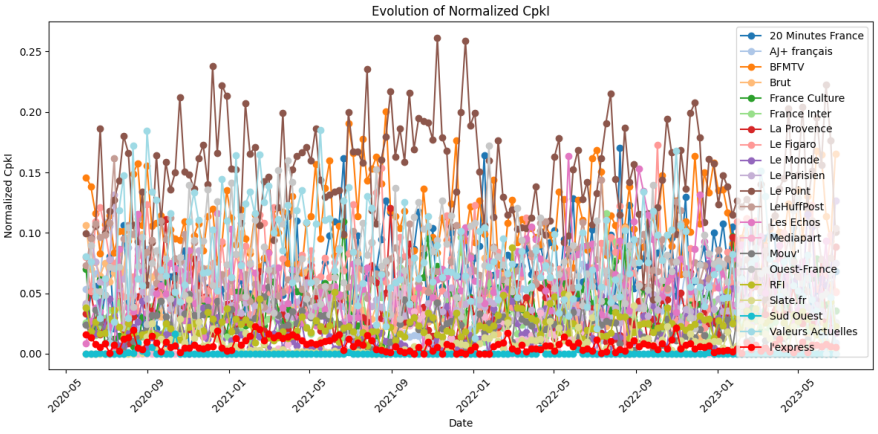


Figure 3.7: Evolution of the popularity measure: normalized CpkI

Again, to better visualize trends, we averaged the measurement over time

$$\hat{\pi}_2^{(i)}(T) = \sum_{t=1}^{T} \pi_2^{(i)}(t), \quad \forall i. \tag{3.4}$$

The figure 3.8 shows a different behavior from that observed by taking into account only views (figure 3.6). In this case, other channels, such as *Le Point* and *Valeurs Actuelles*, are rewarded, while the channels that had high values in terms of $\pi_1(t)$ measure are no longer as attractive (*Le Parisien* and *Le HuffPost*). In fact, linking to what we observed earlier, we expect that more divisive channels such as *Valeurs Actuelles*, which is associated with a

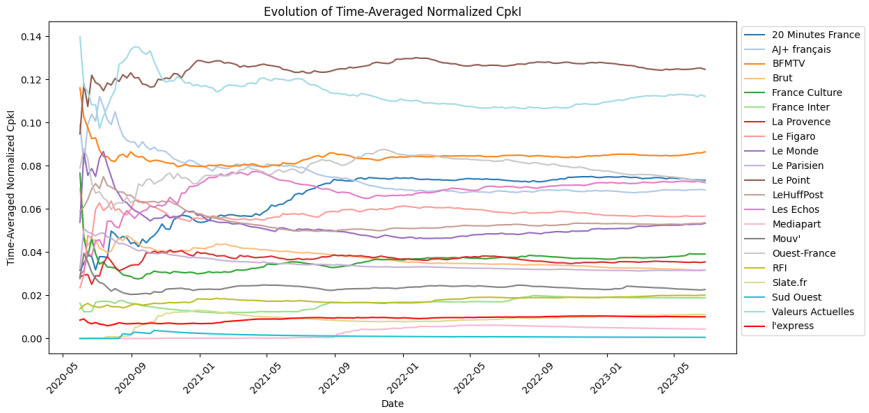


Figure 3.8: Evolution of the time-averaged popularity measure: normalized CpkI

far-right newspaper, would have higher $\pi_2(t)$ values, as they might elicit stronger reactions from users, who are inclined to comment.

We observe, then, that modeling popularity with different measures puts the focus on different effects. We now ask, on what factors might a popularity trend depend?

3.3 External factors

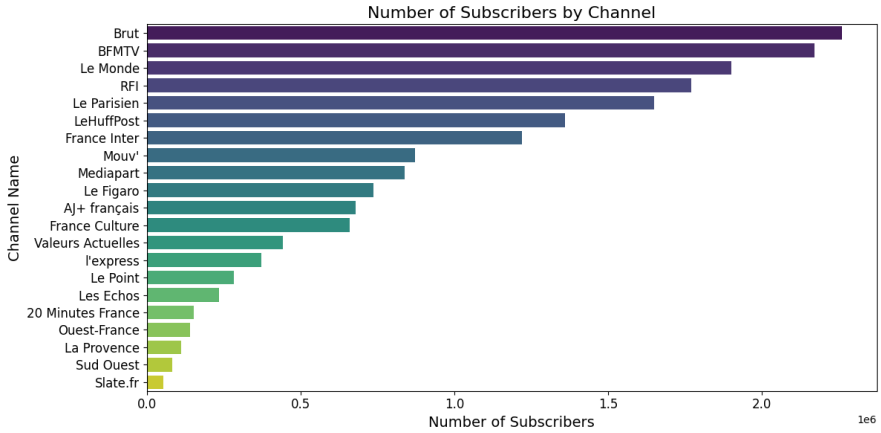
Working in the context of YouTube videos, and specifically wondering how the popularity of the channels associated with them evolves, requires not only the definition of a specific measure of popularity, discussed in the previous section, but also an understanding of what external factors might influence a user in choosing which video to watch. We certainly recognize the presence of a recommendation mechanism of the platform itself but, at this time, we will focus on what might be, based on the literature and our knowledge, the factors external to the network that influence a user's choice.

Several studies have shown that numerous indicators influence the choice to view a video specifically. Many of these, discussed in [39] and [40], are related to the video's content and the speaker's ability to express himself. Since we are reasoning at a higher level (all the aggregated videos of a certain channel) and since all channels refer to the same topics, it does not make much sense to reason about the content, but rather to look for indicators at the level of the channel itself.

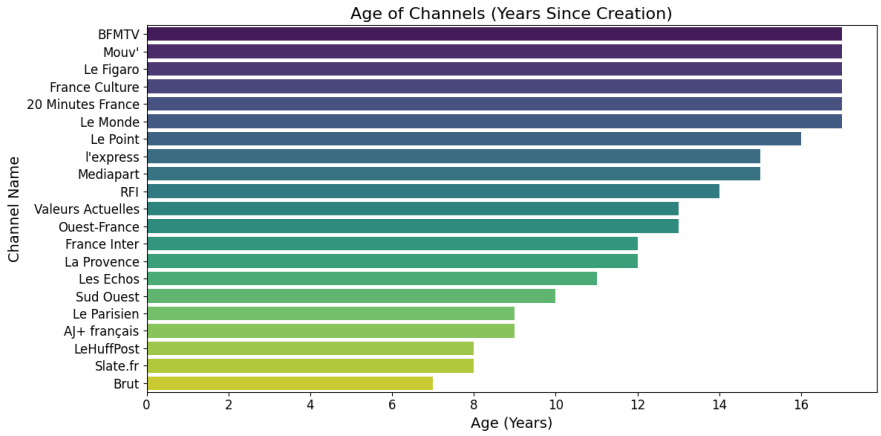
One of the first characteristics that results in increased visibility of a channel is the number of subscribers (Figure 3.9a). Subscribing to a channel on YouTube implies the appearance of videos associated with it on the user's home section, so wondering about subscribers seems natural.

In 2018, Mathias Bärtl published a study [41] highlighting not only the importance of a channel's subscribers but also the channel's age. He observed that older channels that had been on the platform longer were associated with higher visibility. Hence the idea that another indicator can be related to the age of the channel (Figure 3.9b).

Other elements can come into play in the increased visibility of a channel: publication



(a) Subscribers of each channel



(b) Years of activity of each channel

Figure 3.9: Features of channels

rate, publication consistency, and duration of videos [39]. Regarding the former, we refer to how many videos are posted by each channel in a given time interval, which for us turns out to be the week, in order to align with the time window used in the section 3.2.

By *publication consistency*, however, we refer to how the publication activity varies week by week. To understand the strategy for each channel, we have defined a measure that allows us to give an idea of when a certain creator publishes. For every channel i, it is defined as

$$m^{(i)}(w) = \frac{1}{7}\sum_{g=1}^{7} \mathbb{1}_{w_g}^{(i)}, \quad \forall i,$$

where w represents the specific week, w_g is the day of the week and

$$\mathbb{1}_{w_g}^{(i)} = \begin{cases} 1 & \text{if the channel publishes on day } w_g, \\ 0 & \text{otherwise.} \end{cases}$$

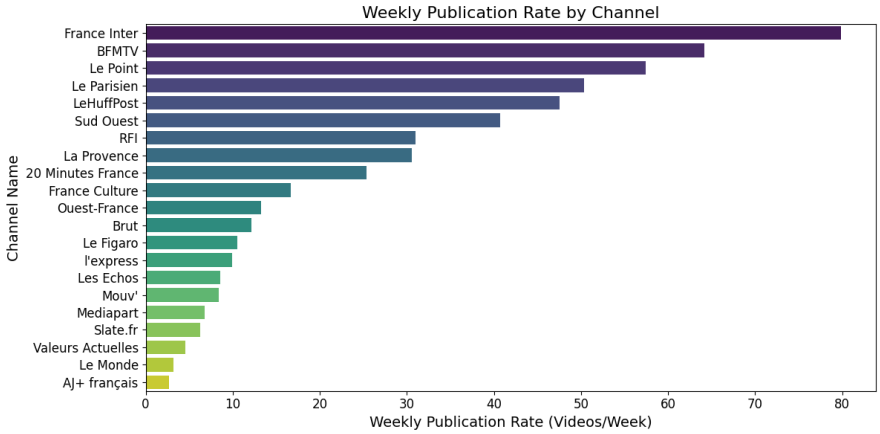


Figure 3.10: Publication rate per week for each channel

In other words, it is the fraction of days of the week in which the channel is active. By considering the standard deviation of this measure as an *index of variation* of publications over time, we obtain information about which ones have the most constant activities. Figure 3.11 shows that all channels have unstable publication activities that vary over time, but some of them remain more constant, like *Le Parisien* and *Le HuffPost*.

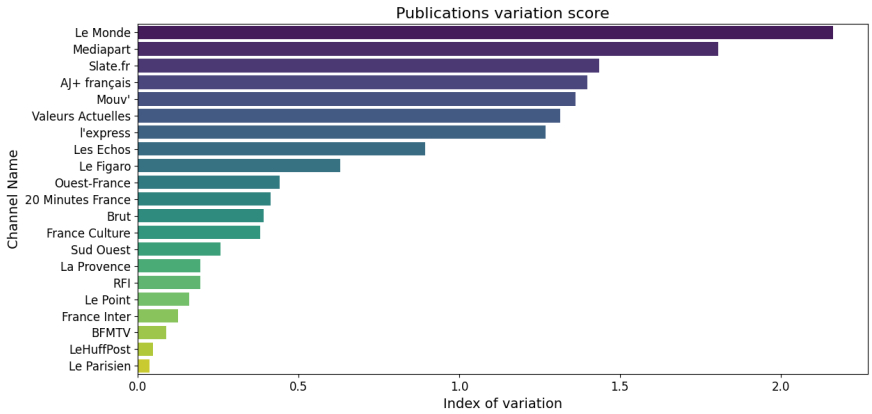


Figure 3.11: Variation index for each channel

Welbourne and Grant in [39] included video duration in the factors that influence a channel's popularity and visibility observing that shorter videos are viewed more than longer videos. In this context, what we studied first was how videos are distributed by duration for each channel: how many videos are published for each type of duration? At the same time, we were interested in knowing how views are distributed with respect to duration for each channel. The figures 3.12a and 3.12b show these distributions, using the natural logarithm of duration for easier reading. It can be seen that most channels show a peak of video publications around 1.3 minutes. Views versus duration, on the other hand,

show different behaviors depending on the channel.

Rather than considering only the length of the videos, as proposed in [39], we observed how one possible indicator might be the alignment between users' expectations (what types of videos in term of duration are most viewed?) and the type of video proposed (what types of videos in term of duration are most published?). To understand how well video production aligns with user preferences, we compared the two distributions (number of videos posted per duration and number of views per duration showed in figures 3.12a and 3.12b) using a statistical test known as *Jensen-Shannon Divergence* (JSD). JSD is a symmetric, bounded metric (ranging from 0 to 1) that quantifies how different two probability distributions are:

- A JSD close to 0 indicates high similarity (better alignment).
- A JSD closer to 1 indicates significant divergence (poor alignment).

Figure 3.13 shows the behavior of channels in terms of this similarity measure: *AJ+ francais* is the channel that most closely aligns with user preferences.

Analysis of correlations

Given the external factors of our interest, it is important to understand what kind of relationship there is, if any, between them and the user engagement measures introduced in the section 3.2.

We begin our study with the first measure of user engagement $\pi_1^{(i)}(t)$, defined in (3.1), which corresponds to the normalized number of views of each channel i. Figure 3.6 showed that each channel converges over time to a specific popularity value or generally exhibits a certain trend. To assess correlation, we took values at convergence of cumulative popularity $\hat{\pi}_1(t)$ (defined in (3.3)), using a threshold of 0.005 or taking the last available value (if there was no true convergence). The limit value of this cumulative popularity will be called $\hat{\pi}_1^{(i)\star}$, for all channel i.

Table 3.1 shows the limit popularity value for each channel.

Table 3.1: Limit Popularity $\hat{\pi}_1^{(i)\star}$ of each channel

(a) Part 1

Channel	Popularity
20 Minutes France	0.0056
AJ+ français	0.0097
BFMTV	0.0849
Brut	0.0780
France Culture	0.0111
France Inter	0.1641
La Provence	0.0107

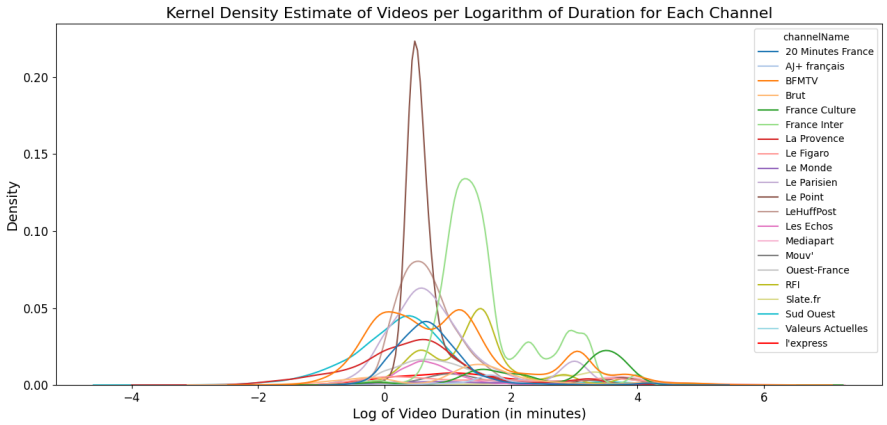
(b) Part 2

Channel	Popularity
Le Figaro	0.0201
Le Monde	0.0301
Le Parisien	0.2955
Le Point	0.0270
LeHuffPost	0.1381
Les Echos	0.0076
Mediapart	0.0206

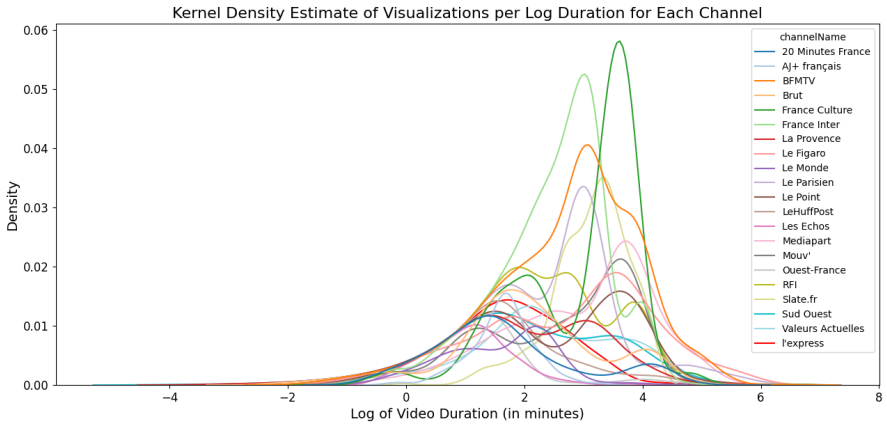
(c) Part 3

Channel	Popularity
Mouv'	0.0087
Ouest-France	0.0055
RFI	0.0313
Slate.fr	0.0011
Sud Ouest	0.0093
Valeurs Actuelles	0.0268
l'express	0.0141

Once these values were obtained, we evaluated the correlation between $\hat{\pi}_1^{(i)\star}, \forall i$ and the external factors listed above. Figure 3.14 shows the results in terms of correlation.



(a) Number of videos published per log of duration



(b) Number of visualizations per log of duration

Figure 3.12: Analysis of duration

It can be seen that the publication rate and the number of subscribers are both fairly positively correlated with $\hat{\pi}_1^{(i)\star}$ while the change in publication practices and channel years have a negative correlation. In any case, as we would expect, none of the external variables has a very very strong correlation with the popularity measure. The alignment index and average video duration are the least influential indicators.

Similarly to what was done for measure $\pi_1^{(i)}(t)$, this analysis can be repeated for the second measure of our interest, namely $\pi_2^{(i)}(t)$ which corresponds to the number of comments for thousand impressions normalized (3.2). We had already found substantial differences between the two measures in the section 3.2, as they seemed to reward different aspects of the channels. For this reason, the limit values of cumulative popularity $\pi_2^{(i)\star}$(obtained as for the $\pi_1^{(i)\star}$ case) are markedly different, as shown in the table 3.2.

Correlation analysis with the various external indices again confirms the differences from the previous popularity measure $\pi_1^{(i)}(t)$. Figure 3.15 shows that in this case the most

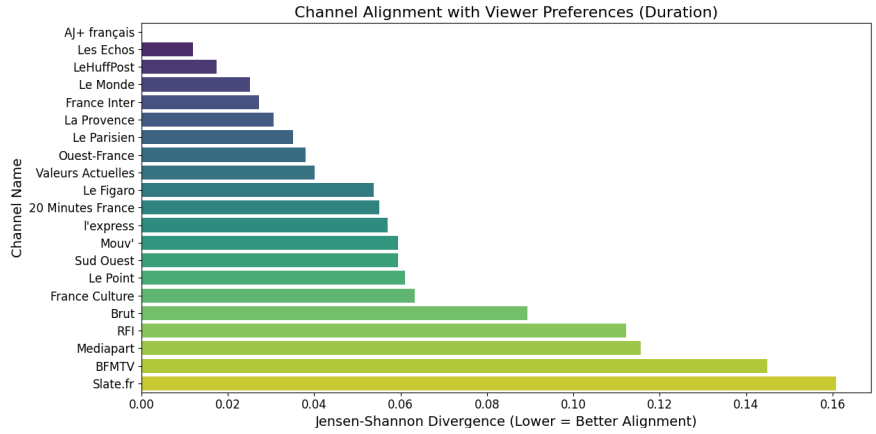


Figure 3.13: Similarity in terms of videos produced and users preferences

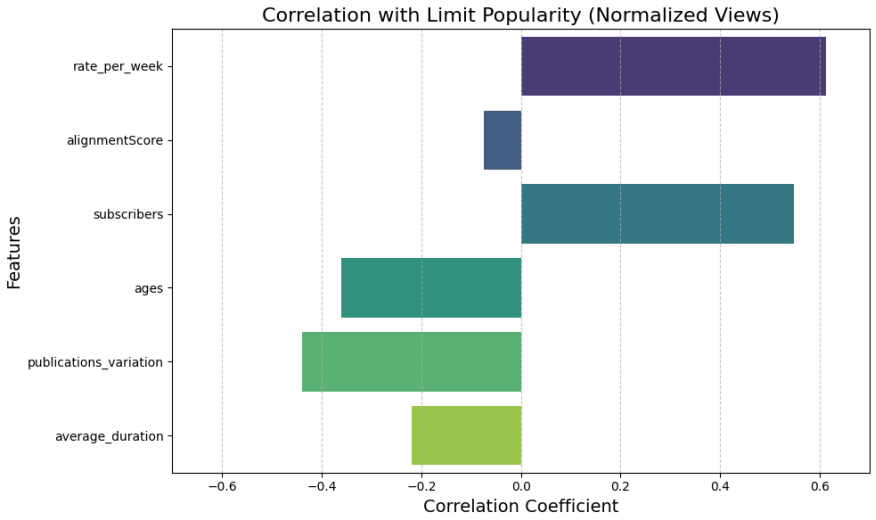


Figure 3.14: Correlations between $\hat{\pi}_1^\star$ and external factors

significant indicators in terms of correlation are the years of channel activity (positive) and the alignment index of the type of videos published in terms of duration (negative). Here again, we do not find high correlations. The results appear significant: the negative correlation with the alignment index and average video duration shows that videos that align less with audience preferences or appear longer and harder to follow encounter a higher value of comments (since $\pi_2^{(i)}(t)$ also rewards the presence of comments).

Features Selection and Linear Regression

Having understood the most interesting indicators for each popularity measure, we worked on a linear combination of these factors to understand whether it is possible to estimate

Table 3.2: Limit Popularity $\hat{\pi}_2^{(i)\star}$ of each channel

(a) Part 1

Channel	Popularity
20 Minutes France	0.0733
AJ+ français	0.0688
BFMTV	0.0865
Brut	0.0318
France Culture	0.0391
France Inter	0.0188
La Provence	0.0355

(b) Part 2

Channel	Popularity
Le Figaro	0.0567
Le Monde	0.0534
Le Parisien	0.0317
Le Point	0.1247
LeHuffPost	0.0536
Les Echos	0.0722
Mediapart	0.0044

(c) Part 3

Channel	Popularity
Mouv'	0.0227
Ouest-France	0.0727
RFI	0.0200
Slate.fr	0.0111
Sud Ouest	0.0005
Valeurs Actuelles	0.1122
l'express	0.0100

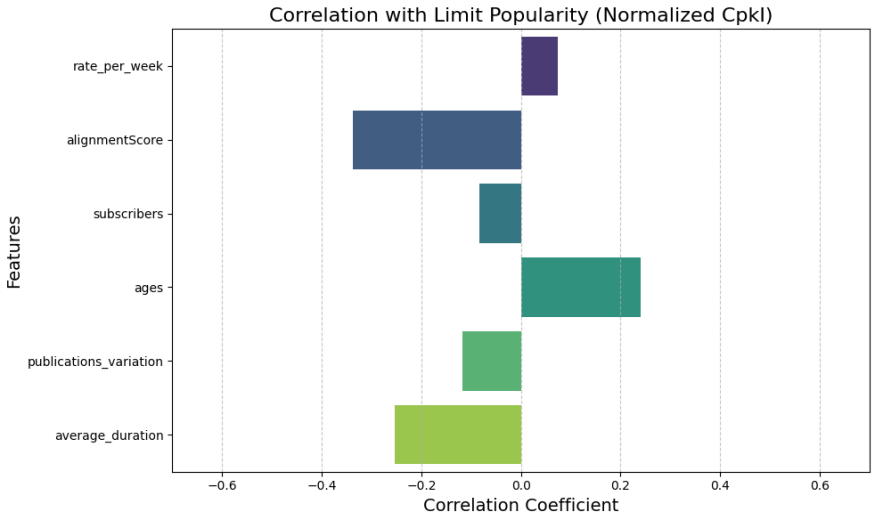


Figure 3.15: Correlations between $\hat{\pi}_2^{\star}$ and external factors

the limit value of popularity. In other words, we implemented a linear regression via *MSE minimization*. Once the general popularity estimation $\hat{\Pi}^{(i)\star}$ is obtained

$$\hat{\Pi}^{(i)\star} = \alpha_1 \text{feature}_1^{(i)} + \alpha_2 \text{feature}_2^{(i)} + \cdots + \alpha_n \text{feature}_n^{(i)},$$

the *Mean Squared Error (MSE)* is defined as

$$\text{MSE}(\hat{\pi}^{\star}, \hat{\Pi}^{\star}) = \frac{1}{|I|} \sum_{i \in \mathcal{I}} (\hat{\pi}^{(i)\star} - \hat{\Pi}^{(i)\star})^2,$$

where $|I|$ represents the number of channels. Thus, our problem for both measures is to find

$$\text{argmin}_{\alpha_1,\ldots,\alpha_n} \text{MSE}(\hat{\pi}^{\star}, \hat{\Pi}^{\star}).$$

Before proceeding with the regression, we carried out a features selection operation, as we were interested in understanding whether not only the indicators presented at the

beginning of this section could be significant, but also their possible interactions or transformations. The latter would have helped to include nonlinearities in the model. To do this we could have proceeded by brute force, performing regression including possible linear combinations and transformations of features in addition to the standard indicators. Since this was a very time-consuming procedure, we tried to use a Machine Learning model, XGBoost, aware that our dataset is very very small and that the results obtained would not have been very meaningful mainly due to overfitting. We then used the results obtained from the XGBoost very cautiously and manually checked that including or excluding a certain feature resulted in substantial changes in terms of MSE. XGBoost (eXtreme Gradient Boosting) is an open-source software library that provides regularizing gradient boosting. It is based on the use of multiple decision trees (boosting) and takes advantage of gradient descent. New trees are created from previous trees based on past errors. It is very useful for obtaining the feature importance score because it is computed for a single decision tree by the amount that each attribute's split point improves the performance measure.

For both limit measures of popularity $\hat{\pi}_1^\star$ and $\hat{\pi}_2^\star$, the only additional feature found to be relevant is the interaction between the number of subscribers and the publication rate, obtained by multiplying the respective indicators. This new feature can be explained by the fact that the publication rate can have different effects depending on whether the audience is larger or smaller. Consequently, it is not enough to consider only individual factors. All other interactions and transformations tried did not lead to interesting results in terms of MSE reduction.

It is worth noting that since the popularity limit value is a normalized value, all features used in the regression model were first normalized using min-max scaling, i.e.

$$\text{feature\_norm}^{(i)} = \frac{\text{feature}^{(i)} - \min_i \text{feature}^{(i)}}{\max_i \text{feature}^{(i)} - \min_i \text{feature}^{(i)}}$$

where $\text{feature}^{(i)}$ represents the value of a generic feature for the channel i. In this way, we ensure that the model considers them on equal scale and prevent features with large ranges from disproportionately influencing the results.

As mentioned earlier, a regression analysis was performed on both popularity measures. In both cases, the features used are as follows: *number of channel subscribers, weekly publication rate, years of activity, publications variation index, average video duration, alignment index in terms of duration, interaction between number of subscribers and publication rate.*

In the case of the second popularity measure $\hat{\pi}_2^\star$, based on the comments per thousand impressions (CpkI), given also the low correlations observed with the various indicators available, we were unable to obtain statistically valid and interesting results. Going further down this path probably requires additional features that we are ignoring and that are not available in this moment.

Instead, we can observe what happens in the case of the first popularity measure, based on normalized views. Table 3.3 shows the results of the regression analysis for $\hat{\pi}_1^\star$: we obtain a small MSE that demonstrates the model can correctly approximate the popularity values.

Table 3.3: Regression Results for $\hat{\pi}_1^\star$

Features	**Value**
Subscribers: $\boldsymbol{\alpha}_1$	0.021
Ages: $\boldsymbol{\alpha}_2$	0.001
Publ. rate: $\boldsymbol{\alpha}_3$	0.038
Index of alignment: $\boldsymbol{\alpha}_4$	0.001
Interactions (subscribers, publ. rate): $\boldsymbol{\alpha}_5$	0.177
Publications variation: $\boldsymbol{\alpha}_6$	0.001
Average duration: $\boldsymbol{\alpha}_7$	0.001
Mean Squared Error (MSE)	0.0024
Spearman Rank Correlation value	0.8013

Among the results, in addition to the values of the alphas that minimize the MSE and that again demonstrate that the most relevant features are the number of subscribers and the publication rate (and their interaction), there is an additional statistical index, known as the *Spearman Correlation Rank*. It measures the rank correlations, namely the statistical dependence between the rankings of two variables (8.1.4 in Appendix). As we are interested in knowing the ranking of channels in terms of popularity from the beginning, the Spearman Correlation index allows us to understand whether this ranking is also maintained by the model. Even if the MSE is small, since the popularity values are all very small and close, reasoning about the ranks allows us to better understand whether the model performs well or not. Obtaining a Spearman Rank correlation value of 0.8, we can conclude that the model associates similar relevance in terms of popularity to the real one.

Figure 3.16 shows even better how the model performs compared to reality: the dashed line shows the trend that would occur if the statistical model performed perfectly and predicted the popularity values perfectly. It can be seen that except for a few cases, there is little deviation from this ideal trend.

3.4 Final Observations

The empirical analysis yielded several interesting results. We used two different measures of popularity from the information that YouTube makes available to us, noting that each leads to the prevalence of certain channels and features rather than others. We reasoned about what factors might be influencing channel popularity, doing a literature review and combining it with available data.

Our analysis was purely exploratory of the dataset at hand, to understand what kind of behaviors channels have over time and how to characterize them. This initial exploration provides a foundation for identifying key patterns and formulating hypotheses that can be rigorously tested. More representative datasets can be used in the future to statistically validate our theoretical model.

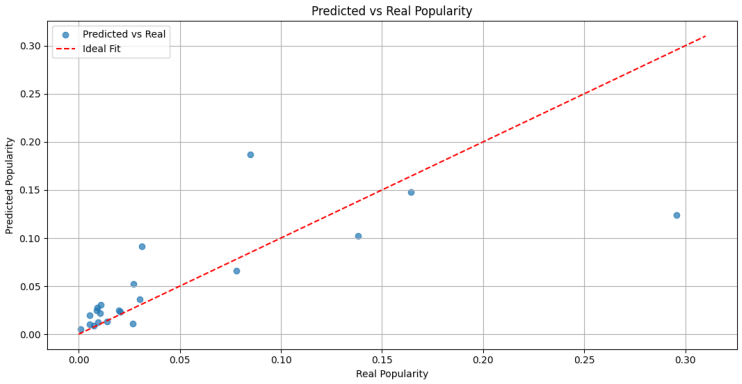


Figure 3.16: Behavior of predicted popularity vs real popularity

Chapter 4

Deterministic Modeling of Popularity Dynamics

This chapter aims to define the deterministic model, which later we will use to approximate the expectation of the general stochastic model. We will analyze the different special cases obtained by canceling one or more terms and obtain theoretical results that will define the conditions for long-run convergence. For each of these, we will analyze the behavior numerically.

4.1 Preliminaries

Mathematical graph theory is fundamental to analyzing structures of interacting agents, as in social networks.

A graph is a mathematical structure defined as

$$G = (\mathcal{V}, \mathcal{E}, \mathcal{W}),$$

where

- $\mathcal{V}$ is the set of *nodes* that represents the units participating in the network.
- $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of *links*. $(i, j) \in \mathcal{E}$ corresponds to the link between the nodes i and j. A link of the type (i, i) is called *self-loop*.
- $\mathcal{W} \in \mathbb{R}^{\mathcal{V} \times \mathcal{V}}$ is the *weight* matrix where $W_{ij} > 0$ indicates the strength of the edge (i, j). $W_{ij} = 0$ if and only if $(i, j) \notin E$, i.e., if there is no edge (i, j).

The meaning of the link is different depending on the graph's application. For example, in the study of social network dynamics, it represents the influence of one agent on another, but it could also represent a flow or other type of communication. Similarly, the link's weight shows the connection's strength, which can still be interpreted differently depending on the context of reference.

In the following sections, we will always refer to graphs that are *undirected* and *unweighted*:

- An *undirected* graph is a graph in which the presence of the edge corresponds to a symmetric relation, i.e., if (i, j) is present, automatically (j, i).
  The weight matrix W associated to an undirected graph is always simmetric.
- An *unweighted* graph is characterized by a weight matrix W s.t. $W_{ij} = 1$ if the edge is present, $W_{ij} = 0$ if it is not. In this case, the weight matrix is also known as *adiacency matrix*. We can rewrite this condition as,

$$W_{ij} = 1 \text{ for every } (i, j) \in \mathcal{E}$$

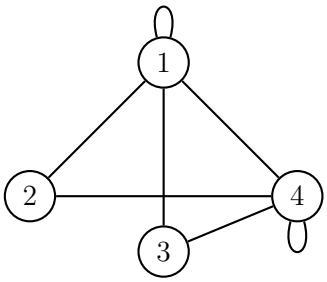


Figure 4.1: Generic undirected and unweighted graph with 4 states

Additional definitions are required to improve the description of a graph:

- *Out-neighborhood*: Set of nodes to which links exit from i, i.e., $\mathcal{N}_i = \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}\}$
- *In-neighborhood*: Set of nodes from which links start towards i, i.e., $\mathcal{N}_i^- = \{j \in \mathcal{V} \mid (j, i) \in \mathcal{E}\}$

In an undirected graph, the two sets coincide. When $\mathcal{N}_i = \{\emptyset, \{i\}\}$, i is called *absorbing state*.

When we work on a graph we can always say that j is reachable from i if there exists a walk from i to j, where a *walk* is defined as $\gamma = (i = i_0, i_1, ..., j = i_l)$ such that $(i_h, i_{h+1}) \in \mathcal{E}$ for all $h = 0, \ldots, l-1$. We can recognize a *strongly connected* graph if for all i and j, i is reachable from j.

A walk $(i_0, i_1, \ldots, i_l)$ such that $i_h \neq i_k$ for all $0 \leq h < k \leq l$, except possibly $i_0 = i_l$ is called *path*. Finally, a closed path of length $l \geq 3$ is known as *cycle*. An *acyclic* graph has no cycles.

A graph can be further characterized in terms of periodicity. Given a node $i \in \mathcal{V}$, the *period* of i is

$$\text{per}(i) := \text{g.c.d.}\{l| \ \exists \text{ closed walk of length } l \text{ in } i\}$$

If the graph $\mathcal{G}$ is strongly connected, the period is the same for each node $i \in \mathcal{V}$, so we can attribute that level of periodicity directly to $\mathcal{G}$.

$$\mathcal{G} \text{ is called } \textit{aperiodic} \text{ if } per(\mathcal{G}) = 1.$$

It is interesting to note that the graph is aperiodic if there are self-loops in $\mathcal{G}$.

Special types of graphs are:

- *Complete graph*: it possesses a link for each pair of nodes; in other words, the nodes of a complete graph are directly connected to all others.
- *Regular graph*: an undirected graph where each vertex has the same number of out-neighbors (or in-neighbors as they coincide in this case).
- *Erdős–Rényi graph*: it is a random graph computed by imposing that a link has a specific probability to be present or absent, independently of the other arcs.

4.2 Mathematical Framework

We consider a set of users $\mathcal{V}$; each of them can access the content of a set $\mathcal{I}$ of channels at each instant of time $t \in \mathbb{N} \cup \{0\}$. The number of users is $n = |\mathcal{V}|$, while the number of channels is $m = |\mathcal{I}|$.

Each user $v \in \mathcal{V}$ is endowed with a scalar value $x_v^{(i)}(t) \in [0,1]$, which quantifies the *interest* by user $v \in \mathcal{V}$ at time t relative to content posted by channel $i \in \mathcal{I}$. Our objective is to develop a model to describe the evolution of this appreciation over time, capturing the interplay of the following factors:

- the network effects, due to the social imitation behavior of agents in a social system;
- the effect of recommendations, proposed by a platform, such as YouTube, that can suggest videos and contents tailored to a users' viewing history, preferences, and interactions;
- external factors, i.e. all factors that are independent of the network or platform.

Modeling network effects

By network effects we mean the phenomenon whereby a user's appreciation of a piece of content is influenced by the direct (because they belong to their own social network) or indirect appreciation of other users due to long-range interaction. This phenomenon emphasizes how individual preferences are not independent, but are instead shaped by social imitation, the communication we are immersed in, and interactions with others.

The idea that users can interact with each other and exchange opinions can be formalized using a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ represent the agents in a social network and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ describe the potential interactions. The type of the graph characterizes the kind of interaction between users (e.g., the *complete graph* corresponds to an interaction of everyone with everyone). We will use a *weighted-adjacency matrix* P, associated with a graph, in order to encode the intensity of these interactions. Formally, we will say that $w \in \mathcal{V}$ influences $v \in \mathcal{V}$ if and only if $P_{vw} > 0$. The matrix P is normalized with respect to the sum on the rows in such a way as to make it rows-stochastic while preserving its structure, i.e. $P\mathbb{1}_n = \mathbb{1}_n$. Figure 4.2 shows an example of an undirected graph to simulate interactions between 6 users. The numbers on the links represent weights associated to the connections.

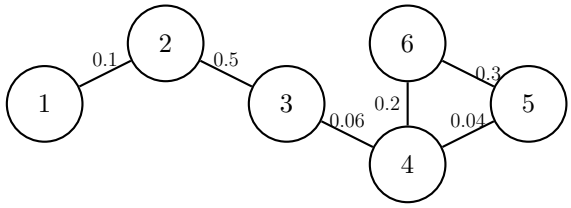


Figure 4.2: Example of graph to simulate users' interactions

Modeling recommendations

Recommendations are suggestions that are algorithmically generated by the platform with the goal of increasing participation and appreciation of the suggested content. Although these algorithms are not disclosed by the platform, they can be expected to leverage user preferences and the appreciation that users have shown in the past toward similar proposed content, i.e. the previous popularity of the channels. Therefore, the recommendation acts as a mechanism known in control theory as the feedback mechanism, which exhibits a feedback effect (Figure 4.3). This mechanism can significantly impact users' appreciation, either guiding them toward new content or reinforcing existing preferences [33].

In our mathematical model, the effect of recommendations will be represented through a metric, a popularity index $\pi^{(i)}(t)$, which is linked to an aggregate measure of the appreciation of previous channel content. Moreover, we propose to use a normalized index in order to take into account the competitive nature of the channels, i.e. a comparative measure of how well a channel performs relative to others. Formally, we will consider $\pi^{(i)}(t) = f(\{x_v^{(i)}(t)\}_{v\in\mathcal{V},i\in\mathcal{I}})$ with the property $\sum_{i\in\mathcal{I}} \pi^{(i)}(t) = 1$ for all $t \in \mathbb{N}\cup\{0\}$.

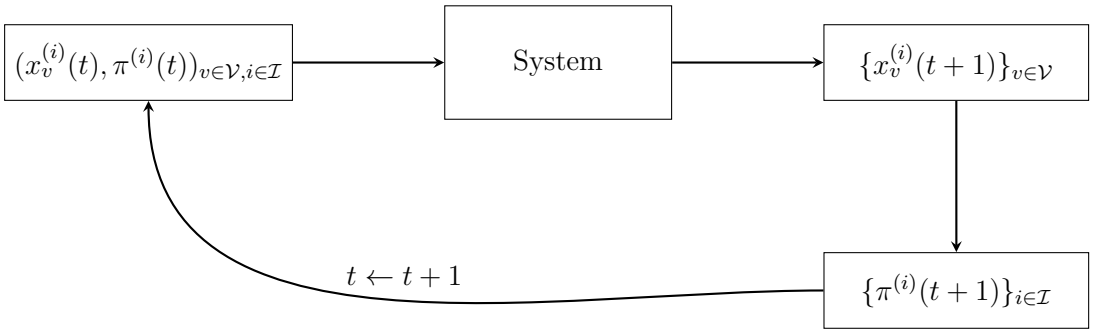


Figure 4.3: Feedback effect of the raccomandation

Modeling the external factors

We suppose that the external factors that influence the dynamics of appreciation are constant and do not depend on the time. This is an oversimplification and considering

time-varying factors will be subject of future work. We suppose that the external factors, such as the competence of content creators posting in the channel, or their expertise, act as persistent input in the dynamics. This measure of quality will be denoted by $q^{(i)} \in [0,1]$.

4.3 Dynamics Definition and Theoretical Analysis

The dynamic of the appreciation $x_v^{(i)}(t)$ evolves according to the following equation for all $v \in \mathcal{V}$:

$$x_v^{(i)}(t+1) = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} x_w^{(i)}(t) + \beta_v \pi^{(i)}(t) + \gamma_v q^{(i)}, \tag{4.1}$$

where for all $i \in \mathcal{I}$ the ranking $\pi^{(i)}(t)$ of channel i at time t is defined as:

$$\pi^{(i)}(t) = \frac{\sum_{v \in \mathcal{V}} x_v^{(i)}(t)}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} x_w^{(j)}(t)}.$$

The contribution of interactions, the impact of the popularity and the influence of expertise are weighted through the coefficients α_v, β_v and γ_v, so that each user v can give higher or lower priority to them. These coefficients must satisfy the following normalization condition for each user $v \in \mathcal{V}$:

$$\alpha_v + \beta_v + \gamma_v = 1.$$

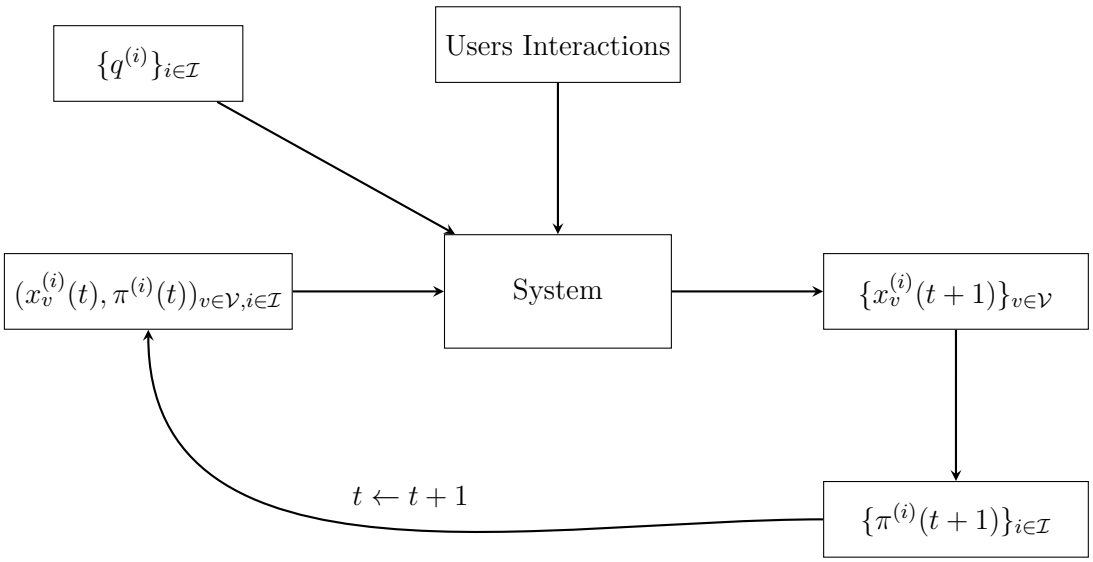


Figure 4.4: Evolution of the dynamics

Figure 4.4 shows the structure and terms that define the model.

In the next sections, we will analyze some particular cases separately. More precisely, we will investigate whether the dynamics admit equilibrium points and eventually study their stability for different parameter values.

- In section 4.3.1 we will analyze the case of $\alpha_v = 0\ \forall v \in \mathcal{V}$. It corresponds to the case where user interaction is not considered in the appreciation's evolution. This first case will show clear behavior at limit, in which the value at the convergence of $\pi^{(i)}(t)$ will depend on the external factor associated with channel i.
- Section 4.3.2 will instead be devoted to the diametrically opposite case, which occurs when $(\alpha_v\, \beta_v, \gamma_v) = (1,0,0)$ for all $v \in \mathcal{V}$, that is, when social network represents the only contribution into the dynamics. The dynamics observed are that of the well-known *French De Groot* model [19], from whose study the stability results will be obtained.
- The case with $\beta_v = 0$ for all $v \in \mathcal{V}$ will be studied in section 4.3.3, and corresponds to considering only the network effect and the external factors into the dynamics. Here again, we will find the dynamics of a well-known model in the literature, *Friedkin Johnsen*'s model [20].
- Section 4.3.4 will be devoted to analyzing the case with $\gamma_v = 0$ for all $v \in \mathcal{V}$, to understand what happens when only network effects and recommendations influence the dynamics. The dynamics will show different behavior from previous cases, but they will still reach convergence. It will be noted that in this case there is a consensus at the limit, obtained from a reformulation of the problem.
- Finally, in section 4.3.5, we will study what happens when all terms give contribution into the dynamics, namely, $\alpha_v \neq 0, \beta_v \neq 0, \gamma_v \neq 0$ for some $v \in \mathcal{V}$, looking at the model (4.1) in its entirety. We will again find convergence at the limit and characterize the values of $\pi^{(i)}(t)$ and $x_v^{(i)}(t)\ \forall v \in \mathcal{V},\ \forall i \in \mathcal{I}$ that are obtained for $t \to \infty$.

It is worth remarking that although we will show the analysis of all the cases listed above, our contribution is mainly related to the study of the cases not known in the literature, i.e., those corresponding to sections 4.3.1, 4.3.4 and 4.3.5.

4.3.1 Asymptotic Behavior Without Network Effects

Within the dynamics described by equation (4.1), each future appreciation value is influenced by user interactions in the network described by P.

Imposing that $\alpha_v = 0\ \forall v \in \mathcal{V}$, it cancels out the contribution of interactions, retaining only the influence from past popularity values and expertise. By taking into account that in this case $\gamma_v = 1 - \beta_v$, we can rewrite the dynamics in (4.1) as follows:

$\forall i \in \mathcal{I}, \forall v \in \mathcal{V}$

$$\begin{aligned} x_v^{(i)}(t+1) &= \beta_v \pi^{(i)}(t) + (1-\beta_v) q^{(i)}, \\ \pi^{(i)}(t+1) &= \frac{\sum_{v\in\mathcal{V}} {x_v}^{(i)}(t)}{\sum_{j\in\mathcal{I}} \sum_{w\in\mathcal{V}} {x_w}^{(j)}(t)}. \end{aligned} \tag{4.2}$$

The following proposition guarantees that for any initial condition the dynamics in (4.2) converges to a limit point.

Proposition 1. *Let $\alpha_v = 0$ for all $v \in \mathcal{V}$. If $\exists\ v \in \mathcal{V}$ s.t. $\beta_v \neq 1$ and $\exists\ i \in \mathcal{I}$ s.t. $q^{(i)} \neq 0$, then, for any initial condition $\pi(0) = \pi_0$, the dynamics in (4.2) is stable. In addition, $\forall i \in \mathcal{I}$, $\forall v \in \mathcal{V}$, we have*

$$\lim_{t\to\infty} \pi^{(i)}(t) = \frac{q^{(i)}}{q_{\text{tot}}}, \quad \lim_{t\to\infty} x_v^{(i)}(t) = \beta_v \frac{q^{(i)}}{q_{\text{tot}}} + (1-\beta_v)q^{(i)},$$

where $q_{tot} = \sum_{i\in\mathcal{I}} q^{(i)}$

Proof. By denoting

$$\overline{\beta} = \frac{1}{|\mathcal{V}|}\sum_{v\in\mathcal{V}} \beta_v, \qquad q_{tot} = \sum_{i\in\mathcal{I}} q^{(i)},$$

and the average value on the set of users $\mathcal{V}$ of $x_v^{(i)}(t)$

$$\overline{x}^{(i)}(t) = \frac{\sum_{v\in V} x_v^{(i)}(t)}{|\mathcal{V}|},$$

it is possible to rewrite $\pi^{(i)}(t)$ as

$$\pi^{(i)}(t) = \frac{\overline{x}^{(i)}(t)}{\sum_{i\in\mathcal{I}} \overline{x}^{(i)}(t)}.$$

Exploiting it in the equation of the dynamics (4.2), we obtain

$$x_v^{(i)}(t+1) = \beta_v \pi^{(i)}(t) + (1-\beta_v)q^{(i)} = \beta_v \frac{\overline{x}^{(i)}(t)}{\sum_{i\in\mathcal{I}} \overline{x}^{(i)}(t)} + (1-\beta_v)q^{(i)},$$

that can be used to express the dynamics for $\pi^{(i)}(t+1)$. In fact, noting that

$$\sum_{v\in\mathcal{V}} x_v^{(i)}(t+1) = \frac{\overline{x}^{(i)}(t)}{\sum_{i\in\mathcal{I}} \overline{x}^{(i)}(t)} \sum_{v\in\mathcal{V}} \beta_v + q^{(i)} \sum_{v\in\mathcal{V}} (1-\beta_v)$$

$$\overline{x}^{(i)}(t+1) = \frac{\overline{x}^{(i)}(t)}{\sum_{i\in\mathcal{I}} \overline{x}^{(i)}(t)} \overline{\beta} + q^{(i)}(1-\overline{\beta}),$$

and, again exploiting the previous equation,

$$\sum_{i\in I} \overline{x}^{(i)}(t+1) = \frac{\sum_{i\in I} \overline{x}^{(i)}(t)}{\sum_{i\in\mathcal{I}} \overline{x}^{(i)}(t)} \overline{\beta} + (1-\overline{\beta}) \sum_{i\in I} q^{(i)}$$

$$\sum_{i\in I} \overline{x}^{(i)}(t+1) = \overline{\beta} + (1-\overline{\beta})q_{tot}.$$

Given these simplifications, it is possible now to write an equation of dynamics for $\pi^{(i)}(t+1)$

$$\pi^{(i)}(t+1) = \frac{\overline{\beta}}{\overline{\beta} + (1-\overline{\beta})q_{tot}} \pi^{(i)}(t) + \frac{(1-\overline{\beta})}{\overline{\beta} + (1-\overline{\beta})q_{tot}} q^{(i)}.$$

We can recall

$$a = \frac{\overline{\beta}}{\overline{\beta} + (1-\overline{\beta})q_{tot}} \quad \text{and} \quad b = \frac{(1-\overline{\beta})}{\overline{\beta} + (1-\overline{\beta})q_{tot}},$$

so the equation becomes

$$\pi^{(i)}(t+1) = a\pi^{(i)}(t) + bq^{(i)},$$

from which recursively we have

$$\pi^{(i)}(t+1) = a(a\pi^{(i)}(t-1) + bq^{(i)}) + bq^{(i)} = a^{t+1}\pi^{(i)}(0) + \sum_{s=0}^{t} a^s bq^{(i)}.$$

It is important to note the conditions to have $|a| < 1$:

$$a = \frac{\overline{\beta}}{\overline{\beta} + (1-\overline{\beta})q_{tot}} < 1 \iff \overline{\beta} < \overline{\beta} + (1-\overline{\beta})q_{tot}$$

$$\iff (1-\overline{\beta})q_{tot} > 0 \iff q_{tot} \neq 0,\ \overline{\beta} < 1$$

$$\iff q_{tot} \neq 0, \sum_{v\in\mathcal{V}} \beta_v < |\mathcal{V}|.$$

To meet the latter two conditions we need only require that an channel i exists s.t. $q^{(i)} \neq 0$ and an user v exists s.t. $\beta_v \neq 1$.

At this point, given the assumptions to have $|a| < 1$, we can easily derive the value of $\pi^{(i)}(t+1)$ for $t \to \infty$ remembering the convergence of a finite geometric series

$$\lim_{t\to\infty} \pi^{(i)}(t+1) = \sum_{s=0}^{\infty} a^s bq^{(i)} = \frac{1}{1-a} bq^{(i)},$$

so in the end, recalling the definition of a and b,

$$\pi^{(i)\star} = \frac{\cancel{\overline{\beta} + (1-\overline{\beta})q_{tot}}}{(1-\overline{\beta})q_{tot}} \frac{(1-\overline{\beta})}{\cancel{\overline{\beta} + (1-\overline{\beta})q_{tot}}} q^{(i)} = \frac{(1-\overline{\beta})}{(1-\overline{\beta})q_{tot}} q^{(i)} = \frac{q^{(i)}}{q_{tot}},$$

and automatically by the dynamics in (4.2), we have

$$x_v^{(i)\star} = \beta_v \frac{q^{(i)}}{q_{tot}} + (1-\beta_v)q^{(i)}.$$

□

Corollary 1. *Under the conditions of Proposition 1, the convergence of $\pi^{(i)}(t)$ to its equilibrium $\pi^{(i)\star}$ occurs at a rate determined by the decay factor $a = \frac{\overline{\beta}}{\overline{\beta}+(1-\overline{\beta})q_{tot}}$. Specifically, we have the following bound*

$$\|\pi^{(i)}(t+1) - \pi^{(i)\star}\| \leq \left(\frac{\overline{\beta}}{\overline{\beta} + (1-\overline{\beta})q_{tot}}\right)^{t+1} \|\pi^{(i)}(0) - \pi^{(i)\star}\|.$$

Similarly, we have

$$\|x_v^{(i)}(t+1) - x_v^{(i)\star}\| \leq \beta_v \left(\frac{\overline{\beta}}{\overline{\beta} + (1-\overline{\beta})q_{tot}}\right)^{t+1} \|\pi^{(i)}(0) - \pi^{(i)\star}\|.$$

Proof. In the proof of the Proposition 1, we observed that :

$$\pi^{(i)}(t+1) = a\pi^{(i)}(t) + bq^{(i)},$$

where $a = \frac{\overline{\beta}}{\overline{\beta}+(1-\overline{\beta})q_{\text{tot}}}$ and $b = \frac{1-\overline{\beta}}{\overline{\beta}+(1-\overline{\beta})q_{\text{tot}}}$. We can define the deviation of $\pi^{(i)}(t)$ from its equilibrium point $\pi^{(i)\star} = \frac{q^{(i)}}{q_{\text{tot}}}$ as

$$\Delta\pi^{(i)}(t) = \pi^{(i)}(t) - \pi^{(i)\star}.$$

Substituting $\pi^{(i)}(t) = \Delta\pi^{(i)}(t) + \pi^{(i)\star}$ into the recursive formula, we have

$$\pi^{(i)}(t+1) = a(\Delta\pi^{(i)}(t) + \pi^{(i)\star}) + bq^{(i)}.$$

Recalling that $\pi^{(i)\star} = \frac{q^{(i)}}{q_{\text{tot}}}$, we can also observe that

$$bq^{(i)} = (1-a)\pi^{(i)\star}.$$

Exploiting it in the definition of $\pi^{(i)}(t+1)$

$$\pi^{(i)}(t+1) = a\Delta\pi^{(i)}(t) + a\pi^{(i)\star} + (1-a)\pi^{(i)\star}.$$

Since $a\pi^{(i)\star} + (1-a)\pi^{(i)\star} = \pi^{(i)\star}$, we have

$$\pi^{(i)}(t+1) = a\Delta\pi^{(i)}(t) + \pi^{(i)\star}.$$

Hence

$$\Delta\pi^{(i)}(t+1) = \pi^{(i)}(t+1) - \pi^{(i)\star} = a\Delta\pi^{(i)}(t).$$

Applying this last equation recursively, we can conclude that

$$\Delta\pi^{(i)}(t+1) = a^{t+1}\Delta\pi^{(i)}(0),$$

hence,

$$\|\pi^{(i)}(t+1) - \pi^{(i)\star}\| \leq |a^{t+1}|\|\pi^{(i)}(0) - \pi^{(i)\star}\|.$$

Let us now analyze what happens for the $x_v^{(i)}(t)$. Similarly to what was done for $\pi^{(i)}(t)$, we can define

$$\Delta x_v^{(i)}(t) = x_v^{(i)}(t) - x_v^{(i)\star}.$$

The dynamics of $x_v^{(i)}(t)$ are given by:

$$x_v^{(i)}(t+1) = \beta_v\pi^{(i)}(t) + (1-\beta_v)q^{(i)}.$$

so we can rewrite it by exploiting $\Delta x_v^{(i)}(t+1) = x_v^{(i)}(t+1) - x_v^{(i)\star}$ and $\pi^{(i)}(t) = \Delta\pi^{(i)}(t) + \pi^{(i)\star}$:

$$\Delta x_v^{(i)}(t+1) = \beta_v(\Delta\pi^{(i)}(t) + \pi^{(i)\star}) + (1-\beta_v)q^{(i)} - x_v^{(i)\star}.$$

substituting the equilibrium value $x_v^{(i)\star} = \beta_v\frac{q^{(i)}}{q_{\text{tot}}} + (1-\beta_v)q^{(i)}$, we obtain:

$$\Delta x_v^{(i)}(t+1) = \beta_v(\Delta\pi^{(i)}(t) + \pi^{(i)\star}) + \cancel{(1-\beta_v)q^{(i)}} - \beta_v\frac{q^{(i)}}{q_{\text{tot}}} - \cancel{(1-\beta_v)q^{(i)}}.$$

Given the expression of $\pi^{(i)\star} = \frac{q^{(i)}}{q_{tot}}$, the deviation from equilibrium becomes:

$$\Delta x_v^{(i)}(t+1) = \beta_v \Delta \pi^{(i)}(t).$$

Since $\Delta \pi^{(i)}(t) = a^t \Delta \pi^{(i)}(0)$, where $|a| < 1$, we can express the deviation for $x_v^{(i)}(t)$ as:

$$\Delta x_v^{(i)}(t+1) = \beta_v a^t \Delta \pi^{(i)}(0),$$

and

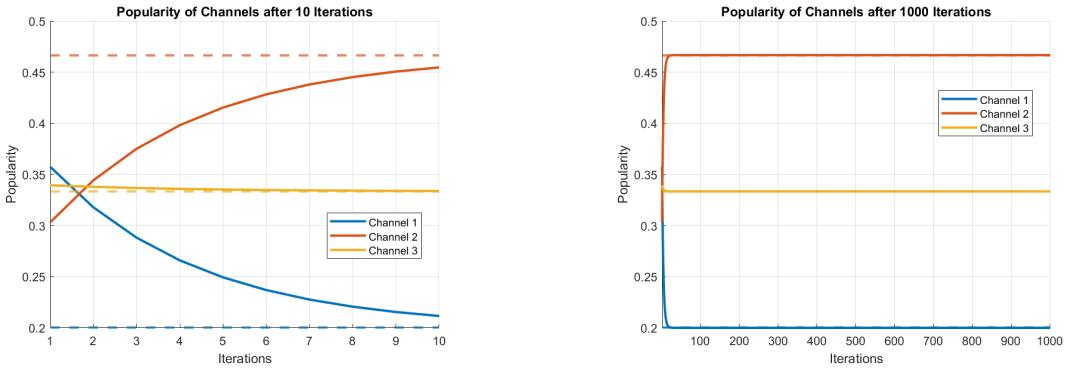
$$\|\Delta x_v^{(i)}(t+1)\| \leq |\beta_v||a^t|\|\Delta \pi^{(i)}(0)\|.$$

□

Numerical simulations confirm the results obtained theoretically, showing that whatever the initial condition on $x_v^{(i)}(0)$ or $\pi^{(i)}(0)$, the system converges to the equilibrium point we derived in Proposition 1. We present a simple example to illustrate the asymptotic behavior.

Example 4.3.1. We consider three channels and we set the parameters β_v $\forall v \in \mathcal{V}$ and the values of $q^{(i)}$, representing the others factors (competence/quality) of channels $\forall i \in \mathcal{I}$, by sampling from a uniform distribution defined in [0,1]. The initial condition of $x_v^{(i)}(0)$ is obtained in the same way $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$: each value $x_v^{(i)}(0)$ is obtained from a uniform distribution defined in [0,1]. The number of users in the social network is equal to 20.

We show in Figure 4.5 and in Figure 4.6 the evolution of popularity and appreciation of the users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and 1000 instants of time (see figure on the right).



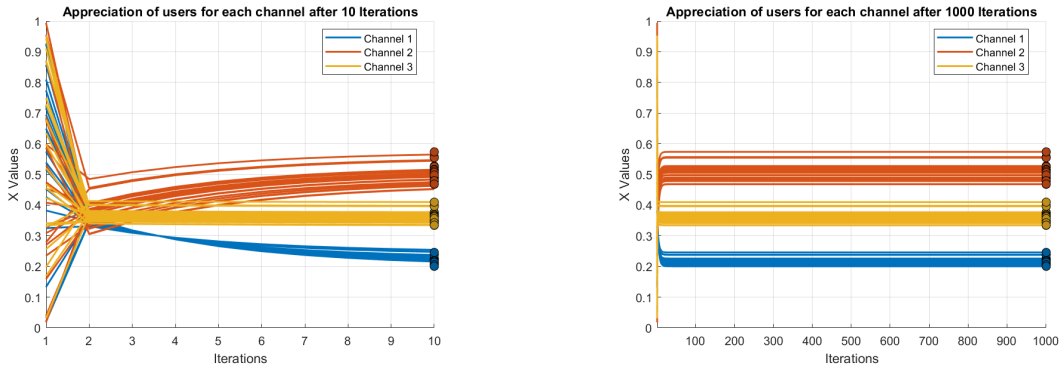
(a) 10 iterations

(b) 1000 iterations

Figure 4.5: Evolution of $\pi^{(i)}(t)$ from (4.2) in the setting of Example 4.3.1. The dotted lines correspond to $\pi^{(i)\star}$.

The dotted lines show the theoretical limit values of the sequence $\{\pi^{(i)}(t)\}_{t\in\mathbb{N}}$, while the colored dots correspond to the limit value of the sequence $\{x_v^{(i)}(t)\}_{t\in\mathbb{N}}$ for all $i \in \mathcal{I}$. The plot of appreciation values is made by evaluating the dynamics of each user's appreciation for each channel.

It can be noticed that the dynamics converge quickly to the theoretical equilibrium points, being the exponential rate of convergence.



(a) 10 iterations

(b) 1000 iterations

Figure 4.6: Evolution of $x_v^{(i)}(t)$ from (4.2) in the setting of Example 4.3.1. The colored dots correspond to $x_v^{(i)\star}$.

4.3.2 French-DeGrooth: Networks effects and consensus dynamics

Let us now consider what happens in the opposite situation, that is when only user interactions contribute to the definition of appreciations (no ranking value and external factors). This setting translates from the perspective of equation (4.1), imposing that $\beta_v = 0, \gamma_v = 0 \ \forall v \in \mathcal{V}$. For simplicity in the calculations, we will assume that $\alpha_v = 1 \ \forall v \in \mathcal{V}$.

Given previous assumptions, the dynamics is reduced to:
$\forall i \in \mathcal{I}, \forall v \in \mathcal{V}$

$$x_v^{(i)}(t+1) = \sum_{w \in \mathcal{V}} P_{vw} x_w^{(i)}(t).$$

In matrix form, starting from an initial condition $x_v^{(i)}(0)$, we have:

$$x_v^{(i)}(t+1) = P x_v^{(i)}(t) = P^{t+1} x_v^{(i)}(0). \tag{4.3}$$

This is the well known *French-De-Groot model* [19].

To analyze the convergence we will use the following Lemma from the theory of the *French de Groot* model:

Lemma 1. *(Theorem 11′ in [42]) The model*

$$x(k+1) = W x(k)$$

converges (i.e. W is regular) to

$$x^\star = \lim_{k \to \infty} x(k) = \lim_{k \to \infty} W^k x(0), \quad \forall \ x(0)$$

if and only if $\lambda = 1$ is the only eigenvalue of W on the unit circle $\{\lambda \in \mathbb{C} : |\lambda| = 1\}$. The model reaches consensus (i.e. W is fully regular) if and only if this eigenvalue is simple, i.e. the corresponding eigenspace is spanned by the vector $\mathbb{1}$.

The following proposition guarantees that the dynamics in (4.3) converges to a limit point.

Proposition 2. *Let $\alpha_v = 1$ for all $v \in \mathcal{V}$ and $\beta_v = \gamma_v = 0$ for all $v \in \mathcal{V}$. Then, for any initial condition $x^{(i)}(0) = x_0^{(i)}$, the dynamics in (4.3) is stable and $\forall i \in \mathcal{I}$ we have*

$$\lim_{t\to+\infty} x^{(i)}(t) = \lim_{t\to+\infty} P^t x^{(i)}(0) = \mathbb{1}_n(\nu^{*T} x^{(i)}(0)).$$

where ν^* is the stationary distribution of the matrix P.

Proof. To use Lemma 1 we first prove that the matrix P satisfies the conditions. Recalling that P is rows-stochastic by definition, from the *Perron Frobenius* theorem (Theorem 1 in Appendix), we know that exists a simple eigenvalue λ_P s.t. $\lambda_P = 1$ while every other eigenvalue λ_i of P is s.t. $|\lambda_i| < 1$.

We can therefore apply the Lemma 1 and conclude that the system is convergent. This condition is equivalent to the condition that P is fully regular, i.e. the limit $P^\infty = \lim_{t\to\infty} P^t$ exists and $P^\infty = \mathbb{1}_n p_\infty^T$ for some $p_\infty \in \mathbb{R}^n$.

Being a rows-stochastic matrix we can say something more: one of the properties of a stochastic matrix is that

$$P^t \to \mathbb{1}_n \nu^{\star T} \quad \text{as} \quad t \to +\infty,$$

where $\nu^\star$ is the stationary distribution defined by the *Perron Frobenius* theorem, namely

$$\nu^\star \geq 0 \text{ s.t. } \mathbb{1}^T \nu^\star = \mathbb{1} \text{ and } P^T \nu^\star = \nu^\star.$$

Combining the last two results, we can characterize the general limit vector p_∞ as $\nu^\star$, and obtain:

$$\lim_{t\to+\infty} x^{(i)}(t) = \lim_{t\to+\infty} P^t x^{(i)}(0) = \mathbb{1}_n(\nu^{\star T} x^{(i)}(0)).$$

□

In other terms, the dynamics lead asymptotically to a consensus: all agents' states converge to the common value $\nu^{\star T} x^{(i)}(0)$, called the *consensus point*, which is a convex combination of the initial states with weights given by the stationary probability components.

Corollary 2. *Under the conditions of the Proposition 2 , the convergence of $x^{(i)}(t)$ to its equilibrium $\lim_{t\to\infty} x^{(i)}(t) = x^{(i)\star}$ occurs at a rate determined by λ_2, the second largest eigenvalue in the absolute value of the matrix P. Specifically, we have the following bound*

$$\|x^{(i)}(t) - x^{(i)\star}\| \leq C\lambda_2^t \|x^{(i)}(0) - x^{(i)\star}\|.$$

Proof. The result comes from the rate of convergence of the P matrix to its stationary distribution $\nu^\star$

$$\|P^t - \nu^\star\| \leq C\lambda_2^t.$$

□

The value of $\pi^{(i)}(t)$ $\forall i \in \mathcal{I}$ at limit will be obtained by normalizing the user consensus values for that specific channel i.

In other words, if we recall the consensus value for all the users for the channel $i \in \mathcal{I}$

$$\mu^{(i)} := \nu^{\star T} x^{(i)}(0),$$

the corresponding $\pi^{(i)\star}$ will be

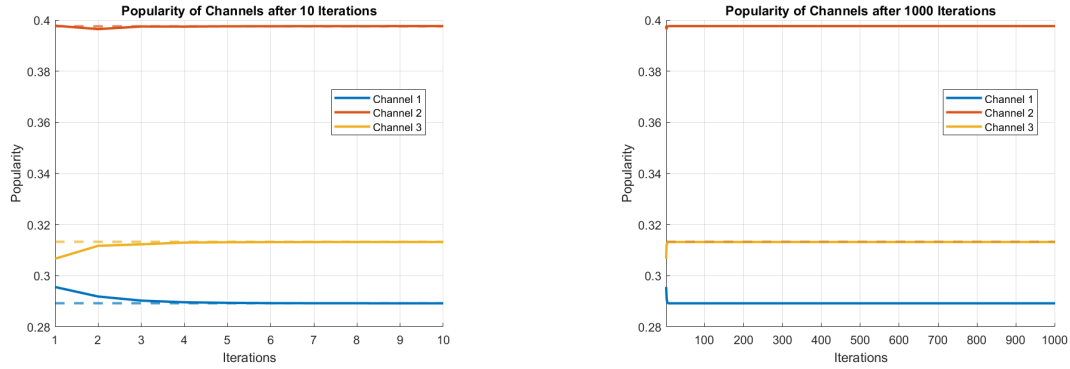
$$\pi^{(i)\star} = \frac{\mathbb{1}_n^T \mathbb{1}_n \mu^{(i)}}{\sum_{i\in\mathcal{I}} \mathbb{1}_n^T \mathbb{1}_n \mu^{(i)}} = \frac{\mu^{(i)}}{\sum_{i\in\mathcal{I}} \mu^{(i)}},$$

as $\mathbb{1}_n^T \mathbb{1}_n = n$.

Numerical simulations confirm the results obtained theoretically, showing that the system converges to the equilibrium point we derived in Proposition 2. We present a simple example to illustrate the asymptotic behavior.

Example 4.3.2. We consider three channels and we set the initial condition of $x_v^{(i)}(0)$ $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$ sampling by a uniform distribution defined in [0,1]. The number of users in the social network is equal to 20. The graph used to simulate the users' interactions is an Erdős–Rényi graph with a probability to have a link of 0.5.

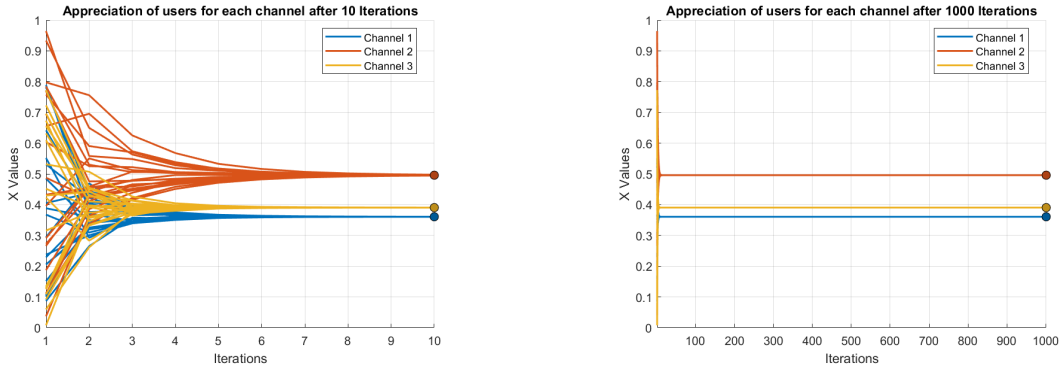
We show in Figure 4.7 and in Figure 4.8 the evolution of popularity and appreciation of the users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and 1000 instants of time (see figure on the right).



(a) 10 iterations

(b) 1000 iterations

Figure 4.7: Evolution of $\pi^{(i)}(t)$ from (4.3) in the setting of Example 4.3.2. The dotted lines correspond to $\pi^{(i)\star}$.



(a) 10 iterations

(b) 1000 iterations

Figure 4.8: Evolution of $x_v^{(i)}(t)$ from (4.3) in the setting of Example 4.3.2. The colored dots correspond to $x_v^{(i)\star}$.

The dotted lines show the theoretical limit values of the sequence $\{\pi^{(i)}(t)\}_{t\in\mathbb{N}}$, while the colored dots correspond to the limit value of the sequence $\{x_v^{(i)}(t)\}_{t\in\mathbb{N}}$ for all $i \in \mathcal{I}$. The plot

of appreciation values is made by evaluating the dynamics of each user's appreciation for each channel.

It can be noticed that the dynamics converge quickly to the theoretical equilibrium points, being the exponential rate of convergence.

4.3.3 Friedkin and Johnsen model: network effects and external factors

Let us now analyze what happens in the case where the appreciation $x_v^{(i)}(t)$ of the generic user v never takes into account the popularity $\pi^{(i)}(t)$ of channel i, but depends only on the latter's other factors (expertise/competence) $q^{(i)}$ and exchange of ideas with other users. This scenario corresponds to imposing $\beta_v = 0 \; \forall v \in \mathcal{V}$ in (4.1).

Noting that $\gamma_v = 1 - \alpha_v$, the system will be:
$\forall i \in \mathcal{I}, \forall v \in \mathcal{V}$

$$\begin{aligned} x_v^{(i)}(t+1) &= \alpha_v \sum_w P_{vw} x_w^{(i)}(t) + (1-\alpha_v) q^{(i)}(t) \\ x^{(i)}(t+1) &= AP x^{(i)}(t) + (I - A) q^{(i)}. \end{aligned} \tag{4.4}$$

where A is the diagonal matrix defined as

$$A = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}.$$

The equations in (4.4) describe an instance of the well-known *Friedkin Jonhsen's model*, [20]. As in the latter, in fact, we recognize *social influence* (through the matrix P), *susceptibility* of the user v to the social influence (through the parameters α_v), as well as the incorporation of initial *prejudices* that here are represented by the other factors' value of the channel i, $q^{(i)}$.

To prove the convergence of the dynamics in (4.4), we need the following Lemma from the theory of *Friedkin Johnsen* model:

Lemma 2. *(Theorem 21 in [43])*
The Friedkin Johnsen model

$$x(k+1) = \Lambda W x(k) + (I - \Lambda) u$$

is asymptotically stable if and only if all agents are directly or indirectly influenced by prejudices. Then, at limit,

$$x^\star = (I - \Lambda W)^{-1}(I - \Lambda).$$

This holds, in particular, if $\Lambda < I_n$ or $\Lambda \neq I_n$ and $\mathcal{G}[W]$ is strongly connected.

We can now prove the main result.

Proposition 3. *Let $\beta_v = 0$ for all $v \in \mathcal{V}$. Assume that in the graph associated to P for all $v \in \mathcal{V}$ there exists a path from node v to the node $w \in \mathcal{V}$ s.t. $\alpha_w < 1$. Then the opinions converge and*

$$x^{(i)\star} = \lim_{t \to \infty} x(t) = (I - AP)^{-1}(I - A) q^{(i)}.$$

Proof. Starting from (4.4) we can rewrite the opinion profile at time $t+1$ as

$$x^{(i)}(t+1) = (AP)^{t+1}x^{(i)}(0) + \sum_{s=0}^{t}(AP)^s(I-A)\mathbf{q}^{(i)}.$$

Taking into account that A is substocastic and P is rows-stochastic, it is easily proved that AP is also substocastic. Given

$$AP = \begin{pmatrix} \alpha_1 P_{11} & \alpha_1 P_{12} & \dots & \alpha_1 P_{1n} \\ \alpha_2 P_{21} & \alpha_2 P_{22} & \dots & \alpha_2 P_{2n} \\ & & \ddots & \\ \alpha_n P_{n1} & \alpha_n P_{n2} & \dots & \alpha_n P_{nn} \end{pmatrix}.$$

For a generic row i

$$\sum_{j\in\mathcal{V}} \alpha_i P_{ij} = \alpha_i \sum_{j\in\mathcal{V}} P_{ij} = \alpha_i \leq 1,$$

given that $\alpha_i \leq 1 \ \forall i \ \in \mathcal{V}$ by definition.

Due to assumptions, all the conditions of the Lemma 2 are satisfied, moreover AP is Schur stable by the Lemma 6 in Appendix. Thus, the dynamics in (4.4) is convergent to $x^{(i)\star}$, that in our case is

$$x^{(i)\star} = (I-AP)^{-1}(I-A)q^{(i)}\mathbb{1}_n, \quad \forall i \in \mathcal{I}.$$

□

The proposition's assumptions imply that there is at least one node where the other factors' contribution is different from 0 and all users are indirectly connected to one (or more) of them.

In addition, since P is stochastic, under the assumptions of the Proposition 3, the matrix $M := (I-AP)^{-1}(I-A)$, called *Total Effect Matrix* in [20], is stochastic too. For this reason, we can rewrite the limit value $x^{(i)\star}$ exploiting the last two observations,

$$x^{(i)\star} = q^{(i)}(I-AP)^{-1}(I-A)\mathbb{1}_n = q^{(i)}\mathbb{1}_n, \quad \forall i \in \mathcal{I}.$$

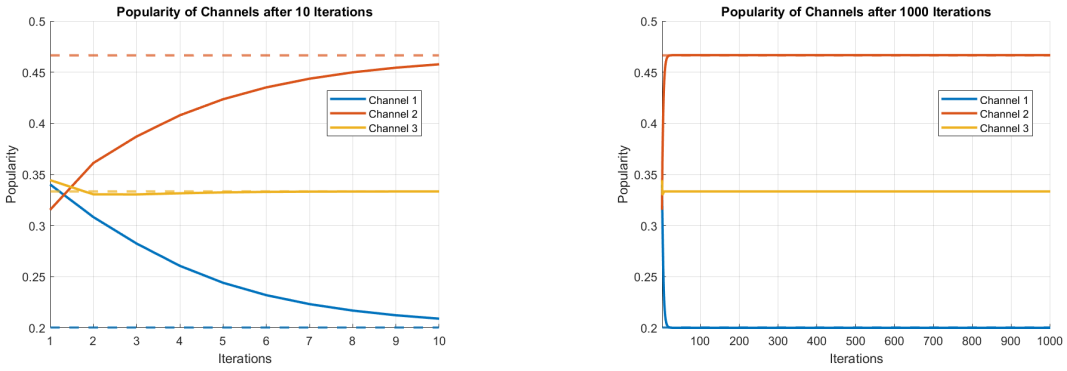
It follows that $\pi^{(i)}(t+1)$ converges to the normalization of the other factors (competence/expertise) of the channel i with respect to the competence of all channels:

$$\pi^{(i)\star} = \frac{q^{(i)}}{\sum_{j\in\mathcal{I}} q^{(j)}}, \quad \forall i \in \mathcal{I}.$$

Numerical simulations confirm the results obtained theoretically, showing that the system converges to the equilibrium point we derived in Proposition 3. We present a simple example to illustrate the asymptotic behavior.

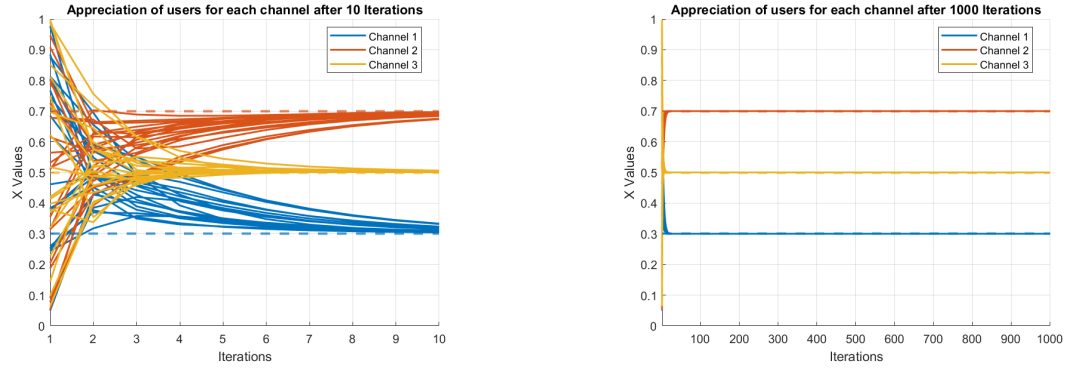
Example 4.3.3. We consider three channels and we set the parameters $\alpha_v \ \forall v \in \mathcal{V}$ and the values of $q^{(i)}$, representing other factors (competence/quality) of channels $\forall i \in \mathcal{I}$, by sampling from a uniform distribution defined in [0,1]. The initial condition of $x_v^{(i)}(0)$ is obtained in the same way $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$: each value $x_v^{(i)}(0)$ is obtained from a uniform distribution defined in [0,1]. The number of users in the social network is equal to 20. The graph used to simulate the users' interactions is an Erdős–Rényi graph with a probability of having a link of 0.5.

We show in Figure 4.9 and in Figure 4.10 the evolution of popularity and appreciation of the users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and 1000 instants of time (see figure on the right).



(a) 10 iterations (b) 1000 iterations

Figure 4.9: Evolution of $\pi^{(i)}(t)$ from (4.4) in the setting of Example 4.3.3. The dotted lines correspond to $\pi^{(i)\star}$.



(a) 10 iterations (b) 1000 iterations

Figure 4.10: Evolution of $x_v^{(i)}(t)$ from (4.4) in the setting of Example 4.3.3. The dotted lines correspond to $x_v^{(i)\star}$.

The dotted lines show the theoretical limit values of the sequence $\{\pi^{(i)}(t)\}_{t\in\mathbb{N}}$, while the colored dots correspond to the limit value of the sequence $\{x_v^{(i)}(t)\}_{t\in\mathbb{N}}$ for all $i \in \mathcal{I}$. The plot of appreciation values is made by evaluating the dynamics of each user's appreciation for each channel.

4.3.4 Network effects and recommendations in absence of external factors

In Sections 4.3.1 and 4.3.3, we have observed that the other factors' value $q^{(i)}$, mainly associated to quality/competence index of each channel, is of great importance in the long-term behavior of appreciation and popularity. It then becomes interesting to understand what the elimination of this contribution from the equation leads to in the dynamics. For this reason, we now consider the case of $\gamma_v = 0\ \forall v \in \mathcal{V}$.

Since $\beta_v = 1 - \alpha_v$, the dynamics in (4.1) becomes:
$\forall i \in \mathcal{I}, \forall v \in \mathcal{V}$

$$x_v^{(i)}(t+1) = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} x_w^{(i)}(t) + (1 - \alpha_v)\pi^{(i)}(t)$$

and in matrix form

$$x^{(i)}(t+1) = APx^{(i)}(t) + (I - A)\pi^{(i)}(t), \tag{4.5}$$

where A is the diagonal matrix defined as

$$A = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix}.$$

We have the following proposition:

Proposition 4. *Let $\gamma_v = 0$ for all $v \in \mathcal{V}$ and define $z(t) = \sum_{i \in \mathcal{I}} x^{(i)}(t)$. Assume that in the graph, associated to P $\forall v \in \mathcal{V}$, there exists a path from v to w with $\alpha_w < 1$. Then $z(t)$ converges and*

$$\mathbb{1}_n^T z^\star = \lim_{t \to \infty} \mathbb{1}_n^T z(t) = n.$$

Proof. The definition of $z(t+1)$ is:

$$z(t+1) := \sum_{i \in \mathcal{I}} x^{(i)}(t+1). \tag{4.6}$$

From equation (4.5), recalling that $\sum_{i \in \mathcal{I}} \pi^{(i)}(t) = 1$ $\forall t$ we obtain,

$$z(t+1) = APz(t) + (I - A)\mathbb{1}_n.$$

It should be noticed that this is a specific instance of the *Friedkin Johnsen model* [20] where the bias in the constant input is given by vector $\mathbb{1}_n$. Exploiting Lemma (2) in section 4.3.3 we know that

$$z^\star = (I - AP)^{-1}(I - A)\mathbb{1}_n.$$

As already noted in the section 4.3.3 in the proof of Proposition (3), the matrix $M := (I - AP)^{-1}(I - A)$ is stochastic, so the limit value for $z(t+1)$ is simply $z^\star = \mathbb{1}_n$. □

Recalling the definition of $\pi^{(i)}(t)$, we can rewrite it in terms of $z(t)$ as

$$\pi^{(i)}(t) = \frac{\sum_{v \in \mathcal{V}} x_v^{(i)}(t)}{\sum_{v \in \mathcal{V}} \sum_{i \in \mathcal{I}} x_v^{(i)}(t)} = \frac{\mathbb{1}_n^T x^{(i)}(t)}{\mathbb{1}_n^T z(t)}. \tag{4.7}$$

It is now advantageous to treat jointly the dynamics of $x^{(i)}(t)$ and that of $\pi^{(i)}(t)$ by treating them as if they were independent, although they are not. For this purpose, we consider the two dynamics, where the second one is obtained from the condition (4.7),

$$x^{(i)}(t+1) = APx^{(i)}(t) + (I - A)\pi^{(i)}(t)\mathbb{1}_n,$$

$$\pi^{(i)}(t+1) = \frac{\mathbb{1}_n^T x^{(i)}(t+1)}{\mathbb{1}_n^T z(t+1)} = \frac{\mathbb{1}_n^T}{\mathbb{1}_n^T z(t+1)}(APx^{(i)}(t) + (I - A)\mathbb{1}_n \pi^{(i)}(t)),$$

and define

$$s^{(i)}(t) := \begin{pmatrix} x^{(i)}(t) \\ \pi^{(i)}(t) \end{pmatrix}. \tag{4.8}$$

The joint dynamics are

$$s^{(i)}(t+1) = U(t)s^{(i)}(t), \tag{4.9}$$

with

$$U(t) = \begin{bmatrix} AP & (I-A)\mathbb{1}_n \\ \frac{\mathbb{1}^T AP}{\mathbb{1}_n^T z(t+1)} & \frac{\mathbb{1}_n^T (I-A)\mathbb{1}_n}{\mathbb{1}_n^T z(t+1)} \end{bmatrix}.$$

The time-dependent matrix $U(t)$ can be rewritten by splitting the contribution into two terms, where the former is time-invariant and the latter is time-varying. The decomposition is $U(t) = \tilde{U} + \Delta U(t)$ with

$$\tilde{U} = \begin{bmatrix} AP & (I-A)\mathbb{1}_n \\ \frac{\mathbb{1}_n^T AP}{n} & \frac{\mathbb{1}_n^T (I-A)\mathbb{1}_n}{n} \end{bmatrix}, \ \Delta U(t) = \left(\frac{1}{\mathbb{1}_n^T z(t+1)} - \frac{1}{n} \right) \begin{bmatrix} 0_n & 0 \\ \mathbb{1}_n^T AP & \mathbb{1}_n^T (I-A)\mathbb{1}_n \end{bmatrix}. \tag{4.10}$$

It is important to note that the time invariant matrix $\tilde{U}$ is rows-stochastic, since $\tilde{U}\mathbb{1}_{n+1} = \mathbb{1}_{n+1}$ and $\tilde{U}_{ij} \geq 0 \ \forall i, \ \forall j$.

To analyze the behavior of the system for $t \to \infty$ we need the following two lemmas:

Lemma 3 (Theorem 2.1 in [44]). *Given $P(t)$ defined as*

$$P(t) := (Q(t) + \Delta Q(t))(Q(t-1) + \Delta Q(t-1)) \dots (Q(1) + \Delta Q(1)) \tag{4.11}$$

where $\|Q(t)\| = 1 \ \forall t$ (with $\|\cdot\|$ any consistent norm on $M_n(\mathbb{C})$, i.e. $\|AB\| < \|A\| \cdot \|B\|$), we have that

$$\sum_t \|\Delta Q(t)\| \text{ converges} \implies P(t) \text{ converges for } t \to \infty$$

if and only if

$$\forall r, \text{ the product } \prod_{\ell=r}^{t} Q(\ell) \text{ converges for } t \to \infty.$$

.

Lemma 4 (Theorem 4.1 in [45]). *Let $\tilde{x}$ the right eigenvector associated to $\tilde{\lambda}_1 = 1$ of the matrix $Q(t)$. If*

$$\sum_{t=1}^{\infty} \|\Delta Q(t)\| < \infty,$$

then for some vector ϕ we have

$$\lim_{t\to\infty} P(t) = \tilde{x}\phi^T.$$

The root convergence index is not greater than

$$\max\{\rho, \sigma\},$$

where ρ is the largest of magnitudes of the subdominant eigenvalues of $Q(t)$ and

$$\sigma = \limsup \|\Delta Q(t)\|_2^{1/t}.$$

The behavior of the system in (4.9) is described by the following proposition:

Proposition 5. *Given the initial condition* $s(0)$ *,* $\forall i \in \mathcal{I}$*,*

$$s^{(i)}(t+1) = U(t)s^{(i)}(t)$$

converges to

$$s^{(i)\star} = (\mathbb{1}_n \phi^T)s^{(i)}(0) \quad \textit{for } t \to \infty,$$

for some non negative $\phi \in \mathbb{R}^{n+1}$*.*

Proof. In the proof passages, we will always use $\|\cdot\|_2$.

Starting from the system in (4.9), explicating recurrence, we quickly come up against a product of time-dependent matrices

$$s^{(i)}(t+1) = U(t)U(t-1)s^{(i)}(t-1) = U(t)U(t-1)\dots U(1)s^{(i)}(0) = P(t)s^{(i)}(0), \ \forall i \in \mathcal{I}, \quad (4.12)$$

where $P(t) := U(t)U(t-1)\dots U(1)$ is the same for all channels i.
After exploiting the decomposition of $U(t)$, introduced in (4.10), for which

$$P(t) = (\tilde{U} + \Delta U(t))(\tilde{U} + \Delta U(t-1))(\tilde{U} + \Delta U(t-2))\dots(\tilde{U} + \Delta U(1)),$$

we can use Lemma 3.

In our case $Q(t) = \tilde{U}\ \forall t \in \mathbb{N}\cup\{0\}$, that is a stochastic matrix as observed; it is therefore trivial to prove that $\|\tilde{U}\|_2 = 1$. Exploiting the *Perron Frobenius* theorem (theorem (1) in Appendix), we also know that

$$P(t,r) = \prod_{l=r}^{t} \tilde{U}^l = \tilde{U}^{t-r+1} \xrightarrow{t\to\infty} \mathbb{1}_n \nu^{\star T},$$

where $\nu^\star$ is the stationary distribution of $\tilde{U}$, i.e.

$$\nu^* \geq 0 \text{ s.t. } \mathbb{1}^T\nu^* = \mathbb{1} \text{ and } \tilde{U}^T\nu^* = \nu^*, \quad (4.13)$$

so the first condition of Lemma 3 is satisfied.

We now check the convergence of $\sum_t \|\Delta U(t)\|$. It can be noticed that

$$\|\Delta U(t)\|_2 \leq \left|\frac{1}{\mathbb{1}_n^T z(t+1)} - \frac{1}{n}\right| \|B\|_2$$

with

$$B = \begin{bmatrix} 0_n & 0 \\ \mathbb{1}_n^T AP & \mathbb{1}_n^T(I-A)\mathbb{1}_n \end{bmatrix}.$$

The matrix B is defined as a null matrix, except for the last row. For this reason, the norm of the matrix coincides with that of the one non-zero vector. Having defined $v = \left(\mathbb{1}_n^T AP, \mathbb{1}_n^T(I-A)\mathbb{1}_n\right)$, we can exploit the inequality between norms, $\|v\|_2 \leq \|v\|_1$ and observe that the $\|v\|_1 = n$. We conclude that $\|v\|_2 \leq n$. As shown in Proposition 4 we have:

$$\mathbb{1}_n^T z(t+1) \to \mathbb{1}_n^T \mathbb{1}_n = n, \ t \to \infty$$

and that

$$\left|\frac{1}{\mathbb{1}_n^T z(t+1)} - \frac{1}{n}\right| \leq c\lambda_1^t,$$

where λ_1 is the maximum eigenvalue of AP that is substochastic (as shown in the proof of Prop. 3), so $\lambda_1 < 1$. Having limited $\|B\|$ with a constant value equal to the size of the graph (as it only depends on the vector v with $\|v\|_2 \leq n$), and having shown that the absolute value of the scalar exhibits an exponential decrease, we can conclude that

$$\sum_i \|\Delta U(t))\| \; converges \implies P(t) \text{ converges for } t \to \infty.$$

So far we have only shown convergence, but we still do not know what $P(t)$ converges to. For this last aspect, we can use Lemma 4.

Since the matrix $\tilde{U}$ is stochastic in our case study, $\tilde{x}$ will be a vector with all components equal to 1. Moreover, we have already proved the convergence of $\|\Delta U(t)\|_2$, so we can use the result of the Lemma 4.

Therefore, we can conclude that the system 4.9, namely

$$s^{(i)}(t+1) = U(t)s^{(i)}(t) = P(t)s^{(i)}(0),$$

for $t \to \infty$ converges to

$$s^{(i)\star} = (\mathbb{1}_n \phi^T)s^{(i)}(0) \; \forall i \in \mathcal{I},$$

where $\mathbb{1}_n \phi^T$ is a matrix with all rows equal to ϕ. □

It can be seen that the dynamic for each channel i converges to a consensus, since the role of each user is the same: in other words, all users converge to the same value of appreciation for the i-th channel, which is the same value as the latter's popularity. Although ϕ is the same $\forall i \in \mathcal{I}$, the consensus value is different depending on the channel i, since it is linked to the initial condition $s^{(i)}(0)$.

Numerical simulations confirm the results obtained theoretically and allow for further characterization of the values obtained at limit. We present a simple example to illustrate the asymptotic behavior.

Example 4.3.4. We consider three channels and we set the parameters $\alpha_v \; \forall v \in \mathcal{V}$ by sampling from a uniform distribution defined in [0,1]. The initial condition of $x_v^{(i)}(0)$ is obtained in the same way $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$: each value $x_v^{(i)}(0)$ is obtained from a uniform distribution defined in [0,1]. The number of users in the social network is equal to 20.

We show in Figure 4.11 and in Figure 4.12 the evolution of popularity and appreciation of the users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and 1000 instants of time (see figure on the right), when the users' interactions are simulated using a complete graph. In Figures 4.13 and 4.14 the same is done when the graph that represents users' interactions is and Erdős–Rényi graph with the probability to have an arc equal to 0.2.

It can be observed that for any channel i the appreciation of all users $x_v^{(i)\star} \; \forall v \in \mathcal{V}$ goes to a consensus, as observed in theory. This value also corresponds to the popularity value of that specific channel i, $\pi^{(i)\star}$. This again confirms the theory, since consensus is obtained on the joint system, described in the equation (4.9). Analysis of behavior in simulation allows this consensus to be further defined.

From the Theorem 5 we know that

$\forall i \in \mathcal{I}$

$$s^{(i)\star} = (\mathbb{1}_n \phi^T)s^{(i)}(0), \quad \text{for some non negative } \phi \in \mathbb{R}^{n+1}.$$

Simulations allow us to observe that, barring an error that will be discussed later, this vector ϕ is nothing but the stationary distribution of the time-invariant matrix $\tilde{U}$, called $\nu^\star$ and already

defined in (4.13). This suggests to us that over long periods of time the model described in (4.9) is approximated by the model

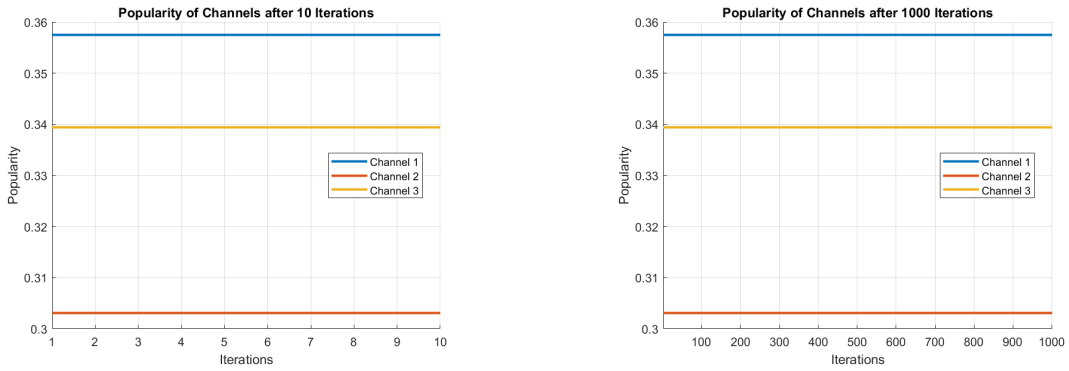
$\forall i \in \mathcal{I}$,

$$\hat{s}^{(i)}(t+1) = \tilde{U}\hat{s}^{(i)}(t), \tag{4.14}$$

$$\hat{s}(t) := \begin{pmatrix} \hat{x}^{(i)}(t) \\ \hat{\pi}^{(i)}(t) \end{pmatrix}.$$

since the term $\Delta U(t)$ very quickly becomes smaller and smaller and $U(t) \xrightarrow{t\to\infty} \tilde{U}$ (we obtained exponential speed in the proof of Proposition 5).

The error involved in this approximation depends on the topology of the graph used to simulate the interactions between users: while for graphs with high levels of connection between users (as many arcs per number of nodes), such as the complete graph, the error is very small, in the case of graphs with fewer arcs the errors accumulate more and result in a larger deviation between the model in (4.9) and that in (4.14).



(a) 10 iterations

(b) 1000 iterations

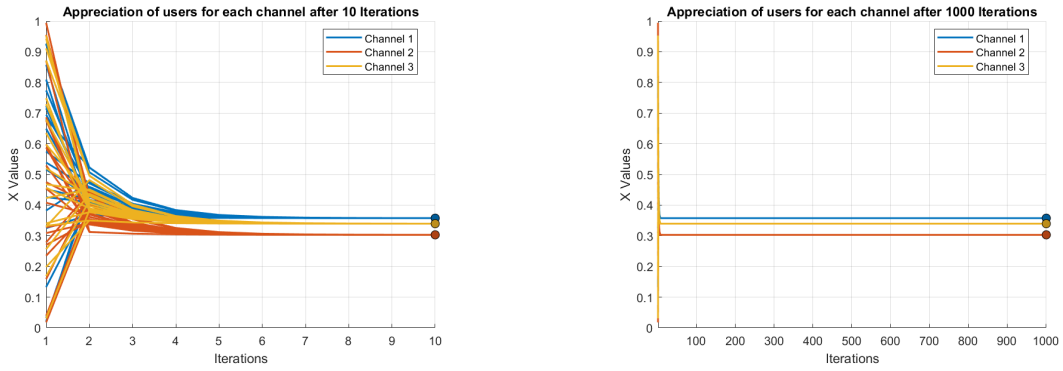
Figure 4.11: Evolution of $\pi^{(i)}(t)$ from (4.5) on a complete graph, in the setting of Example (4.3.4). The dotted lines correspond to $\hat{\pi}^{(i)\star}$ from (4.14).

The dotted lines in case of $\pi^{(i)}(t)$ and the colored dots in case of $x_v^{(i)}(t)$ in the plots represent the limit values $\hat{s}^\star$ of the model described in (4.14).

Observation 1. In the case of the complete graph the errors between the value obtained by iterating the dynamics and the theoretical ones associated with the model (4.14), settle around the order of 10^{-16}, hence the two models are equivalent. In the case of the Erdős–Rényi graph the error rises to the order of 10^{-4}, confirming the relationship between topology (degree of connectivity) of the graph and deviation between the two models described in (4.9) and (4.14).

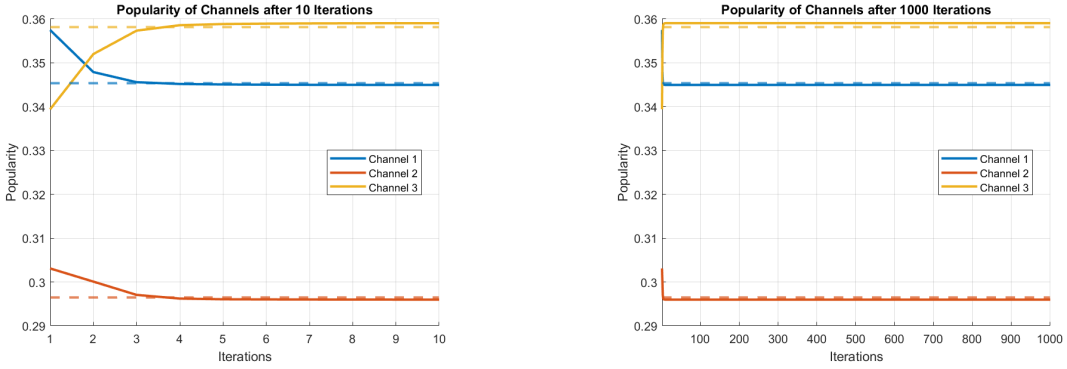
4.3.5 Asymptotic behavior with all effects

So far we have analyzed what happens by canceling one or more terms in the equation (4.1), or, in other words, nullifying some dependence, whether it is on the platform, on the network or on other factors' contribution. It's time to return to the general case of (4.1), to understand how these terms interact with each other. The dynamics are described by:



(a) 10 iterations (b) 1000 iterations

Figure 4.12: Evolution of $x_v^{(i)}(t)$ from (4.5) on a complete graph, in the setting of Example (4.3.4). The colored dots correspond to $\hat{x}_v^{(i)\star}$ from (4.14).



(a) 10 iterations (b) 1000 iterations

Figure 4.13: Evolution of $\pi^{(i)}(t)$ from (4.5) on an Erdős–Rényi graph, in the setting of Example (4.3.4). The dotted lines correspond to $\hat{\pi}^{(i)\star}$ from (4.14).

$\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$,

$$x_v^{(i)}(t+1) = \alpha_v \left(\sum_w P_{vw} x_w^{(i)}(t) \right) + \beta_v \pi^{(i)}(t) + \gamma_v q^{(i)},$$

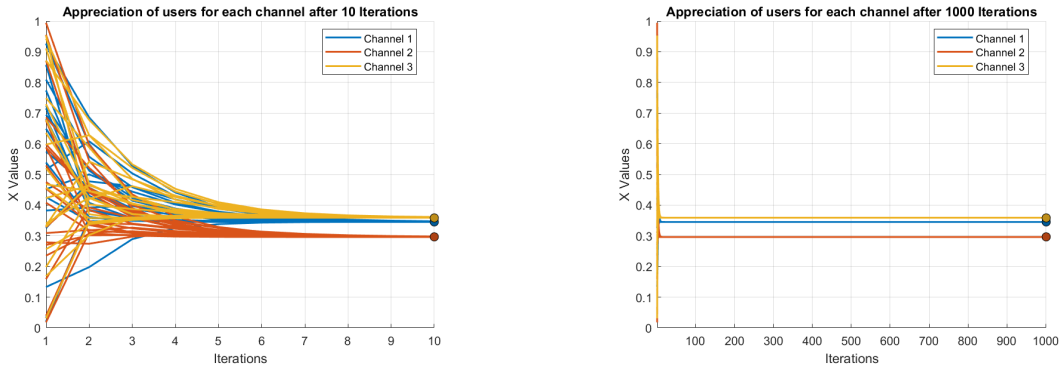
that we can rewrite in matrix form as

$\forall i \in \mathcal{I}$,

$$x^{(i)}(t+1) = APx^{(i)}(t) + B\pi^{(i)}(t) + (I - A - B)q^{(i)}, \tag{4.15}$$

where A and B are diagonal matrices, respectively

$$A = \begin{pmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \beta_1 & & & \\ & \beta_2 & & \\ & & \ddots & \\ & & & \beta_n \end{pmatrix}.$$



(a) 10 iterations

(b) 1000 iterations

Figure 4.14: Evolution of $x_v^{(i)}(t)$ from (4.5) on an Erdős–Rényi, in the setting of Example 4.3.4. The colored dots correspond to $\hat{x}_v^{(i)\star}$ from (4.14).

We have the following proposition:

Proposition 6. *Given the model described in (4.1), let's define* $z(t) = \sum_{i\in\mathcal{I}} x^{(i)}(t)$. *Assume that in the graph associated to* P *for all* $v \in \mathcal{V}$ *there exists a path from* v *to* w *such that* $\gamma_w > 0$. *Then* $z(t)$ *converges and*

$$\mathbb{1}_n^T z^* = \lim_{t\to\infty} \mathbb{1}_n^T z(t) = \mathbb{1}_n^T (I - AP)^{-1}(B\mathbb{1}_n + q_{tot}(I - A - B))\mathbb{1}_n.$$

Proof. The joint contribution in the equation of the α_v- and β_v-dependent terms means that the model has similarities with the one just discussed in Section 4.3.4 (case $\gamma_v = 0\ \forall v \in \mathcal{V}$), so the mathematical formulation will be partly similar to it, with the appropriate complications given by the presence of γ_v.

From equation (4.15), recalling that $\sum_{i\in\mathcal{I}} \pi^{(i)}(t) = 1$ and defining $q_{tot} := \sum_{i\in\mathcal{I}} q^{(i)}$, we obtain

$$z(t+1) = \sum_{i\in\mathcal{I}} x^{(i)}(t+1) = AP \sum_{i\in\mathcal{I}} x_w^{(i)}(t) + B\mathbb{1}_n + (I - A - B)q_{tot}\mathbb{1}_n \tag{4.16}$$

$$z(t+1) = APz(t) + B\mathbb{1}_n + q_{tot}(I - A - B)\mathbb{1}_n.$$

By explicating recurrence

$$z(t+1) = (AP)^{t+1} z(0) + \sum_{k=0}^{t} (AP)^k \left(B\mathbb{1}_n + q_{\text{tot}}(I - A - B)\mathbb{1}_n\right).$$

Imposing that $\forall v \in \mathcal{V}$ exists a path from v to w with $\gamma_w > 0$ guarantees that AP is substochastic matrix (generalization of the proof of Proposition 3 in section 4.3.3). Consequently, stability is maintained and it is observed that at limit the model converges to

$$z^\star = \lim_{t\to\infty} z(t) = (I - AP)^{-1}(B\mathbb{1}_n + q_{tot}(I - A - B))\mathbb{1}_n.$$

□

Recalling the definition in (4.7) of $\pi^{(i)}(t)$ in terms of $z(t)$, we can write the dynamics of $x^{(i)}(t)$ and $\pi^{(i)}(t)$ as

$$x^{(i)}(t+1) = APx^{(i)}(t) + B\pi^{(i)}(t)\mathbb{1}_n + (I - A - B)q^{(i)}\mathbb{1}_n,$$

$$\pi^{(i)}(t+1) = \frac{\mathbb{1}_n^T x^{(i)}(t+1)}{\mathbb{1}_n^T z(t+1)} = \frac{\mathbb{1}_n^T}{\mathbb{1}_n^T z(t+1)}(APx^{(i)}(t) + B\pi^{(i)}(t)\mathbb{1}_n + (I - A - B)q^{(i)}\mathbb{1}_n).$$

As in section 4.3.4, we define

$$s^{(i)}(t) := \begin{pmatrix} x^{(i)}(t) \\ \pi^{(i)}(t) \end{pmatrix}$$

and express the joint dynamics

$\forall i \in \mathcal{I}$,

$$s^{(i)}(t+1) = U(t)s^{(i)}(t) + q^{(i)}c(t), \tag{4.17}$$

where

$$U(t) = \begin{bmatrix} AP & B\mathbb{1}_n \\ \frac{\mathbb{1}_n^T AP}{\mathbb{1}_n^T z(t+1)} & \frac{\mathbb{1}_n^T B\mathbb{1}_n}{\mathbb{1}_n^T z(t+1)} \end{bmatrix}, \quad c(t) = \begin{pmatrix} (I - A - B)\mathbb{1}_n \\ \frac{\mathbb{1}_n^T (I-A-B)\mathbb{1}_n}{\mathbb{1}_n^T z(t+1)} \end{pmatrix}. \tag{4.18}$$

Lemma 5. *If $q_{tot} \geq 1$ and $z_v(0) = \sum_{i\in\mathcal{I}} x_v^{(i)}(0) \geq 1$, $\forall v \in \mathcal{V}$, then the time-variant matrix $U(t)$ is substochastic $\forall t \in \mathbb{N}$.*

Proof. From assumptions of Proposition 6, we have that for all $v \in \mathcal{V}$ there exists a path from $v \in \mathcal{V}$ to $w \in \mathcal{V}$ such that $\gamma_w > 0$. To this, follows that the condition of substochasticity is satisfied for the first n rows of $U(t)$. Thus, focusing on the last row of the matrix $U(t)$, the condition to be satisfied for substocasticity at each time step t is

$$\frac{\mathbb{1}_n^T AP}{\mathbb{1}_n^T z(t)} + \frac{\mathbb{1}_n^T B\mathbb{1}_n}{\mathbb{1}_n^T z(t)} < 1,$$

that corresponds, being $\mathbb{1}_n^\top z(t) > 0$ $\forall t$, to

$$\sum_i \alpha_i + \sum_i \beta_i < \mathbb{1}_n^T z(t), \ \forall t. \tag{4.19}$$

We now prove that $z(t) \geq \mathbb{1}_n$ by induction on $t \in \mathbb{N}_0$. By hypothesis $z_v(0) = \sum_{i\in\mathcal{I}} x_v^{(i)}(0) \geq 1$ $\forall v \in \mathcal{V}$ and $q_{tot} \geq 1$. Now, we have

$$\begin{aligned} z(t+1) &= APz(t) + B\mathbb{1}_n + q_{\text{tot}}(I - A - B)\mathbb{1}_n \\ &> AP\mathbb{1}_n + B\mathbb{1}_n + (I - A - B)\mathbb{1}_n \\ &\geq A\mathbb{1}_n + B\mathbb{1}_n + (I - A - B)\mathbb{1}_n = \mathbb{1}_n. \end{aligned}$$

where the last inequality follows from the stochasticity of the matrix P.

Recalling that $\alpha_v + \beta_v + \gamma_v = 1$, we have $\sum_i \alpha_i + \sum_i \beta_i \leq n < \mathbb{1}_n^T z(t)$, $\forall t$. □

Proposition 7. *The dynamics in (4.17),*

$$s^{(i)}(t+1) = U(t)s^{(i)}(t) + q^{(i)}c(t), \ \forall i \in \mathcal{I}$$

is convergent for $t \to \infty$.

Proof. We can express $U(t) = \tilde{U} + \Delta U(t)$ with

$$\tilde{U} = \begin{bmatrix} AP & B\mathbb{1}_n \\ \frac{\mathbb{1}_n^T AP}{\mathbb{1}_n^T z^*} & \frac{\mathbb{1}_n^T B\mathbb{1}_n}{\mathbb{1}_n^T z^*} \end{bmatrix} \quad \Delta U(t) = \left(\frac{1}{\mathbb{1}_n^T z(t+1)} - \frac{1}{\mathbb{1}_n^T z^*}\right) \begin{bmatrix} 0_n & 0 \\ \mathbb{1}_n^T AP & \mathbb{1}_n^T B\mathbb{1}_n \end{bmatrix},$$

and $c(t) = \tilde{c} + \Delta c(t)$ with

$$\tilde{c} = \begin{pmatrix} (I - A - B)\mathbb{1}_n \\ \frac{\mathbb{1}_n^T (I-A-B)\mathbb{1}_n}{\mathbb{1}_n^T z^*} \end{pmatrix} \quad \Delta c(t) = \begin{pmatrix} 0 \\ \mathbb{1}_n^T (I - A - B)\mathbb{1}_n (\frac{1}{\mathbb{1}_n^T z(t+1)} - \frac{1}{\mathbb{1}_n^T z^*}) \end{pmatrix}.$$

Explicating recurrence, the dynamics in (4.17) becomes

$$s^{(i)}(t+1) = \left(\prod_{k=0}^{t} U(k)\right) s^{(i)}(0) + q^{(i)} \sum_{k=0}^{t} \left(\prod_{j=k+1}^{t} U(j)\right) c(k). \tag{4.20}$$

Let's start analyzing

$$\left(\prod_{k=1}^{t} U(k)\right) s^{(i)}(0),$$

which represents the cumulative effect of the dynamics over time, involving the product of the matrices $U(k)$. Since each matrix $U(k)$ is substochastic according to Lemma (5), we expect that the product of these matrices will not grow unbounded.

We know that the time dependence of the sequence $\{U(t)\}_{t>0}$ is only linked to the term $\Delta U(t)$. In turn, the term $\Delta U(t)$ depends on $\left(\frac{1}{\mathbb{1}_n^T z(t+1)} - \frac{1}{\mathbb{1}_n^T z^\star}\right)$. From Proposition 4, that describes the evolution of $\{z(t)\}_t$, follows that the convergence of $z(t)$ to $z^\star$ is exponential, so

$$\left| \frac{1}{\mathbb{1}_n^T z(t+1)} - \frac{1}{\mathbb{1}_n^T z^\star} \right| \leq c\lambda_1^t \tag{4.21}$$

where λ_1 is the maximum eigenvalue of AP that is substochastic (as shown in the proof of Prop. (3)), so $\lambda_1 < 1$.

We can conclude that $\{U(t)\}_{t>0}$ converges to the matrix $\tilde{U}$ as $t \to \infty$ (since $\Delta U(t) \to 0$, when $t \to \infty$, exponentially), which is still a substochastic matrix (the eigenvalues of $\tilde{U}$ are strictly less than 1 in absolute value). Then the product $\prod_{k=1}^{t} U(k)$ will tend to zero. In fact, using the submultiplicative property of norms,

$$\| \prod_{k=1}^{t} U(k)\|_2 \leq \prod_{k=1}^{t} \|U(k)\|_2 \leq \prod_{k=1}^{t} (\|\tilde{U}\|_2 + \|\Delta U(k)\|_2).$$

For sufficiently large t, $\|\Delta U(k)\|_2$ becomes arbitrarily small. Let $\epsilon_k = \|\Delta U(k)\|_2$. Then:

$$\| \prod_{k=1}^{t} U(k)\|_2 \leq \prod_{k=1}^{t} (\|\tilde{U}\|_2 + \epsilon_k).$$

Define $\eta = \|\tilde{U}\|_2 < 1$. For large t, ϵ_k is small enough to satisfy $\|\tilde{U}\|_2 + \epsilon_k < 1$. This ensures:

$$\prod_{k=1}^{t} (\|\tilde{U}\|_2 + \epsilon_k) \to 0 \quad \text{as} \quad t \to \infty.$$

Hence:

$$\| \prod_{k=1}^{t} U(k)\|_2 \to 0 \quad \text{as} \quad t \to \infty.$$

Taking advantage of the definition of $c(t)$, the second term of the equation (4.20) can instead be divided into two parts,

$$q^{(i)} \sum_{k=0}^{t} \left(\prod_{j=k+1}^{t} U(j) \right) (\tilde{c} + \Delta c(k)) = q^{(i)} \sum_{k=0}^{t} \left(\prod_{j=k+1}^{t} U(j) \right) \tilde{c} + q^{(i)} \sum_{k=0}^{t} \left(\prod_{j=k+1}^{t} U(j) \right) \Delta c(k)$$

For the constant component $\tilde{c}$:

$$\sum_{k=0}^{t} \left(\prod_{j=k+1}^{t} U(j) \right) \tilde{c}.$$

Since $\tilde{c}$ is constant, each term satisfies:

$$\left\| \left(\prod_{j=k+1}^{t} U(j) \right) \tilde{c} \right\|_2 \leq \left\| \prod_{j=k+1}^{t} U(j) \right\|_2 \|\tilde{c}\|_2 .$$

We have already proved that $\left\| \prod_{j=k+1}^{t} U(j) \right\|_2 \to 0$ as $t \to \infty$ and that each term $\Delta U(j)$ tends to 0 exponentially fast since (4.21). From the latter, we can conclude that also the convergence of $\left\| \prod_{j=k+1}^{t} U(j) \right\|_2$ is exponential and it ensures that the series:

$$\sum_{k=0}^{\infty} \left\| \prod_{j=k+1}^{t} U(j) \right\|_2 \|\tilde{c}\|_2$$

converges.

As far as it concerns the time-dependent component $\Delta c(k)$:

$$\sum_{k=0}^{t} \left(\prod_{j=k+1}^{t} U(j) \right) \Delta c(k).$$

we can see that each term satisfies:

$$\left\| \left(\prod_{j=k+1}^{t} U(j) \right) \Delta c(k) \right\|_2 \leq \left\| \prod_{j=k+1}^{t} U(j) \right\|_2 \|\Delta c(k)\|_2 .$$

The term $\Delta c(k)$ depends again on $\frac{1}{\mathbb{1}_n^T z(k+1)} - \frac{1}{\mathbb{1}_n^T z^\star}$, which vanishes as $k \to \infty$ exponentially fast (as shown in (4.21)). For this reason, the product $\left\| \prod_{j=k+1}^{t} U(j) \right\|_2 \|\Delta c(k)\|_2$ is composed by two terms that go to 0 exponentially fast.

We can conclude that the infinite series:

$$\sum_{k=0}^{t} \prod_{j=k+1}^{t} U(j) \Delta c(k)$$

is a summable series and converges for $t \to \infty$, since both factors decay sufficiently rapidly.

Given the convergence of all terms, we can conclude that the dynamics in (4.17),

$$s^{(i)}(t+1) = U(t)s^{(i)}(t) + q^{(i)}c(t), \ \forall i \in \mathcal{I}$$

is convergent for $t \to \infty$. □

Example 4.3.5. We consider three channels and we set the parameters α_v, β_v and γ_v $\forall v \in \mathcal{V}$ by sampling each of them from a uniform distribution defined in [0,1] and then normalizing to have $\alpha_v + \beta_v + \gamma_v = 1$. The initial condition of $x_v^{(i)}(0)$ is obtained in the same way $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$: each value $x_v^{(i)}(0)$ is obtained from a uniform distribution defined in [0,1]. Again, the $q^{(i)}$ value for each channel is obtained by sampling from the same distribution. The number of users in the social network is equal to 20.

We show in Figure 4.15 and in Figure 4.16 the evolution of popularity and appreciation of the users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and 1000 instants of time (see figure on the right), when the users' interactions are simulated using a complete graph. In Figures 4.17 and 4.18 the same is done when the graph that represents users' interactions is and Erdős–Rényi graph with the probability to have an arc equal to 0.2.

In this case, the dotted lines in the plots of $\pi^{(i)}(t)$ (Figures 4.15 and 4.17) and the colored dots in the plots of $x_v^{(i)}(t)$ (Figures 4.16 and 4.18), represent the limit values of $\hat{\pi}^{(i)}(t)$ and $\hat{x}_v^{(i)}(t)$ respectively, obtained from the model:

$\forall i \in \mathcal{I}$,

$$\hat{s}^{(i)}(t+1) = \tilde{U}\hat{s}^{(i)}(t) + q^{(i)}\tilde{c}, \quad \forall i \in \mathcal{I}, \tag{4.22}$$

$$\hat{s}(t) := \begin{pmatrix} \hat{x}^{(i)}(t) \\ \hat{\pi}^{(i)}(t) \end{pmatrix}.$$

The goal, in fact, is to find out how the dynamics described by (4.17) differ from those of the model in (4.22), in which the time dependence does not appear, but we have only the time indipendent matrix $\tilde{U}$ and the time indipendent vector $\tilde{c}$. Of course, model in (4.22) is easier to study, as, given that the matrix $\tilde{U}$ is substochastic, it is stable and an equilibrium point exists (from *Friedkin-Johnsen model* theory [20]), i.e.

$$\lim_{t\to\infty} \hat{s}^{(i)}(t) = \hat{s}^{(i)\star} = (I - \tilde{U})^{-1} q^{(i)}\tilde{c}, , \quad \forall i \in \mathcal{I}. \tag{4.23}$$

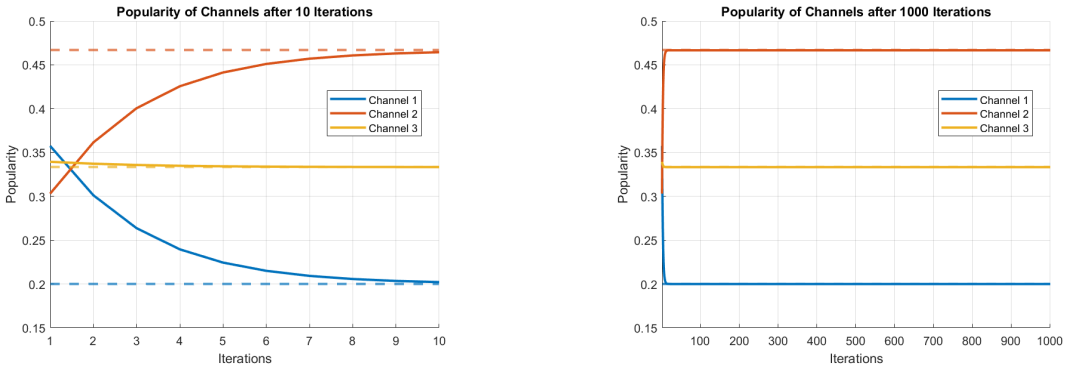
Observation 2. Simulations in Matlab show that at the limit these two models appear to have the same behavior, as the dynamics described in (4.17) converges to the limit values just presented, in (4.23). Moreover, in this case there does not seem to be a relevant dependence on the graph, as the error on the convergence value settles on 10^{-16} for both graphs, hence we can conclude that the two models are equivalent.

Observation 3. The limit values of $\hat{\pi}^{(i)\star}$, to which $\{\pi^{(i)}(t)\}_{t>0}$ simulatively converge, show to be only dependent on the other factors' value of each channel, i.e.

$$\hat{\pi}^{(i)*} = q^{(i)}/q_{tot}, \quad \forall i \in \mathcal{I}.$$

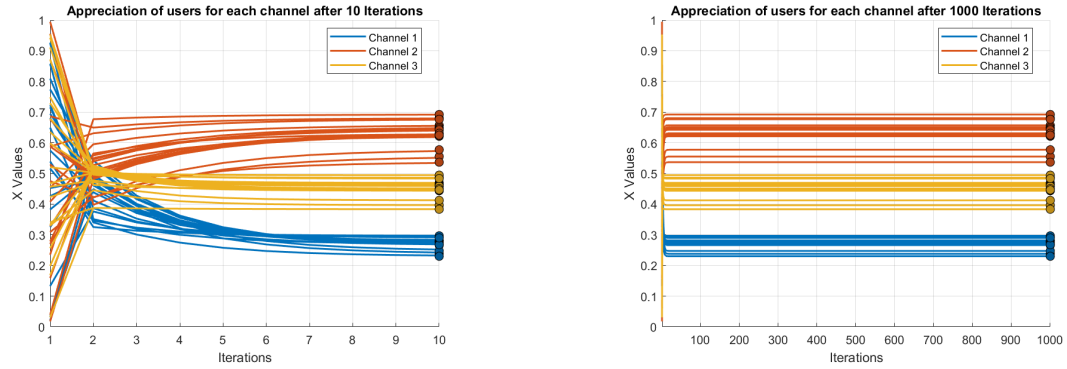
This means that each channel's popularity over time settles on a value that depends on its "expertise"/"experience".

Remark 1. Convergence of dynamics also occurs when the conditions of Lemma 5, i.e. $q_{tot} \geq 1$, $\sum_{i\in\mathcal{I}} x_v^{(i)}(0) \geq 1$ are not met. Figures (4.19) and (4.20) show the evolution of dynamics of $\pi^{(i)}(t)$ and $x_v^{(i)}(t)$ when these assumptions are not satisfied: there is no difference from the case where the conditions are valid.



(a) 10 iterations (b) 1000 iterations

Figure 4.15: Evolution of $\pi^{(i)}(t)$ from (4.15) on a complete graph, in the setting of Example 4.3.5. The dotted lines correspond to $\hat{\pi}^{(i)\star}$ from (4.23).



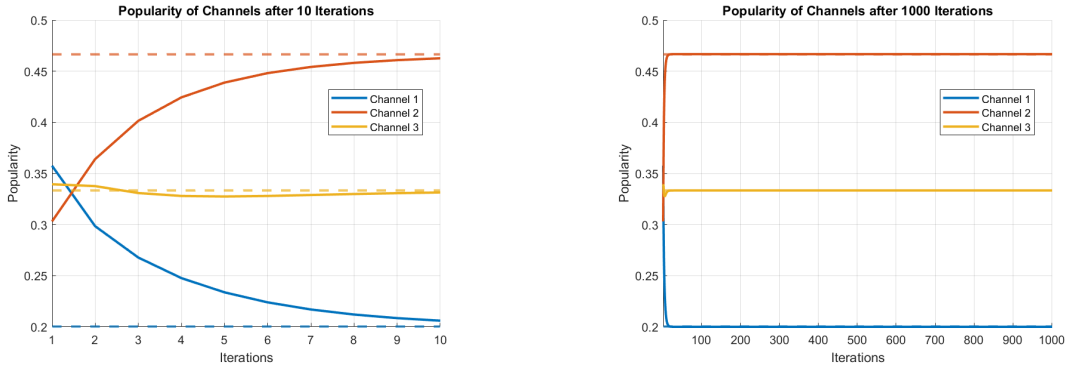
(a) 10 iterations (b) 1000 iterations

Figure 4.16: Evolution of $x_v^{(i)}(t)$ from (4.15) on a complete graph, in the setting of Example 4.3.5. The colored dots correspond to $\hat{x}_v^{(i)\star}$ from (4.23).

4.4 Final Observations

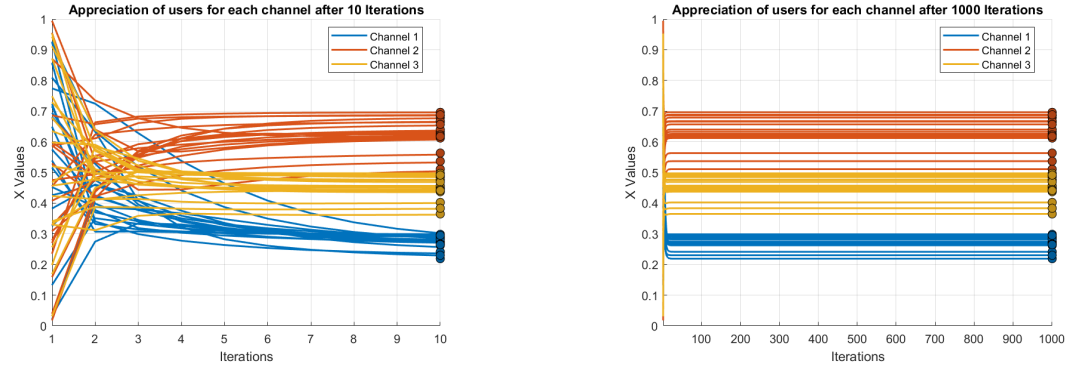
The analysis of the dynamics in (4.1), as the weights associated with each contribution varied, yielded interesting results. First, we were able to find (non-restrictive) conditions that guaranteed convergence in all cases. Except for the cases known in the literature (4.3.3 and 4.3.2), all others describe dynamics that we have studied for the first time.

Second, we can conclude that whenever the term defining the external factors enters the dynamics, it has some dependence on the obtained equilibrium values: the limit popularity for each channel is often represented by the normalized $q^{(i)}$ value respect to all channels (4.3.1 and 4.3.3). When this term is not present (all channels are considered equally), however, the dynamics tend to show a consensus among all users on the importance given to a certain channel, where the ranking associated with different channels depends on the initial conditions $x_v^{(i)}(0)$ (4.3.4 and 4.3.2).



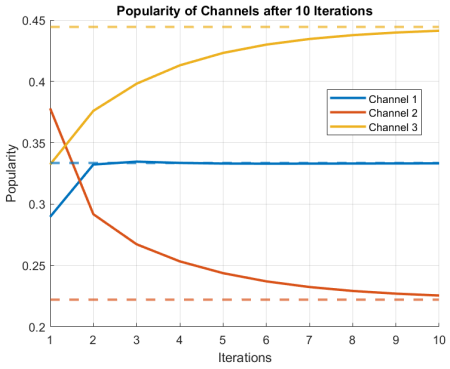
(a) 10 iterations (b) 1000 iterations

Figure 4.17: Evolution of $\pi^{(i)}(t)$ from (4.15) on an Erdős–Rényi graph, in the setting of Example (4.3.5). The dotted lines correspond to $\hat{\pi}^{(i)\star}$ from (4.23).

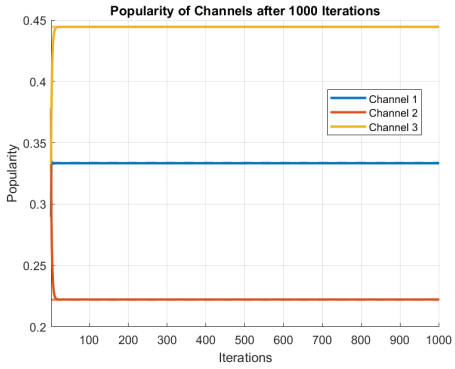


(a) 10 iterations (b) 1000 iterations

Figure 4.18: Evolution of $x_v^{(i)}(t)$ from (4.15) on an Erdős–Rényi graph, in the setting of Example (4.3.5). The colored dots correspond to $\hat{x}_v^{(i)\star}$ from (4.23).

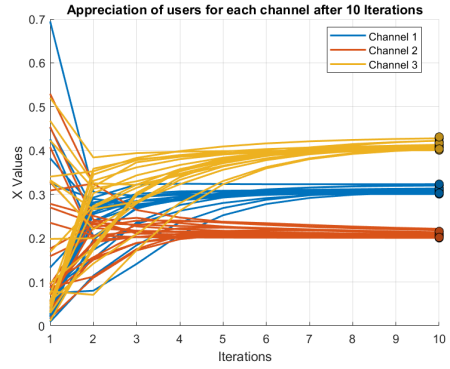


(a) 10 iterations

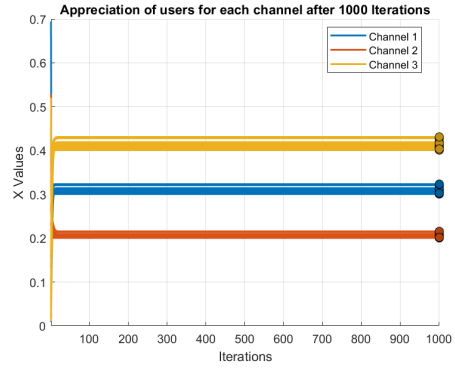


(b) 1000 iterations

Figure 4.19: Evolution of $\pi^{(i)}(t)$ from (4.15) on a complete graph, in the setting of Example 4.3.5 when assumptions are not satisfied. The dotted lines correspond to $\hat{\pi}^{(i)\star}$ from (4.23).



(a) 10 iterations



(b) 1000 iterations

Figure 4.20: Evolution of $x_v^{(i)}(t)$ from (4.15) on a complete graph, in the setting of Example 4.3.5 when assumptions are not satisfied. The colored dots correspond to $\hat{x}_v^{(i)\star}$ from (4.23).

Chapter 5

Stochastic Modeling of Popularity Dynamics

This chapter aims to study the stochastic nonlinear model in its entirety and to analyze its behaviors as the weights associated with its terms change. Conditions and theorems aimed at understanding steady-state behavior will be obtained, and these results once again will be supported by simulations. As anticipated, it will be proved that the dynamics described in the chapter 4 approximates its expected dynamics.

5.1 Mathematical Framework

The mathematical framework of the system is not particularly different from what was presented in section 4.2. We keep all introduced definitions consistent, including user set $\mathcal{V}$, channel set $\mathcal{I}$, user content appreciation measure $x_v^{(i)}(t)\ \forall v,\ \forall i,\ t \in \mathbb{N} \cup \{0\}$, and ranking $\pi^{(i)}(t),\ \forall i$, which aligns with the channel's popularity value.

What changes in this case is the introduction of a random variable into the appreciation measure, which makes the dynamics more realistic, showing how the success of the content is not entirely predictable. We will call this new key variable $Y_v^{(i)}(t)$ and associate with it the meaning of the success of the content proposed by channel i for user v at time t. It satisfies the condition:

$$\mathrm{E}\left[Y_v(t) \mid x_v(t)\right] = x_v(t), \tag{5.1}$$

indicating that the success of the content, in expectation, aligns with the user's appreciation of it. In our case, each $Y_v^{(i)}(t)$ is formalized with a Bernoulli variable, then $Y_v^{(i)}(t) \in \{0,1\}\ \forall v,\ \forall i$ and:

- with probability $x_v^{(i)}(t)$, $Y_v^{(i)}(t) = 1$ (the content of i is successful for the user v);
- with probability $1 - x_v^{(i)}(t)$, $Y_v^{(i)}(t) = 0$ (the content of i is not successful for the user v).

All properties of the terms that enter the dynamic, user interactions, the role of the recommendation platform, and external factors associated with the channel $q^{(i)}$, are the same as set in the section 4.2, so we refer to that section for additional information.

5.2 Dynamics Definition and Theoretical Analysis

The dynamic of the appreciation $x_v^{(i)}(t)$ evolves according to the following equation for all $v \in \mathcal{V}$: $\forall v,\ \forall i$

$$x_v^{(i)}(t+1) = \alpha_v \sum_w P_{vw} Y_w^{(i)}(t) + \beta_v \pi^{(i)}(t) + \gamma_v q^{(i)}. \tag{5.2}$$

where for all $i \in \mathcal{I}$ the ranking $\pi^{(i)}(t)$ of channel i at time t is defined as:

$$\pi^{(i)}(t) = \frac{\sum_{v\in\mathcal{V}} Y_v^{(i)}(t)}{\sum_{j\in\mathcal{I}} \sum_{w\in\mathcal{V}} Y_w^{(j)}(t)}. \tag{5.3}$$

Eq. (5.3) shows that the ranking of a specific channel i is defined by the success of its content normalized against the success of the content of all other channels. In this case, therefore, the ranking value $\pi^{(i)}(t)$ is not constructed from user appreciation $x_v^{(i)}(t)$, but rather from the success of the content $Y_v^{(i)}(t)$, which as we know from (5.1), aligns in expectation with the appreciation $x_v^{(i)}(t)$.

Being a stochastic model, studying the dynamics theoretically is a bit more complex. To do this, we will use some slightly finer mathematical tools. Subsection 5.2.1 will show how to define a Markov chain which will help us analyze the dynamics behavior over time. In this manner, the mathematical framework of this structure will be a reference for rigorous results. Following subsection 5.2.2, will clarify a delicate aspect, namely how we manage to approximate the expected dynamics of the stochastic model described in (5.2), with the deterministic model studied in the chapter 4. After that, similarly to what we saw in the section 4.2, we will proceed by studying different cases of the model in (5.2), associated with different values of parameters, showing and analyzing the results obtained in terms of equilibrium points and stability.

- In section 5.2.3, we will study what happens in the general case, when all terms give contribution into the dynamics, namely, $\alpha_v \neq 0, \beta_v \neq 0, \gamma_v \neq 0$ for some $v \in \mathcal{V}$, looking at the model (5.2) in its entirety. We will analyze the properties of the Markov Chain and obtain relations between the time-averaged dynamics and the expected dynamics.
- In sections 5.2.4 and 5.2.5 we will analyze the case of $\alpha_v = 0\ \forall v \in \mathcal{V}$ and $\beta_v = 0$ for all $v \in \mathcal{V}$. They correspond to the case where user interaction is not considered in the appreciation's evolution and the case in which we have only the network effect and the external factors, respectively. We will exploit the general case theory to get some interesting results.
- Section 5.2.6 will be devoted to analyzing the case with $\gamma_v = 0$ for all $v \in \mathcal{V}$, to understand what happens when only network effects and recommendations influence the dynamics. In this case the chain will no longer be ergodic, so we will show the presence of absorbing states.

5.2.1 How to establish a Markov chain

The model described in (5.2) has a dependence on the random variables $Y_v^{(i)}(t)$, which are the only source of randomness. We know that each of these variables is defined as a discrete Bernoulli variable, so $Y_v^{(i)}(t) \in \{0,1\}\ \forall t,\ \forall v,\ \forall i$. We can model the system as a Markov chain, where these Bernoulli random variables determine the states. In fact, the system exhibits the *Markov property* because the future state $x_v^{(i)}(t+1)$ depends only on $Y_v^{(i)}(t)$, and not on the entire history of the system.

To better show the presence of stochasticity in the model, we can rewrite the dynamics in (5.2) as

$\forall v,\ \forall i$

$$x_v^{(i)}(t+1) = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} Y_w^{(i)}(t) + \beta_v \frac{\sum_{v \in \mathcal{V}} Y_v^{(i)}(t)}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} Y_w^{(j)}(t)} + \gamma_v q^{(i)}, \tag{5.4}$$

where we exploited the definition of ranking $\pi^{(i)}(t)$ introduced in (5.3).

To define the states of the Markov Chain we must first note that, by fixing the channel i, all the $Y_v^{(i)}(t)$ express a combination of 0 and 1 depending on what the outcome of his content is (successful or not) for each user. Thus, we can have 2^n configurations for each channel, of the form $(Y_1^{(i)}, Y_2^{(i)}, \dots, Y_n^{(i)})$, where each $Y_v^{(i)} \in \{0,1\}$, $\forall v$ (the time dependence no longer appears as we reason about the configurations). At the same time, however, we must take into account the possible coupling of the various channels, which is related to the fraction term associated with the definition of ranking $\pi^{(i)}(t)$. This implies that the possible configurations are 2^{nm}, where m is the number of channels in our system, that include the successful outcome of all channels by all users.

The states of the Markov chain are precisely defined as these configurations, and each of them can be represented as

$$S_l = \begin{bmatrix} 0 & 0 & 1 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \cdots & 1 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & 1 & \cdots & 0 & 1 \\ 0 & 0 & 1 & \cdots & 1 & 1 \end{bmatrix} \in \mathbb{R}^{n \times m},$$

where on each column of this matrix, we find the success variables for the content of channel i for each user, namely $(Y_1^{(i)}, Y_2^{(i)}, \dots, Y_n^{(i)})$.

We can now define the transition matrix W of the Markov chain, where $W_{l' \leftarrow l}$ represents the probability of transitioning from state S_l to state $S_{l'}$:

$$W_{l' \leftarrow l} = \mathbb{P}(S(t+1) = S_{l'} \mid S(t) = S_l).$$

To define correctly the transition probability, we first have to introduce all the tools we need. For a specific channel i, $\tilde{S}_l^{(i)}$ is the subset of users v for whom $(S_l)_{(v,i)} = 1$. These are the users for whom the content of the channel i is successful, formally

$$\tilde{S}_l^{(i)} = \{v \in \mathcal{V} \mid (S_l)_{(v,i)} = 1\}.$$

The probability of transitioning from S_l to $S_{l'}$ depends on the success/failure outcomes for each user v on each channel i. This is determined by the dynamics of the system in (5.4):

$$(\theta_l)_v^i = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} (S_l)_{(w,i)} + \beta_v \frac{\sum_{v \in \mathcal{V}} (S_l)_{(v,i)}}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} (S_l)_{(w,j)}} + \gamma_v q^{(i)}.$$

The transition probability $W_{l' \leftarrow l}$ accounts for all users v and channels i:

- If $v \in \tilde{S}_{l'}^{(i)}$, then $(S_{l'})_{(v,i)} = 1$, and we use the success probability $(\theta_l)_v^i$.
- If $v \notin \tilde{S}_{l'}^{(i)}$, then $(S_{l'})_{(v,i)} = 0$, and we use the failure probability $1 - (\theta_l)_v^i$.

Thus, $W_{l' \leftarrow l}$ is the product of probabilities for all users v and channels i:

$$W_{l' \leftarrow l} = \prod_{i=1}^{m} \prod_{v \in \tilde{S}_{l'}^{(i)}} (\theta_l)_v^i \prod_{v \notin \tilde{S}_{l'}^{(i)}} \left(1 - (\theta_l)_v^i\right). \tag{5.5}$$

Since at each step the Bernoulli variables are independent, we are only interested in whether in the state $S_{l'}$ we have 1's or 0's, and choose accordingly the probability of success or failure calculated in the previous state.

The table 5.1 summarizes the notations used while figure 5.1 shows an example of a small part of the Markov Chain when $n = 2,\ m = 2$.

Notation	Description
S_l	Markov chain state matrix, representing success/failure configurations.
$(S_l)_{(v,i)}$	Element in position (v, i) of the matrix S_l.
$\tilde{S}_l^{(i)}$	Users for whom content from channel i is successful in state S_l.
$W_{l' \leftarrow l}$	Probability of transitioning from state S_l to $S_{l'}$.
$(\theta_l)_v^i$	Success probability for user v in channel i given S_l.

Table 5.1: Summary of notation used in the Markov chain model.

5.2.2 Alignment of Deterministic and Stochastic Dynamics

Recalling the dynamics in (5.2) we can compute their expectation, as they describe a random system. In order to do this, we use the *Law of Total Expectation* (in *Appendix*), namely:

$$\mathbb{E}[x_v^{(i)}(t+1)] = \mathbb{E}[\mathbb{E}[x_v^{(i)}(t+1)|\mathcal{F}_t]], \tag{5.6}$$

with $\mathcal{F}_t$ the filtration at time t.

Applying the first expected value with respect to filtration, we obtain

$$\mathbb{E}[x_v^{(i)}(t+1)|\mathcal{F}_t] = \alpha_v \sum_w P_{vw} \mathbb{E}[Y_w^{(i)}(t) \mid \mathcal{F}_t] + \beta_v \mathbb{E}[\pi^{(i)}(t) \mid \mathcal{F}_t] + \gamma_v q^{(i)}$$

$$= \alpha_v \sum_w P_{vw} x_w^{(i)}(t) + \beta_v \mathbb{E}[\pi^{(i)}(t) \mid \mathcal{F}_t] + \gamma_v q^{(i)} =$$

$$= \alpha_v \sum_w P_{vw} x_w^{(i)}(t) + \beta_v \mathbb{E}\left[\frac{\sum_{v \in \mathcal{V}} Y_v^{(i)}(t)}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} Y_w^{(j)}(t)} \mid \mathcal{F}_t\right] + \gamma_v q^{(i)},$$

where we exploited the condition in (5.1), i.e. $\mathrm{E}\left[Y_v(t) \mid x_v(t)\right] = x_v(t)$.

We need to calculate:

$$\mathbb{E}\left[\frac{\sum_{v \in \mathcal{V}} Y_v^{(i)}(t)}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} Y_w^{(j)}(t)} | \mathcal{F}_t\right], \ \forall i$$

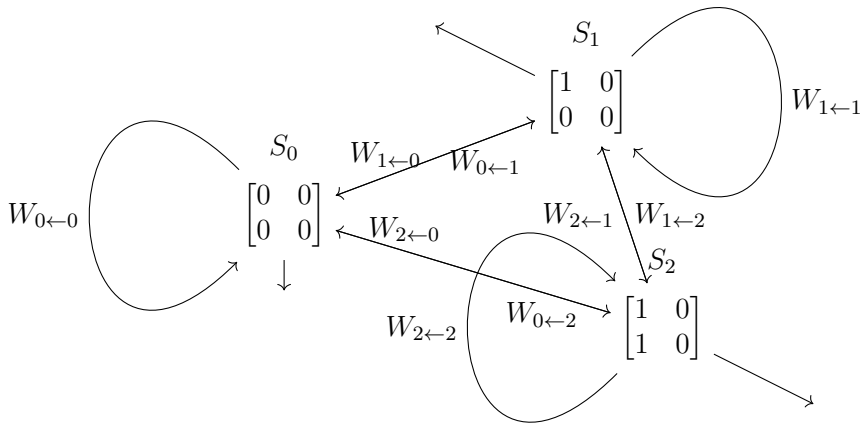


Figure 5.1: Example of a part of the Markov Chain for $n = 2$ and $m = 2$, with transitions extending to absent states.

which is the expected value of the ratio of two random variables. To continue in the discussion we need to understand how this expectation differs from the ratio of expectations.

Recalling (5.1), we can recognize that at both numerator and denominator, there are generalizations of binomials, as sums of Bernoulli variables with different averages, also known as *Poisson Binomials*. Given that each variable $Y_v^{(i)}(t)\ \forall v \in \mathcal{V}, \forall i \in \mathcal{I}$ is independent of the others, we have

$$\mathbb{E}\left[\sum_{v\in\mathcal{V}} Y_v^{(i)}(t)\right] = \sum_{v\in\mathcal{V}} \mathbb{E}\left[Y_v^{(i)}(t)\right] = \sum_{v\in\mathcal{V}} x_v^{(i)}(t),\ \forall i$$

and

$$\mathbb{E}\left[\sum_{j\in\mathcal{I}}\sum_{w\in\mathcal{V}} Y_w^{(j)}(t)\right] = \sum_{j\in\mathcal{I}}\sum_{w\in\mathcal{V}} \mathbb{E}\left[Y_w^{(j)}(t)\right] = \sum_{j\in\mathcal{I}}\sum_{w\in\mathcal{V}} x_w^{(j)}(t).$$

From the theory of probability calculus, we know that the relationship for which the expected ratio is the ratio of expected values does not hold, but there are different theoretical approximations of the former using the latter, depending on the random variables in question.

Since obtaining an accurate theoretical relationship with which to estimate the error involved in approximating the expected ratio to the ratio of expectations is a lengthy process that is beyond the scope of this thesis, we have limited ourselves to proving that the approximation is valid numerically, using Monte Carlo simulations (we include a strategy of theoretical proof in Section 8.2 of Appendix). In other words, what we did was to simulate with Monte Carlo the

value of the expected ratio, such as

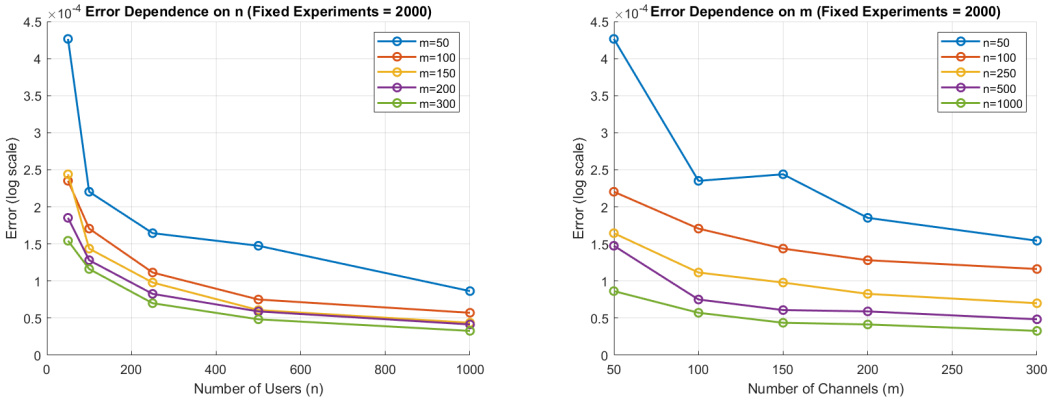
$$exp\_ratio^{(i)} = \frac{1}{K_{tot}} \sum_{k=1}^{K_{tot}} \left[\frac{\sum_{v \in \mathcal{V}} Y_v^{(i)}(k)}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} Y_w^{(j)}(k)} \right] \quad \forall i,$$

where at each step k, $Y_v^{(i)}(k)$ is sampled from a Bernoulli of mean $x_v^{(i)}$, and then to compare this value with the ratio of expectations, that in this case is

$$real\_ratio^{(i)} = \frac{\sum_{v \in \mathcal{V}} x_v^{(i)}}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} x_w^{(j)}}, \quad \forall i.$$

Since they are two vectors of $|\mathcal{I}|$ components, to evaluate the error we calculated $\|real\_ratio - exp\_ratio\|_2$.

Figure 5.2 shows how the estimation improves as n and m, the number of users and channels respectively, increase when the experiments are fixed to 2000.



(a) Fixed m

(b) Fixed n

Figure 5.2: Variation of the error associated to Monte Carlo simulations for different combinations of n and m.

We can therefore agree that

$$\mathbb{E}\left[\frac{\sum_{v \in \mathcal{V}} Y_v^{(i)}(t)}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} Y_w^{(j)}(t)} | \mathcal{F}_t \right] \approx \frac{\mathbb{E}[\sum_{v \in \mathcal{V}} Y_v^{(i)}(t) | \mathcal{F}_t]}{\mathbb{E}[\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} Y_w^{(j)}(t) | \mathcal{F}_t]} = \frac{\sum_{v \in \mathcal{V}} x_v^{(i)}(t)}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} x_w^{(j)}(t)}, \forall i.$$

Returning to $\mathbb{E}[x_v^{(i)}(t+1)|\mathcal{F}_t]$ we can conclude that its dynamics can be approximated by deterministic dynamics, namely

$\forall v, \forall i$

$$\mathbb{E}[x_v^{(i)}(t+1)|\mathcal{F}_t] \approx \alpha_v \sum_w P_{vw} x_w^{(i)}(t) + \beta_v \frac{\sum_{v \in \mathcal{V}} x_v^{(i)}(t)}{\sum_{w \in \mathcal{V}} \sum_{i \in \mathcal{I}} x_v^{(i)}(t)} + \gamma_v q^{(i)}.$$

The trick is to note that the latter dynamics define exactly the system studied in the chapter 4, whose conditions and convergence values are already known. This approximation will be very useful in the following sections, because it will be exploited to visualize behavior at the limit of expected dynamics, albeit with some unknown, but still small, error term.

Remark 2. Once this approximation is assumed, the second expectation indicated in (5.6) has no longer relevance because the dynamics used to approximate are already all deterministic.

5.2.3 Asymptotic behavior with all effects

We begin our analysis from the general case, reasoning about the properties of the Markov chain defined in section 5.2. Let us recall the dynamics:

$\forall v,\ \forall i$

$$x_v^{(i)}(t+1) = \alpha_v \sum_w P_{vw} Y_w^{(i)}(t) + \beta_v \pi^{(i)}(t) + \gamma_v q^{(i)}. \tag{5.7}$$

Knowing that we can approximate the expected dynamics $\mathbb{E}[x_v^{(i)}(t+1)]$ with the deterministic model described in section 4.3.5, it is important to understand the relationship between the stochastic model in (5.7) and its expectation.

Proposition 8. *Assuming that* $0 < q^{(i)} < 1\ \forall i \in \mathcal{I}$*, the Markov chain associated with the dynamics in* (5.7) *is irreducible and aperiodic, thus ergodic.*

Proof. We begin the proof by proving the irreducibility of the chain.

We know that

$$W_{l' \leftarrow l} = \prod_{i=1}^{m} \prod_{v \in \tilde{S}_{l'}^{(i)}} (\theta_l)_v^i \prod_{v \notin \tilde{S}_{l'}^{(i)}} \left(1 - (\theta_l)_v^i\right),$$

where

$$(\theta_l)_v^i = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} (S_l)_{(w,i)} + \beta_v \frac{\sum_{v \in \mathcal{V}} (S_l)_{(v,i)}}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} (S_l)_{(w,j)}} + \gamma_v q^{(i)}.$$

Since $\gamma_v \neq 0\ \forall v$, and assuming that $q^{(i)} > 0\ \forall i$, we can easily observe that $(\theta_l)_v^i > 0\ \forall v,\ \forall i,\ \forall l$. At the same time, by imposing that $q^{(i)} < 1\ \forall i$, we can see that

$$(\theta_l)_v^i = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} (S_l)_{(w,i)} + \beta_v \frac{\sum_{v \in \mathcal{V}} (S_l)_{(v,i)}}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} (S_l)_{(w,j)}} + \gamma_v q^{(i)} \leq \alpha_v + \beta_v + \gamma_v q^{(i)} < 1.$$

From the latter observations we can conclude that

$$W_{l' \leftarrow l} > 0\ \forall l, l',$$

so the chain is irreducible and aperiodic in that selfloops exist. Moreover, it is possible to go from one state to any other directly, so the chain possesses all possible connections.

Since it is therefore finite, aperiodic and irreducible, we conclude that the Markov chain associated with dynamics is ergodic. □

Having proved that the chain is ergodic, it is possible to derive the invariant distribution of the system. The invariant distribution μ satisfies the equation:

$$\mu^T W = \mu,$$

where W is the transition matrix of the chain.

In an ergodic system, the invariant distribution gives the long-term probabilities of being in each state of the Markov chain. Over a long period, the proportion of time spent in state S_l will

converge to μ_l (that is the l-component of the distribution). At the same time, we can compute the expected time spent in that state, given we start in S_l. It is equal to:

$$\mathbb{E}[T_l] = \frac{1}{\mu_l},$$

with T_l the time spent in state S_l before transitioning to a different state. This relationship indicates that states with a higher invariant probability are expected to be visited more frequently.

As the system is ergodic we can exploit the Ergodic Theorem (Theorem 6 in Appendix), concluding that the limit of the time-averaged stochastic dynamics coincides with the spatial expectation of the limit value, computed respect to the invariant distribution.

$$\mathbb{E}_\mu[x_v^{(i)\star}] = \lim_{T\to\infty} \frac{1}{T} \int_0^T x_v^{(i)}(t)\, dt,$$

where $x_v^{(i)\star} = \lim_{t\to\infty} x_v^{(i)}(t)$.

Example 5.2.1. We consider three channels and we set the parameters α_v $\forall v \in \mathcal{V}$ and the values of external factors $q^{(i)}$, representing competence/quality of channels $\forall i \in \mathcal{I}$, by sampling from a uniform distribution defined in (0,1). The initial condition of $x_v^{(i)}(0)$ is obtained in the same way $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$: each value $x_v^{(i)}(0)$ is obtained from a uniform distribution defined in [0,1]. The number of users in the social network is equal to 20. The graph used to simulate the users' interactions is a complete graph.

We show in Figure 5.3 and in Figure 5.4 the evolution of popularity and appreciation of the users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and 5000 instants of time (see figure on the right). In both figures, the upper part shows stochastic dynamics, and the lower part shows time-averaged dynamics, i.e.

$$\bar{x}_v^{(i)}(T) = \frac{1}{T} \sum_{t=0}^{T} x_v^{(i)}(t)$$

and

$$\bar{\pi}^{(i)}(T) = \frac{1}{T} \sum_{t=0}^{T} \pi^{(i)}(t).$$

In the lower part, the dotted lines show the limit values of the dynamics in (4.15) in terms of popularity (which we have proven that approximate the expected dynamics $\{\mathbb{E}[\pi^{(i)}(t)]\}_{t\in\mathbb{N}}$), while the colored dots correspond to the limit values of the dynamics in (4.15) in terms of appreciation (which we have proven that approximate the expected dynamics $\mathbb{E}\{[x_v^{(i)}(t)]\}_{t\in\mathbb{N}}$ for all $i \in \mathcal{I}$). As a consequence of the ergodic theorem, the plots show that the time-averaged dynamics approach these limit values, although there is always an error term since these are not the true expected dynamics but an approximation of them.

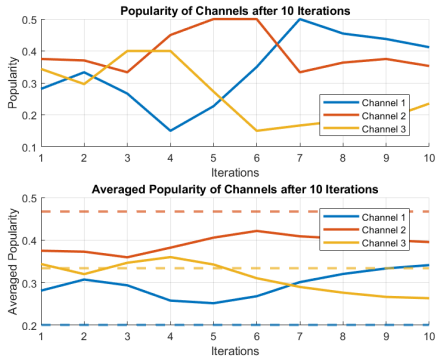
5.2.4 Asymptotic Behavior without Network Effect

Let's analyze now what happens by imposing that $\alpha_v = 0$ $\forall v \in \mathcal{V}$, that is, canceling interactions between users.

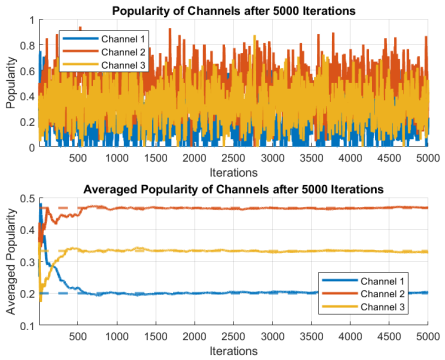
The dynamics in (5.2) becomes:

$\forall v,\ \forall i$

$$x_v^{(i)}(t+1) = \beta_v \pi^{(i)}(t) + (1-\beta_v) q^{(i)}. \tag{5.8}$$

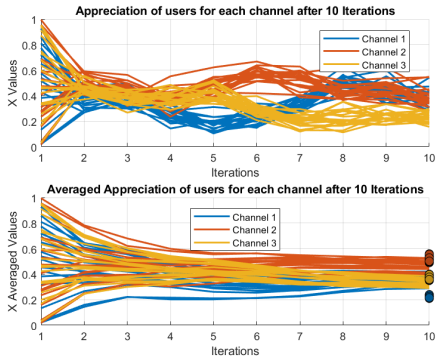


(a) 10 iterations

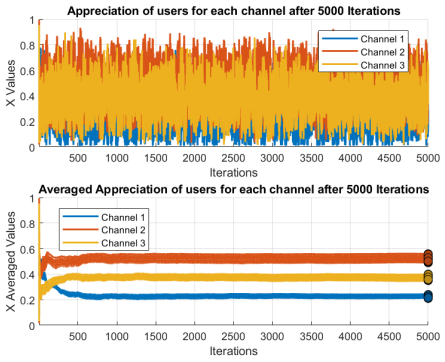


(b) 5000 iterations

Figure 5.3: Evolution of $\pi^{(i)}(t)$ from (5.7) and $\bar{\pi}^{(i)}(t)$ in the setting of Example 5.2.1. The dotted lines correspond to $\hat{\pi}^{(i)\star}$ from (4.22).



(a) 10 iterations



(b) 5000 iterations

Figure 5.4: Evolution of $x_v^{(i)}(t)$ from (5.7) and $\bar{x}_v^{(i)}(t)$ in the setting of Example 5.2.1. The dotted lines correspond to $\hat{x}_v^{(i)\star}$ from (4.22).

Knowing that we can approximate the expected dynamics $\mathbb{E}[x_v^{(i)}(t+1)]$ with the deterministic model described in section 4.3.1, it is important to understand the relationship between the stochastic model in (5.8) and its expectation.

Since we assume that $0 < q^{(i)} < 1 \ \forall i$ otherwise the influence of the external factor would cancel out, we can conclude that the Proposition 8 is of the section 5.2.3 is still valid and thus the Markov chain ergodic.

Let's see what happens in simulation with a simple example.

Example 5.2.2. We consider three channels and we set the parameters $\alpha_v \ \forall v \in \mathcal{V}$ and the values of external factors $q^{(i)}$, representing competence/quality of channels $\forall i \in \mathcal{I}$, by sampling from a uniform distribution defined in (0,1). The initial condition of $x_v^{(i)}(0)$ is obtained in the same way $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$: each value $x_v^{(i)}(0)$ is obtained from a uniform distribution defined in [0,1]. The number of users in the social network is equal to 20.

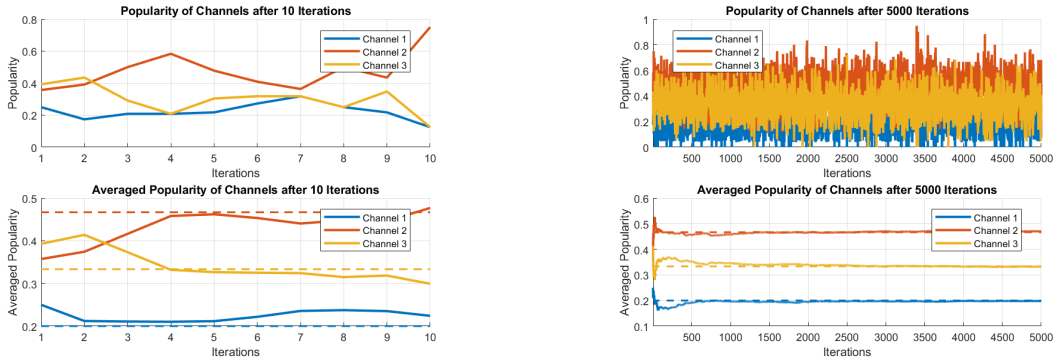
We show in Figure 5.5 and in Figure 5.6 the evolution of popularity and appreciation of the

users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and 5000 instants of time (see figure on the right). In both figures, the upper part shows stochastic dynamics, and the lower part shows time-averaged dynamics, i.e.

$$\bar{x}_v^{(i)}(T) = \frac{1}{T}\sum_{t=0}^{T} x_v^{(i)}(t)$$

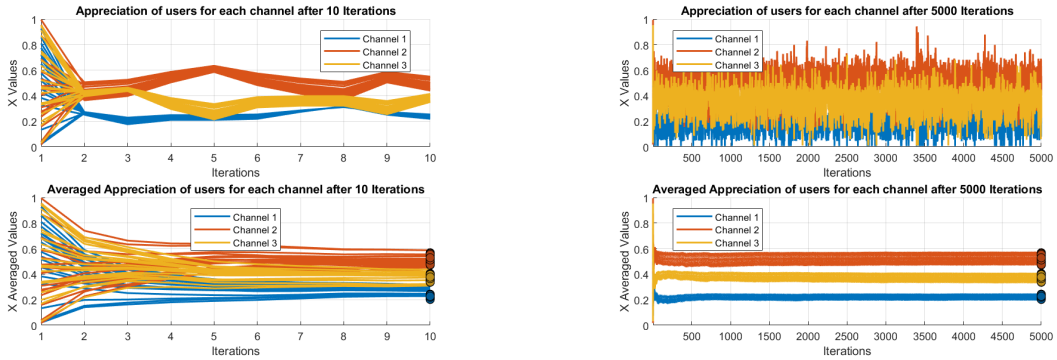
and

$$\bar{\pi}^{(i)}(T) = \frac{1}{T}\sum_{t=0}^{T} \pi^{(i)}(t).$$



(a) 10 iterations (b) 5000 iterations

Figure 5.5: Evolution of $\pi^{(i)}(t)$ from (5.8) and $\bar{\pi}^{(i)}(t)$ in the setting of Example 5.2.2. The colored dots correspond to $\pi^{(i)\star}$ from (4.2).



(a) 10 iterations (b) 5000 iterations

Figure 5.6: Evolution of $x_v^{(i)}(t)$ from (5.8) and $\bar{x}_v^{(i)}(t)$ in the setting of Example 5.2.2. The dotted lines correspond to $x_v^{(i)\star}$ from (4.2).

In the lower part, the dotted lines show the limit values of the dynamics in (4.2) in terms of popularity (which we have proven that approximate the expected dynamics $\{\mathbb{E}[\pi^{(i)}(t)]\}_{t\in\mathbb{N}}$),

while the colored dots correspond to the limit values of the dynamics in (4.2) in terms of appreciation (which we have proven that approximate the expected dynamics $\mathbb{E}\{[x_v^{(i)}(t)]\}_{t\in\mathbb{N}}$ for all $i \in \mathcal{I}$). As a consequence of the ergodic theorem, the plots show that the time-averaged dynamics approach these limit values, although there is always an error term since these are not the true expected dynamics but an approximation of them.

5.2.5 Asymptotic Behavior Without Platform's Recommendation

The absence of the recommendation platform translates into the equation (5.2) by imposing that $\beta_v = 0,\ \forall v$.

The model is

$\forall v,\ \forall i$

$$x_v^{(i)}(t+1) = \alpha_v \sum_{w\in\mathcal{V}} P_{vw} Y_w^{(i)}(t) + \gamma_v q^{(i)}. \tag{5.9}$$

Regarding the expected dynamics, along the lines of what was done in section 5.2.2, we can observe that since the ratio term is not present, in this case we have not only an approximation but an equality between the two dynamics:

$$\begin{aligned}\mathbb{E}[x_v^{(i)}(t+1) \mid \mathcal{F}_t] &= \mathbb{E}[\alpha_v \sum_{w\in\mathcal{V}} P_{vw} Y_w^{(i)}(t) + \gamma_v q^{(i)} \mid \mathcal{F}_t] \\ &= \alpha_v \sum_{w\in\mathcal{V}} P_{vw} \mathbb{E}[Y_w^{(i)}(t)|\mathcal{F}_t] + (1-\alpha_v) q^{(i)} \\ &= \alpha_v \sum_{w\in\mathcal{V}} P_{vw} x_w^{(i)}(t) + (1-\alpha_v) q^{(i)},\end{aligned}$$

which is already deterministic, so applying the expectation again makes no difference. As expected, the obtained dynamics are exactly those of the section 4.3.3, defined in (4.4).

Knowing how the expected dynamics behave, we must understand the relationship between the latter and the initial stochastic model.

Once again, assuming that $0 < q^{(i)} < 1\ \forall i$ we can say that the Proposition 8 still holds and therefore the Markov chain is ergodic.

Let's see what happens simulatively.

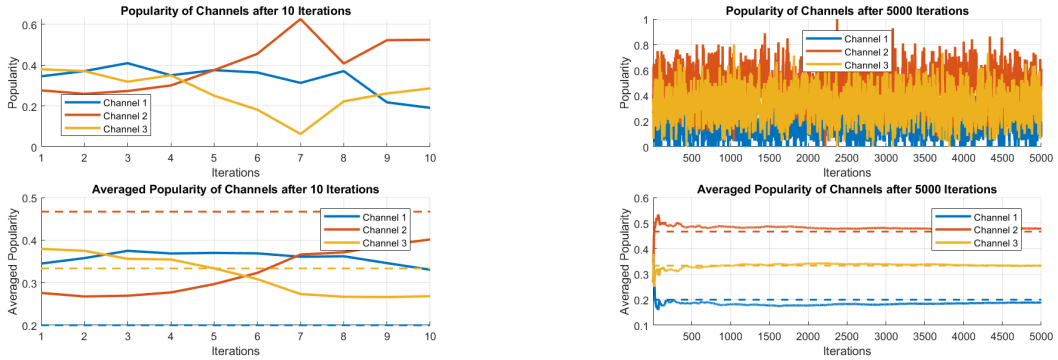
Example 5.2.3. We consider three channels and we set the parameters $\alpha_v\ \forall v \in \mathcal{V}$ and the values of external factors $q^{(i)}$, representing competence/quality of channels $\forall i \in \mathcal{I}$, by sampling from a uniform distribution defined in (0,1). The initial condition of $x_v^{(i)}(0)$ is obtained in the same way $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$: each value $x_v^{(i)}(0)$ is obtained from a uniform distribution defined in [0,1]. The number of users in the social network is equal to 20. The graph used to simulate the users' interactions is a complete graph.

We show in Figure 5.7 and in Figure 5.8 the evolution of popularity and appreciation of the users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and 5000 instants of time (see figure on the right). In both figures, the upper part shows stochastic dynamics, and the lower part shows time-averaged dynamics, i.e.

$$\bar{x}_v^{(i)}(T) = \frac{1}{T}\sum_{t=0}^{T} x_v^{(i)}(t)$$

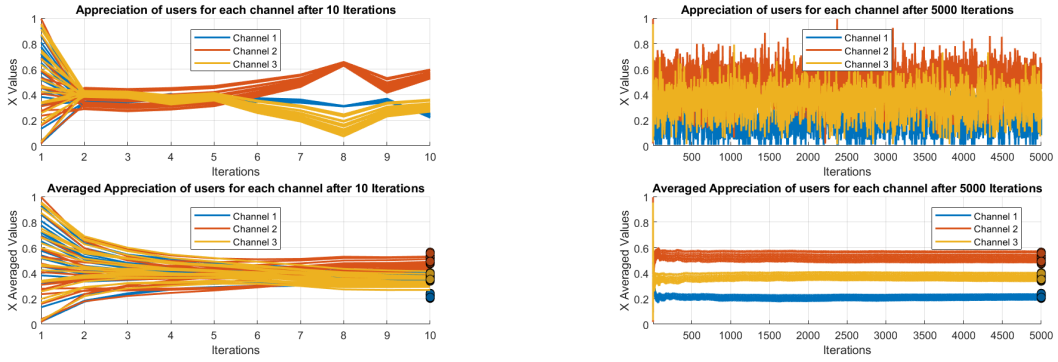
and

$$\bar{\pi}^{(i)}(T) = \frac{1}{T}\sum_{t=0}^{T}\pi^{(i)}(t).$$



(a) 10 iterations (b) 5000 iterations

Figure 5.7: Evolution of $\pi^{(i)}(t)$ from (5.9) and $\bar{\pi}^{(i)}(t)$ in the setting of Example 5.2.3. The dotted lines correspond to $\pi^{(i)\star}$ from (4.4).



(a) 10 iterations (b) 5000 iterations

Figure 5.8: Evolution of $x_v^{(i)}(t)$ from (5.9) and $\bar{x_v}^{(i)}(t)$ in the setting of Example 5.2.3. The colored dots correspond to ${x_v}^{(i)\star}$ from (4.4).

In the lower part, the dotted lines show the limit values of the dynamics in (4.4) in terms of popularity (which we have proven to be coincident with the expected dynamics $\{\mathbb{E}[\pi^{(i)}(t)]\}_{t\in\mathbb{N}}$), while the colored dots correspond to the limit values of the dynamics in (4.4) in terms of appreciation (which we have proven to be coincident with the expected dynamics $\mathbb{E}\{[x_v^{(i)}(t)]\}_{t\in\mathbb{N}}$ for all $i \in \mathcal{I}$). The plots demonstrate the convergence of time-averaged dynamics at these limit values. In other words, it is a confirmation of the validity of the ergodic theorem.

5.2.6 Network effects and recommendations in the absence of external factors

The last case to be studied is obtained by canceling the contribution of external factors, that is, by imposing that $\gamma_v = 0\ \forall v \in \mathcal{V}$. It is also the case that returns the most peculiar results, obtained from a certain Markov chain structure we have not encountered before.

Let's introduce the dynamics, starting with the general model in (5.2). Imposing that $\gamma_v = 0\ \forall v \in \mathcal{V}$, we have

$\forall v,\ \forall i$

$$x_v^{(i)}(t+1) = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} Y_w^{(i)}(t) + (1-\alpha_v)\pi^{(i)}(t). \tag{5.10}$$

Knowing that we can approximate the expected dynamics $\mathbb{E}[x_v^{(i)}(t+1)]$ with the deterministic model described in section 4.3.4, it is important to understand the relationship between the stochastic model in (5.10) and its expectation.

Let us return to the definition of the Markov chain introduced in section 5.2.1. The structure of the chain, as it was defined, and the transition probabilities are still valid. Unlike the cases where the β_v- and α_v-dependent terms are canceled, in this case canceling the contribution of external factors leads to the appearance of absorbing states, which do not make the chain irreducible.

We have the following proposition:

Proposition 9. *The Markov chain associated with the dynamics in (5.10) when $\gamma_v = 0\ \forall v$ has absorbing states, so it is not irreducible. Specifically, for each $i \in \mathcal{I}$, define the states $S_{i\star}$ as*

$$S_{i\star} : \begin{cases} (S_{i\star})_{(v,i)} = 1, & \forall v \in \mathcal{V}, \\ (S_{i\star})_{(v,j)} = 0, & \forall v \in \mathcal{V}, \forall j \neq i,\ j \in \mathcal{I}. \end{cases} \tag{5.11}$$

These states $S_{i\star}$ are absorbing, meaning that the transition probabilities satisfy

$$W_{i\star \leftarrow i\star} = 1, \quad \forall i \in \mathcal{I}, \tag{5.12}$$

$$W_{l' \leftarrow i\star} = 0, \quad \forall\ l' \neq i\star,\ \forall i \in \mathcal{I}, \tag{5.13}$$

and for any other state $S_l \neq S_{i\star}$, we have

$$W_{i\star \leftarrow l} > 0. \tag{5.14}$$

Proof. Recalling that the states of the chain are of the form

$$S_l = \begin{bmatrix} 0 & 0 & 1 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \cdots & 1 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & 1 & \cdots & 0 & 1 \\ 0 & 0 & 1 & \cdots & 1 & 1 \end{bmatrix} \in \mathbb{R}^{n \times m},$$

we can consider the state $S_{i\star}$ defined in this way and show that it is absorbing

$$S_{i\star} : \begin{cases} (S_{i\star})_{(v,i)} = 1 & \forall v \in \mathcal{V} \\ (S_{i\star})_{(v,j)} = 0 & \forall v \in \mathcal{V}, \forall j \neq i,\ j \in \mathcal{I}. \end{cases}$$

In other words, in the $S_{i\star}$ state, the content of channel i is successful for all users, while the content of all other channels $j \neq i$ is unsuccessful for all users. An example when $i = 3$ is:

$$S_{3\star} = \begin{bmatrix} 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{n \times m}.$$

. Once we end up in these states we can observe that:

- For the channel i,

$$(\theta_{i\star})_v^i = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} \cdot 1 + (1 - \alpha_v) \cdot \frac{n}{n} = \alpha_v + (1 - \alpha_v) = 1, \quad \forall v \in \mathcal{V}$$

  since the matrix P is rows-stochastic.

- For all the other channels $j \neq i$,

$$(\theta_{i\star})_v^j = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} \cdot \frac{0}{n} + (1 - \alpha_v) \cdot 0 = 0, \quad \forall v \in \mathcal{V}.$$

Recalling that the probability to go from S_l to $S_{l'}$ is defined as

$$W_{l' \leftarrow l} = \prod_{i=1}^{m} \prod_{v \in \tilde{S}_{l'}^{(i)}} (\theta_l)_v^i \prod_{v \notin \tilde{S}_{l'}^{(i)}} \left(1 - (\theta_l)_v^i\right),$$

we can observe that the only way to have a nonzero probability and exactly equal to 1 starting from $i\star$, is to stay in $i\star$. In that case, in fact, we have the product of all 1's. In all other cases zeros appear in the product that cancel the probability. It shows that the states of the form $S_{i\star}$ are absorbing states.

When the starting state $S_l \neq S_{i\star}$ $\forall i$, $0 < (\theta_l)_v^j < 1$ $\forall l$, $\forall v$, $\forall j$, so $W_{l' \leftarrow l} > 0$, $\forall l'$, and consequently also $W_{i\star \leftarrow l} > 0$. □

Obviously in this case the ergodic theorem no longer applies, and it is not possible to tie the behavior of the time-averaged dynamics to the expected dynamics $\mathbb{E}[x_v^{(i)}(t)]$. What happens, instead, is that the dynamics will quickly concentrate in one of these absorbing states. Unless we start directly from an absorbing state, the dynamics will concentrate in one or the other randomly.

Example 5.2.4. We consider three channels and we set the parameters α_v $\forall v \in \mathcal{V}$ by sampling from a uniform distribution defined in [0,1]. The initial condition of $x_v^{(i)}(0)$ is obtained in the same way $\forall v \in \mathcal{V}, \forall i \in \mathcal{I}$: each value $x_v^{(i)}(0)$ is obtained from a uniform distribution defined in [0,1]. The number of users in the social network is equal to 20. The graph used to simulate the users' interactions is a complete graph.

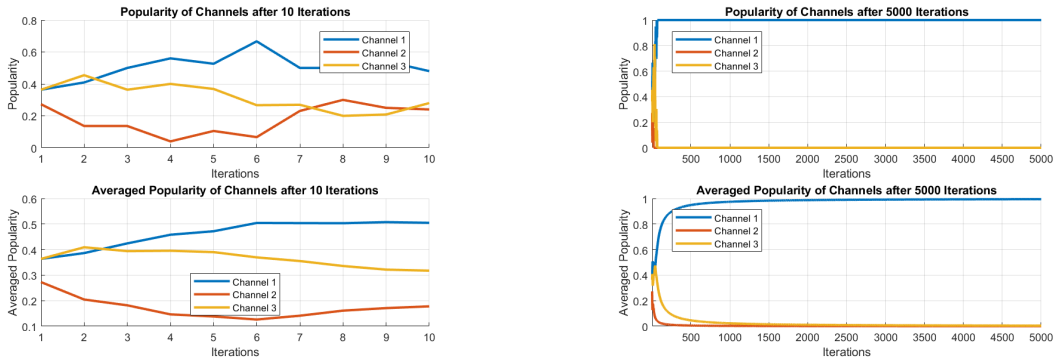
We show in Figure 5.9 and in Figure 5.10 the evolution of popularity and appreciation of the users, respectively, as a function of time, iterating the updates for 10 (see figure on the left) and

5000 instants of time (see figure on the right). In both figures, the upper part shows stochastic dynamics, and the lower part shows time-averaged dynamics, i.e.

$$\bar{x}_v^{(i)}(T) = \frac{1}{T}\sum_{t=0}^{T} x_v^{(i)}(t)$$

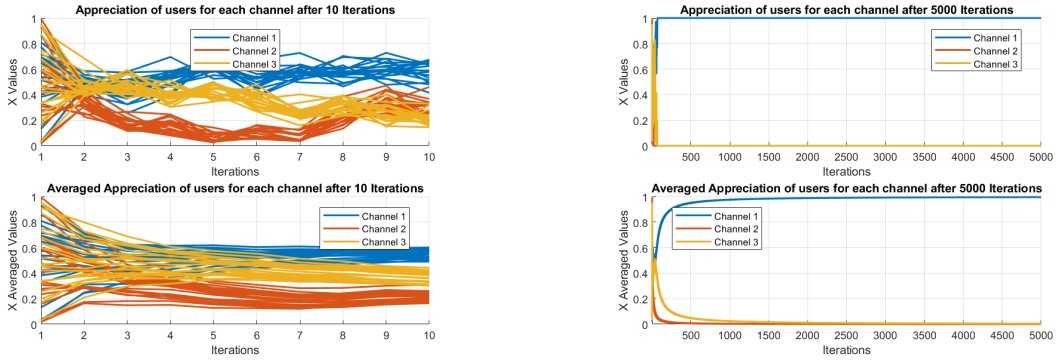
and

$$\bar{\pi}^{(i)}(T) = \frac{1}{T}\sum_{t=0}^{T} \pi^{(i)}(t).$$



(a) 10 iterations (b) 5000 iterations

Figure 5.9: Evolution of $\pi^{(i)}(t)$ from (5.10) and $\bar{\pi}^{(i)}(t)$ in the setting of Example 5.2.4.



(a) 10 iterations (b) 5000 iterations

Figure 5.10: Evolution of $x_v^{(i)}(t)$ from (5.10) and $\bar{x_v}^{(i)}(t)$ in the setting of Example 5.2.4.

As observed theoretically, the dynamics rapidly concentrate all the popularity on a single, randomly chosen channel: the chain enters an absorbing state.

5.3 Final Observations

The analysis of the model in (5.2) showed different results depending on the weights associated with each contribution. First, we observed that it is always possible to construct a Markov chain associated with the dynamics, whose states depend on the configurations assumed by $Y_v^{(i)}(t)$, $\forall i$, $\forall v$. Such a chain is not always ergodic, but its ergodicity depends on the presence of the term in $q^{(i)}$: when it is not present the chain shows absorbing states. So, while in all cases where $q^{(i)}$ appears the temporal mean and the spatial mean obtained using the stationary distribution coincide, by canceling $q^{(i)}$ the popularity is all concentrated on one channel completely at random.

Moreover, by being able to approximate the expected stochastic dynamics with the one studied in the chapter 4, we were able to get an idea of the values at the limit of the time-averaged dynamics, which are then the values around which the stochastic dynamics oscillates.

Chapter 6

Exploring Low Gamma Dynamics

This chapter aims to analyze interesting cases that have not been explored in depth in previous chapters. In particular, we will simulatively analyze the case in which the other factors' value has an increasingly smaller impact in defining appreciation.

6.1 Mathematical Framework and Initial Considerations

In the chapter 4 we discussed the evolution of the stochastic dynamics of appreciation (5.2), by analyzing what happens in the various cases when one of the contributions cancels out, or when all terms interact with each other.

Beyond the cases already studied both theoretically and numerically, one case that was of particular interest when we simulated it is related to requiring that the contribution of the other factors' value for each channel be minimal. To do this, we impose that γ_v, the weight associated with the contribution $q^{(i)}$, thus the external factor, for all users, is about 0 ($\gamma_v \approx 0 \ \forall v \in \mathcal{V}$). The reason for interest can already be clarified qualitatively, before focusing on simulations.

In section 5.2.6, we observed what happens in the case where the contribution of $q^{(i)}$ is completely canceled: absorbing states arise, leading the appreciation to concentrate in one or another channel randomly. This is a different behavior from that in which the $q^{(i)}$ factor is present, studied in the section 5.2.3, in which instead the chain is ergodic and we have the convergence of the time-averaged dynamics to the limit values of the expected dynamics. Moreover, the latter, being an approximation of the deterministic dynamics studied in the section 4.3.5, in terms of popularity has limit values that depend precisely on the values of the external factor for each channel $q^{(i)}$. This means that over time the popularity fluctuates around this factor (barring normalization) and that its time average will converge to these values plus a small error term. (Refer to the section 4.3.5 to go into specifics). Consequently, the case where gamma is small for all users, $\gamma_v \approx 0 \ \forall v$, becomes an intermediate case between this behavior so related to external factors and the presence of the absorbing states.

Recalling the dynamics in 5.2.3:

$\forall i \in \mathcal{I}$,

$$x_v^{(i)}(t+1) = \alpha_v \sum_{w \in \mathcal{V}} P_{vw} Y_w^{(i)}(t) + \beta_v \pi^{(i)}(t) + \gamma_v q^{(i)}, \tag{6.1}$$

with

$$\pi^{(i)}(t) = \frac{\sum_{v \in \mathcal{V}} Y_v^{(i)}(t)}{\sum_{j \in \mathcal{I}} \sum_{w \in \mathcal{V}} Y_w^{(j)}(t)},$$

we already know that imposing $\gamma_v \approx 0 \ \forall v$, does not cause changes in the theoretical formulation, from which the validity of the ergodic theorem is maintained, as shown in the section 4.3.5. Instead, let us observe what happens at the simulative level.

6.2 Simulations and Numerical Results

Simulations were carried out for different orders of magnitude of γ_v: from larger values (order of 10^{-2}) to smaller values (order of 10^{-4}) in order to observe what happens in the dynamics gradually.

6.2.1 Exploring Small-Scale Influence: Gamma on a Moderate Scale

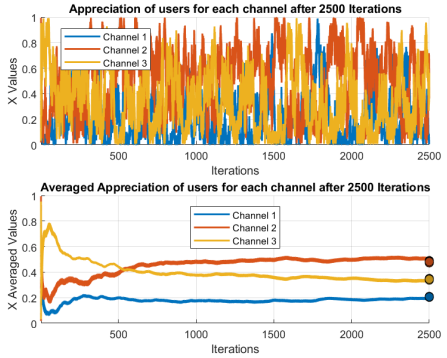
The first interesting case is the one in which all $\gamma_v \ \forall v$ are of the order of 10^{-2}; in this way, we can analyze what happens by imposing that the influence of $q^{(i)}$ is small but not entirely minimal.

Figure 6.1 shows the behavior of the dynamics defined in 5.2.3 (top plot) and of its time average (bottom plot), both in terms of appreciation $x_v^{(i)}(t)$ (Figure 6.1a) and popularity $\pi^{(i)}(t)$ (Figure 6.1b).

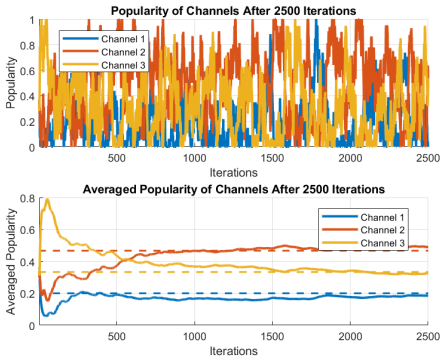
The plots are obtained considering a total number of iterations equal to 2500 and using an Erdos Renyi graph with a probability of having a link equal to 0.5. The initial $x_v^{(i)}(0)$-values are sampled from a uniform distribution between 0 and 1 $\forall v \in \mathcal{V}, \ \forall i \in \mathcal{I}$. The same is done for each $q^{(i)}$. The $\gamma_v \ \forall v$ values are randomly extracted from a uniform distribution defined between 0.01 and 0.09 while α_v from a uniform distribution between 0 and 0.5. The values of β_v are obtained from those of γ_v and $\alpha_v \ \forall v \in \mathcal{V}$ by requiring normalization ($\beta_v = 1 - \alpha_v - \gamma_v \ \forall v$).

In this case, as is visible from the plot, we do not have behaviors that are particularly different from the general case of the section 5.2.3, since here again we have a clear equilibrium point to which the time-averaged dynamics tends. What begins to emerge, however, is a smaller oscillation in stochastic dynamics (visible especially on $\pi^{(i)}(t)$). In other words, it appears that $\pi^{(i)}(t)$ begins to spend more time in configurations where popularity is concentrated on only one channel, and it appears that the preferred channels are those associated with higher $q^{(i)}$ values.

To further clarify this result and see it from another point of view, we constructed a heatmap, shown in the figure 6.2, that illustrates the trend of the stochastic $\pi^{(i)}(t)$ dynamics over time. Having only 3 channels, we can see a symplex. The heatmap is built using a 100*100 grid, and its axes represent the popularity values of channels 1 and 2. From these then we can derive the popularity value of the third channel. The point (1,0) is associated with a concentration of popularity entirely on the first channel, the point (0,1) with complete concentration on the second, and the point (0,0) similarly on the third. The density of the heatmap is represented by colors: the most colored grid dots represent the most visited configurations. The bar next to it is obtained by log-scaling the density in order not to run into too high values and visualize it



(a) Evolution of $x_v^{(i)}(t)\ \forall v \in \mathcal{V}, \forall i \in \mathcal{I}$



(b) Evolution of $\pi^{(i)}(t)\ \forall i \in \mathcal{I}$

Figure 6.1: Dynamics of $x_v^{(i)}(t)$ and $\pi^{(i)}(t)$ from (6.1) obtained by requiring γ_v of the order of 10^{-2}.

better. Again, the number of iterations is 2500. The red dot represents the equilibrium point to which the expected dynamics tend, which in figure 6.1b was represented by the dotted lines.

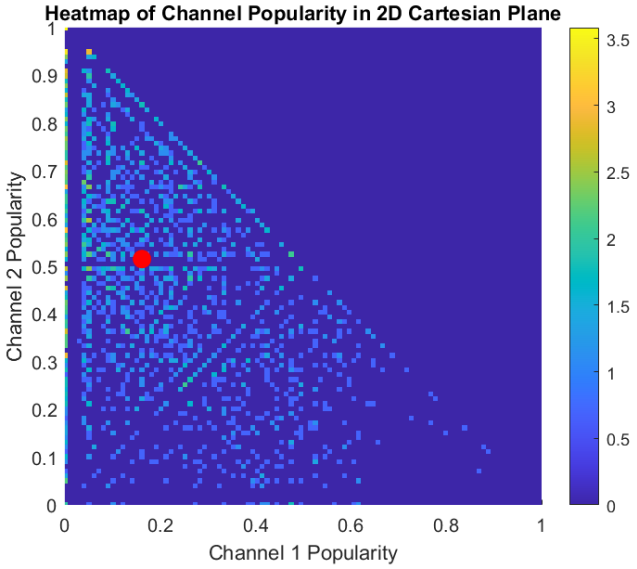


Figure 6.2: Heatmap for evolution of $\pi^{(i)}(t)$ from (6.1) associated to γ_v of the order of 10^{-2}.

It is observed that the dynamics are mostly concentrated in a configuration with higher values of channels 2 and 3 as well as we had visualized from the other plots.

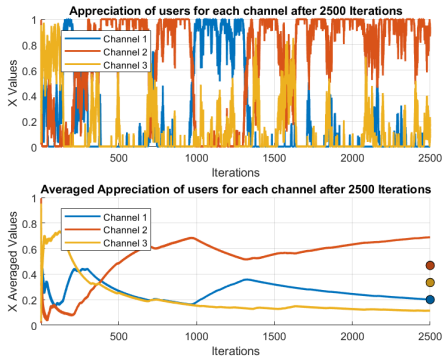
6.2.2 Exploring Finer Adjustments: Gamma on a Smaller Scale

In the previous section, we observed that a reduction of gamma to the second decimal order does not result in large changes but only in some slight differences. Therefore, let us further

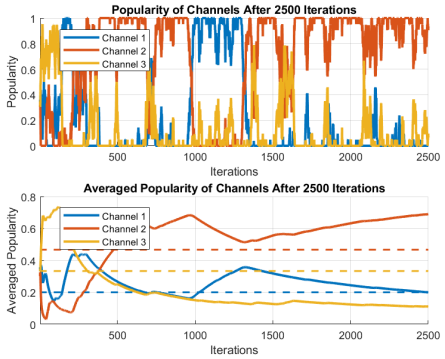
reduce gamma by imposing that $\gamma_v \ \forall v \in \mathcal{V}$ are of the 10^{-3} order of magnitude and observe what happens.

Figure 6.3 shows the behavior of the dynamics defined in 5.2.3 (top plot) and of its time average (bottom plot), both in terms of appreciation $x_v^{(i)}(t)$ (Figure 6.3a) and popularity $\pi^{(i)}(t)$ (Figure 6.3b).

The plots are obtained considering a total number of iterations equal to 2500 and using an Erdos Renyi graph with a probability of having a link equal to 0.5. The initial $x_v^{(i)}(0)$-values are sampled from a uniform distribution between 0 and 1 $\forall v \in \mathcal{V}$, $\forall i \in \mathcal{I}$. The same is done for each $q^{(i)}$. The $\gamma_v \ \forall v$ values are randomly extracted from a uniform distribution defined between 0.001 and 0.009 while α_v from a uniform distribution between 0 and 0.5. The values of β_v are obtained from those of γ_v and $\alpha_v \ \forall v \in \mathcal{V}$ by requiring normalization ($\beta_v = 1 - \alpha_v - \gamma_v \ \forall v$).



(a) Evolution of $x_v^{(i)}(t) \ \forall v \in \mathcal{V}, \forall i \in \mathcal{I}$



(b) Evolution of $\pi^{(i)}(t) \ \forall i \in \mathcal{I}$

Figure 6.3: Dynamics of $x_v^{(i)}(t)$ and $\pi^{(i)}(t)$ from (6.1) obtained by requiring γ_v of the order of 10^{-3}.

Here we observe a clear difference from the previous analysis. Both $\pi^{(i)}(t)$ and $x_v^{(i)}(t)$ begin to spend even more time in configurations in which popularity is associated with only one channel. Configurations in which there is a combination of all three popularities are increasingly rare. The averaged dynamics, on the other hand, still seem to have a pattern of convergence to the equilibrium point, but to be certain of this the model would have to be iterated for a longer time. It is clear that the equilibrium point exists, but the dynamics begin to show similarities to the case where absorbing states appear.

It is natural to ask at this point which channels capture all the popularity for the longest time: to do this we use the heatmap, already presented in the section 6.2.1.

The heatmap shown by the figure 6.4 can give insight into how much a certain configuration is visited. First, although little can be seen from the plot, as the grid is quite dense with points, we can observe that the most visited configurations (yellow squares) are those associated with the concentration of popularity on one channel. Moreover, in most cases, to move from one of these cases to the other, $\pi(t)$ moves along the edges of the simplex, that is, using configurations in which only 2 channels appear and not all 3 (as we had already observed in the figure 6.3b). Going then to analyze the points on the grid, we understand that the most visited configuration (820 times) is the one for which the popularity is all concentrated on the second channel. This is followed by the one where the popularity is concentrated on the first (375 times). Indeed, the channel that dominates users' interest is precisely the one associated with a higher external

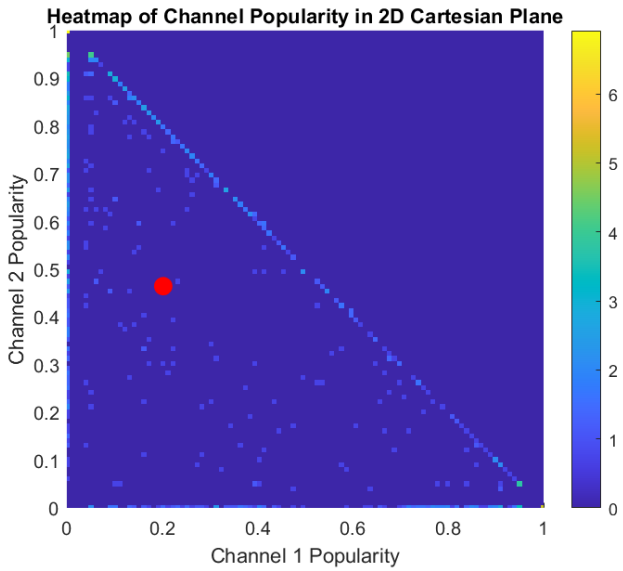


Figure 6.4: Heatmap for evolution of $\pi^{(i)}(t)$ from (6.1) associated to γ_v of the order of 10^{-3}.

factor $q^{(i)}$.

6.2.3 Examining Minimal Influence: Gamma on a Micro Scale

Let us now analyze what happens when we reduce the gamma contribution even more, requiring that $\gamma_v \ \forall v$ is on the order of 10^{-4}.

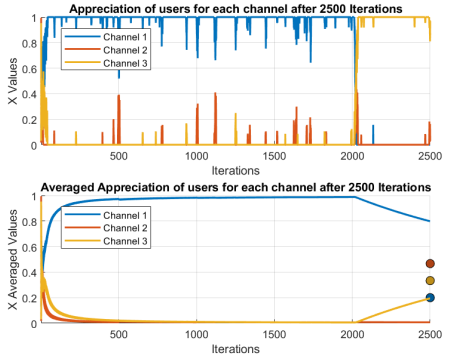
Given already the behavior observed in the section 6.2.2, we expect dynamics that are even more reminiscent of that of the case with $\gamma_v = 0 \ \forall v$, damped occasionally by the effect of the external factor $q^{(i)}$, which although very small, is nevertheless present.

Figure 6.5 shows the behavior of the dynamics defined in 5.2.3 (top plot) and of its time average (bottom plot), both in terms of appreciation $x_v^{(i)}(t)$ (Figure 6.5a) and popularity $\pi^{(i)}(t)$ (Figure 6.5b).

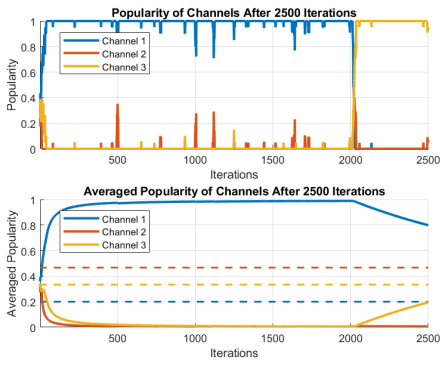
The plots are obtained considering a total number of iterations equal to 2500 and using an Erdos Renyi graph with a probability of having a link equal to 0.5. The initial $x_v^{(i)}(0)$-values are sampled from a uniform distribution between 0 and 1 $\forall v \in \mathcal{V}$, $\forall i \in \mathcal{I}$. The same is done for each $q^{(i)}$. The $\gamma_v \ \forall v$ values are randomly extracted from a uniform distribution defined between 0.0001 and 0.0009 while α_v from a uniform distribution between 0 and 0.5. The values of β_v are obtained from those of γ_v and $\alpha_v \ \forall v \in \mathcal{V}$ by requiring normalization ($\beta_v = 1 - \alpha_v - \gamma_v \ \forall v$).

As anticipated, the pattern is characterized by less frequent oscillations, both in terms of $\pi^{(i)}(t)$ and $x_v^{(i)}(t)$. The average dynamics show exactly what we expected, although there is a tendency to concentrate all the popularity on one of the channels, the presence of small $q^{(i)}$ generates an effect that leads the dynamics to remember the behavior of the general pattern 5.2.3, and thus to try to converge. Since this effect is too weak, however, the dynamics immediately return to concentrating popularity on one channel.

Again, the heatmap in Figure 6.6 allows us to make further considerations about this type of



(a) Evolution of $x_v^{(i)}(t)\ \forall v \in \mathcal{V}, \forall i \in \mathcal{I}$



(b) Evolution of $\pi^{(i)}(t)\ \forall i \in \mathcal{I}$

Figure 6.5: Dynamics of $x_v^{(i)}(t)$ and $\pi^{(i)}(t)$ from (6.1) obtained by requiring γ_v of the order of 10^{-4}.

behavior. Compared to section 6.2.2, we find a greater concentration of dynamics in configurations where the popularity is all focused on one channel: the yellow (more concentrated) squares are precisely those associated with such configurations. Once again, to move from one of these states to the others, configurations in which one of the channels is completely absent (edges of the simplex) are preferred.

What is very interesting to note is that when analyzing the values within the heatmap, the most visited configuration is not the one associated with channel 2 with the largest external factor $q^{(i)}$ (800 visits) but is the one in which all the popularity is associated with channel 1 (1400 visits). Obviously to give more precise information one would have to iterate the dynamics for longer and more times, but it is interesting to observe that for the same number of iterations with the section 6.2.2, the external factor also matters less in making one channel or another prevail.

6.3 Final Observations

Reducing the intensity of the external factor contribution has a clear impact on the model described in (5.2). The dynamics formally is characterized by an ergodic chain, but what happens, in reality, is that as $\gamma_v\ \forall v$ decreases, it exhibits behavior more and more similar to the limit case, which occurs when $\gamma_v = 0\ \forall v$. Since the latter case is characterized by absorbing states in which popularity is concentrated entirely on a single channel, decreasing range tends to concentrate dynamics for longer and longer in these configurations. After a while, it manages to find the strength to get out of it because although the influence of the external factor $q^{(i)}$ is small, nevertheless it is present and manages to give the input to get out of that state.

The simulations also show, that while for values of gamma that are not excessively small (Section 6.2.2), the dynamics choose which channel to focus on based on how much the value of its external factor $q^{(i)}$ is, by reducing gamma they choose more and more randomly which channel to focus on (Section 6.2.3), until they do so totally randomly when $\gamma_v = 0\ \forall v \in \mathcal{V}$ (Section 5.10).

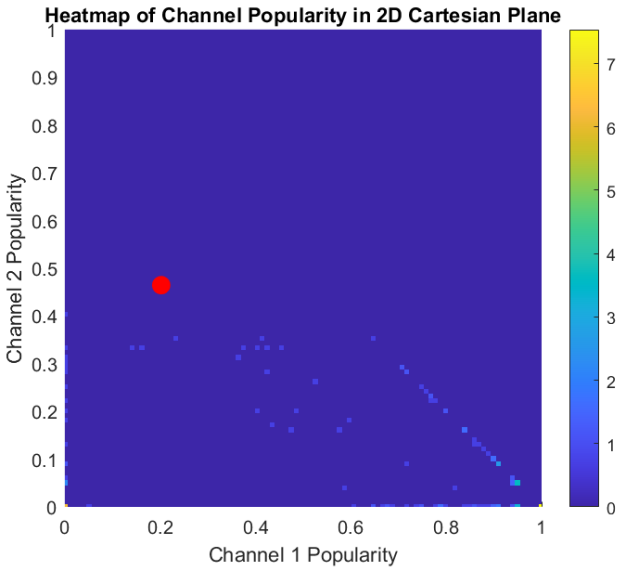


Figure 6.6: Heatmap for evolution of $\pi^{(i)}(t)$ from (6.1) associated to γ_v of the order of 10^{-4}.

Chapter 7

Concluding remarks and future developments

Nowadays, social networks are at the center of our lives, occupying numerous hours of our days. Given their increasingly unwieldy presence, it has become a key issue to understand how the popularity of influencers (creators/channels) evolves over time. Scientific research in this field is trying to define models that emphasize some characteristics over others: many of them adopt epidemiological models, others look for answers in statistics, and some use refined stochastic processes.

This master's thesis aimed to formalize a new model for understanding the evolution of content appreciation on social networks, and consequently to figure out the dynamics of popularity. The nonlinear model insists on three key terms that define popularity: user interactions and the exchange of impressions, role of recommendation platforms inherent to the social network, and possible other factors that depend on an external context.

We started our study with an empirical analysis of data associated with the social network YouTube, which we cleaned and interpreted in Python. We jammed the available data together in order to reconstruct metrics that would model user engagement, and thus the popularity of the various channels. We analyzed the trends, finding that the dynamics all show interesting convergence, and we wondered what potential external factors might or might not influence that kind of trends. At this stage, our analysis remains purely exploratory, aiming to understand the patterns that emerge from the data rather than establish definitive causal relationships. While certain external factors appear to correlate with variations in popularity, their exact role remains an open question, requiring further investigation with more targeted studies.

The theoretical analysis of the model was conducted in order to analyze the contribution of each of the terms listed above, and understand how strong their presence is in the evolution of the dynamics. The study initially focused on a deterministic dynamic of users' appreciation of content, which was shown to always converge, whatever the combinations of weights associated with each factor. While in some cases we fell back on already established models, our contribution was mainly related to the study of cases not known in the literature, which we have shown to converge at specific equilibrium points. Simulations in Matlab of each case analyzed allowed us to support the evidence of the theoretical results: in most cases, the long-term value of the popularity of a specific channel, depends on the external factors associated with it.

Such deterministic dynamics not only allowed us to highlight results not known in the literature but also to approximate the expectation of more general stochastic dynamics. The latter is

characterized by the presence of random variables that make it even more applicable to reality. Theoretical study using more sophisticated tools, such as mainly Markov chains, led us to derive conditions to ensure the ergodicity of the system and thus show convergence of the time average. Again, numerical simulations supported the theoretical formulation.

Our study stands as the starting point of a path aimed at understanding a new dynamic of popularity evolution. As such, the open points and future directions that can be taken are several. Starting from the current open points, the central object of study is to theoretically formalize the approximation between the analyzed deterministic dynamics and the expected stochastic dynamics, rigorously quantifying the committed error term. At the same time, the characterization of the equilibrium points of the general case of deterministic dynamics (Section 4.3.5) requires some extra effort to obtain a closed formula, which so far has only been deduced from simulations. Another interesting point is to do a concentration study for the particular cases studied in Chapter 6, to understand which channels the dynamics focus on the most. Future directions, in addition to closing these open points, could be more focused on further complicating the dynamics by including external factors that have time-dependent and user-dependent value, as well as imposing a more complex recommendation modeling than the one used in this work.

Chapter 8

Appendix

8.1 Theoretical Background

In this section we will recall several theorems and definitions useful for understanding the theory of the previous chapters.

8.1.1 Network Dynamics

Theorem 1. *(Perron-Frobenius Theorem) Given P a non-negative, rows-stochastic and irreducible square matrix, then:*

- $\exists\ \lambda_P = 1$, *where* λ_P *is an eigenvalue of* P*;*
- $\exists\ \nu^* \geq 0$ *s.t.* $\mathbb{1}'\nu^* = \mathbb{1}$ *and* $P'\nu^* = \nu^*$, *where* ν^* *is known as stationary probability distribution of* P*;*
- *every eigenvalue* μ *of* P *is characterized by* $|\mu| \leq 1$*;*
- *if* P *is s.t.* $\mathcal{G}_P$ *(the graph associated to the matrix P) is a strongly connected and aperiodic graph, then every eigenvalue* $\mu \neq 1$ *is s.t.* $|\mu| < 1$ *and* λ_P *is geometrically and algebraically simple.*

Lemma 6. *Consider a substochastic matrix* $A \in \mathbb{R}^{\mathcal{V}\times\mathcal{V}}$. *If in the graph associated to A there is a path from every node to a node linked to a row of A that sums to less than one, A is Schur Stable.*

8.1.2 Useful Definitions

Schur stable matrix

A matrix A is *Schur stable* matrix if its eigenvalues are located in the open unit disk in the complex plane.

Filtration $\mathcal{F}_t$

In the theory of stochastic processes, *filtrations* $\mathcal{F}_t$ are ordered collections of subsets that are used to model the information that is available at a given point.

8.1.3 Markov Theory

Markov Property in Discrete Time

Let $(X_n)_{n=0}^{\infty}$ be a discrete time process in a discrete space S.
$(X_n)_{n=0}^{\infty}$ is a *discrete time Markov Chain* (DTMC) if for any $y_j \in \{1, \ldots, S\}$, the Markov property holds:

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) =$$
$$\mathbb{P}(X_{n+1} = j \mid X_n = i) \stackrel{\text{def}}{=} p(i \mid j).$$

This condition states that the probabilities associated with a future state only depend upon the current state X_n, not on the process's full history.
The values $p(i \mid j)$ are called the (one-step) *transition probabilities*, and they form the *transition matrix*.
In a case where the transition probabilities are not time (n) dependent, the chain is called temporally homogeneous.

If the state space S is finite (say of cardinality k), the transition probabilities can be organized into a $k \times k$ matrix:

$$P = \begin{pmatrix} p(1 \mid 1) & p(1 \mid 2) & \cdots & p(1 \mid k) \\ p(2 \mid 1) & p(2 \mid 2) & \cdots & p(2 \mid k) \\ \vdots & \vdots & \ddots & \vdots \\ p(k \mid 1) & p(k \mid 2) & \cdots & p(k \mid k) \end{pmatrix}$$

with non-negative entries such that the row sums are 1:

$$\sum_{j \in S} P_{ij} = 1, \quad \forall i \in S.$$

Matrices of this kind are called stochastic matrices.
Exploiting the Markov property, to compute the probability of observing a trajectory $X_0, X_1, \ldots, X_m$ it is sufficient to know the transition matrix and the initial distribution $\pi(i) = \mathbb{P}(X_0 = i)$:

$$\mathbb{P}(X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) =$$
$$= p(i_{n-1} \mid i_n) p(i_{n-2} \mid i_{n-1}) p(i_0 \mid i_1) \mathbb{P}(X_0 = i_0).$$

Chapman-Kolmogorov equations

It is possible to generalize the previous theory to calculate transition probabilities in two steps, e.g.,

$$p^{(2)}(i \mid j) = \mathbb{P}(X_{n+2} = j \mid X_n = i).$$

Again, conditioning on $X_{n+1} = k$ and applying the law of total probability, the Markov property, and temporal homogeneity, we have:

$$p^{(2)}(i \mid j) = \mathbb{P}(X_{n+2} = j \mid X_n = i) =$$

$$= \sum_k \mathbb{P}(X_{n+2} = j \mid X_{n+1} = k, X_n = i)\mathbb{P}(X_{n+1} = k \mid X_n = i) =$$

$$= \sum_k p(i \mid k)p(k \mid j).$$

This formula has an interesting geometric interpretation, as the sum of the probabilities of all the paths that lead from i to j in two steps.

The two steps transition can also be expressed in terms of the transition matrix P:

$$P^{(2)} = P \cdot P = P^2.$$

More generally, the Chapman-Kolmogorov equations state:

$$P^{(m+n)} = P^{(m)} \cdot P^{(n)} = P^{m+n}.$$

Component-wise, this means:

$$P^{(m+n)}(i \mid j) = \sum_k P^{(m)}(i \mid k)P^{(n)}(k \mid j).$$

Classification of states and communication classes

We say that state j is *reachable* from i (and write $i \to j$) if there exists an m such that

$$p^{(m)}(i, j) > 0.$$

If both $i \to j$ and $j \to i$, we say that state i *communicates* with state j, and write $i \leftrightarrow j$.
A Markov chain is called *irreducible* if all the states communicate with each other. A subset $A \subseteq S$ of the state space is called *absorbing* (or a *trap*), if when the chain is started in A, there is no way to get out of A.

Let T_y be the time of the first visit to y, without counting X_0.

$$T_y = \min\{n \geq 1 : X_n = y\},$$

T_y is called the *hitting time* of y, and if the chain starts in $X_0 = y$, it is the *return time* to y. T_y is a random variable expressing how many steps are needed to visit y.

Let

$$x_y = \mathbb{P}_x(T_y < \infty) = \mathbb{P}(\text{we will visit } y \text{ again } \mid X_0 = x)$$

be the *probability of returning* to y in a finite time if we start at y.
There are two distinct types of states:

- y is *recurrent* if $y_y = 1$;
- y is *transient* if $y_y < 1$.

The random variable $N_n(y) = \sum_{i=1}^{n} \mathbf{1}(X_i = y)$ represents the number of visits to y before time n. We define $N(y) = \lim_{n\to\infty} N_n(y)$.
We can easily prove that

$$\mathbb{E}_x[N(y)] = \sum_{n=1}^{\infty} p^{(n)}(x, y),$$

indeed, the sequence $N_n(y)$ is almost surely increasing, and monotone convergence holds. Therefore,

$$\mathbb{E}_x[N(y)] = \sum_{n=1}^{\infty} \mathbb{P}_x(X_n = y) = \sum_{n=1}^{\infty} p^{(n)}(x, y).$$

It can be proved that y is recurrent if and only if

$$\mathbb{E}_y[N(y)] = \sum_{n=1}^{\infty} p^{(n)}(y, y) = \infty.$$

Recurrence is a class property, i.e., if $i \to j$ and i is recurrent, so is j.
Transience is also a class property, i.e., if $i \to j$ and i is transient, so is j.

A state i is *positive recurrent* if

$$\mathbb{E}_i(T_i) = \mathbb{E}(T_i \mid X_0 = i) < \infty.$$

It can be proved that positive recurrence is a class property, i.e., if $i \to j$ and i is positive recurrent, so is j.

A state y is *periodic* with period N if $p^{(k)}(y, y)$ can be strictly positive only if k is a multiple of N.

Stationary Distributions

A *stationary distribution* satisfies $\pi = \pi P$ requiring

$$\sum_i \pi(i) = 1.$$

This means the system will no longer change once this distribution is reached.
A Markov Chain (MC) is *positive recurrent* if all states are.

Theorem 2. *Suppose a Markov chain is irreducible and positive recurrent. Then there exists a stationary distribution π with $\pi(i) > 0$ for all $i \in S$.*

Theorem 3. *Suppose a Discrete-Time Markov Chain (DTMC) is irreducible, all states are aperiodic, and there is a stationary distribution π. Then, for all $i, j \in S$,*

$$\lim_{n\to\infty} \mathbb{P}(X_n = j \mid X_0 = i) = \lim_{n\to\infty} p^{(n)}(i, j) = \pi(j).$$

It means that the limit distribution coincides with the stationary distribution.

Theorem 4. *Suppose a DTMC is irreducible and recurrent (all states). Then, almost surely*

$$\lim_{n\to\infty} \frac{N_n(y)}{n} = \frac{1}{\mathbb{E}_y[T_y]}.$$

This is an interpretation of positive recurrence.

Theorem 5. *If a DTMC is irreducible and a stationary distribution exists, then*

$$\pi(y) = \frac{1}{\mathbb{E}_y[T_y]}.$$

It is unique, and the chain is positive recurrent.

Ergodic Theorem

Theorem 6. *Suppose a DTMC is irreducible and a stationary distribution exists. Assume that for some function f,*

$$\sum_x f(x)\pi(x) < \infty.$$

Then, almost surely,

$$\frac{1}{n}\sum_{m=1}^{n} f(X_m) \to \sum_x f(x)\pi(x),$$

i.e., time averages (left) equal space averages in the stationary regime (right).

The function $f(i)$ can be interpreted as the reward (or cost) you get for visiting i. The theorem is a generalization of the law of large numbers for discrete random variables when the random variables we are averaging are not necessarily independent.

8.1.4 Statistical Tools

Mean Squared Error

In statistics, the mean squared error (MSE) of an estimator measures the average of the squares of the errors—that is, the average squared difference between the estimated values and the true value. If a vector of n predictions is generated from a sample of n data points on all variables, and Y is the vector of observed values of the variable being predicted with $\hat{Y}$, then

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

Spearman's Rank Correlation Coefficient

In statistics, Spearman's rank correlation coefficient ρ is a nonparametric measure of rank correlation (statistical dependence between the rankings of two variables). It assesses how well the relationship between two variables can be described using a monotonic function. The Spearman correlation between two variables is equal to the Pearson correlation between the rank values of those two variables.

Given n raw scores (X_i, Y_i), converted to ranks $R[X_i], R[Y_i]$:

$$\rho = Corr(R[X], R[Y]) = \frac{Cov(R[X], R[Y])}{\sqrt{Var(R[X])Var(R[Y])}}.$$

8.2 Alternative Proofs

We include a theoretical proof of the approximation of the ratio analyzed in section 5.2.2.

Given the definition of $\pi^{(i)}(t)$, we need to calculate:

$$\mathbb{E}[\pi^{(i)}|\mathcal{F}_t] = \mathbb{E}\left[\frac{\sum_{v\in\mathcal{V}} Y_v^{(i)}(t)}{\sum_{j\in\mathcal{I}}\sum_{w\in\mathcal{V}} Y_w^{(j)}(t)}|\mathcal{F}_t\right],$$

which is the expected value of the ratio of two random variables. We are interested in understanding how this expectation differs from the ratio of expectations. Recalling that

$$\mathbb{E}[Y_v^{(i)}(t)|\mathcal{F}_t] = x_v^{(i)}(t),$$

we can recognize that at both numerator and denominator, there are generalizations of binomials, as sums of Bernoulli variables with different averages, also known as *Poisson Binomials*. Given that each variable $Y_v^{(i)}(t)\ \forall v \in \mathcal{V}, \forall i \in \mathcal{I}$ is independent of the others, we have

$$\mathbb{E}\left[\sum_{v\in\mathcal{V}} Y_v^{(i)}(t)\right] = \sum_{v\in\mathcal{V}} \mathbb{E}\left[Y_v^{(i)}(t)\right] = \sum_{v\in\mathcal{V}} x_v^{(i)}(t)$$

and

$$\mathbb{E}\left[\sum_{j\in\mathcal{I}}\sum_{w\in\mathcal{V}} Y_w^{(j)}(t)\right] = \sum_{j\in\mathcal{I}}\sum_{v\in\mathcal{V}} \mathbb{E}\left[Y_w^{(j)}(t)\right] = \sum_{j\in\mathcal{I}}\sum_{v\in\mathcal{V}} x_w^{(j)}(t).$$

Defining

$$\sum_{v\in\mathcal{V}} Y_v^{(i)}(t) := R,$$

$$\sum_{j\in\mathcal{I}}\sum_{w\in\mathcal{V}} Y_w^{(j)}(t) := S,$$

we can exploit the second order *Taylor expansion* of $f(R,S) = \frac{R}{S}$ around the point (μ_R, μ_S), where $\mu_R = \mathbb{E}[R]$ and $\mu_S = \mathbb{E}[S]$.

The second-order Taylor expansion of a function $f(R,S)$ around (μ_R, μ_S) is:

$$f(R,S) \approx f(\mu_R,\mu_S) + \left.\frac{\partial f}{\partial R}\right|_{(\mu_R,\mu_S)}(R-\mu_R) + \left.\frac{\partial f}{\partial S}\right|_{(\mu_R,\mu_S)}(S-\mu_S)$$

$$+\frac{1}{2}\left[\left.\frac{\partial^2 f}{\partial R^2}\right|_{(\mu_R,\mu_S)}(R-\mu_R)^2 + 2\left.\frac{\partial^2 f}{\partial R\partial S}\right|_{(\mu_R,\mu_S)}(R-\mu_R)(S-\mu_S)\right]$$

$$\frac{1}{2}\left[+\left.\frac{\partial^2 f}{\partial S^2}\right|_{(\mu_R,\mu_S)}(S-\mu_S)^2\right].$$

The derivatives of $f(R,S) = \frac{R}{S}$ are

- First derivatives:

$$\frac{\partial f}{\partial R} = \frac{1}{S}, \quad \frac{\partial f}{\partial S} = -\frac{R}{S^2}.$$

- Second derivatives:
$$\frac{\partial^2 f}{\partial R^2} = 0, \quad \frac{\partial^2 f}{\partial R \partial S} = -\frac{1}{S^2}, \quad \frac{\partial^2 f}{\partial S^2} = 2\frac{R}{S^3}.$$

Substituting the partial derivatives, we get:

$$f(R,S) \approx \frac{\mu_R}{\mu_S} + \frac{1}{\mu_S}(R-\mu_R) - \frac{\mu_R}{\mu_S^2}(S-\mu_S)$$
$$-\frac{1}{2}\left[\frac{2}{\mu_S^3}\mu_R(S-\mu_S)^2 + 2\frac{1}{\mu_S^2}(R-\mu_R)(S-\mu_S)\right].$$

Taking the expectation of both sides we have

$$\mathbb{E}\left[\frac{R}{S}\right] \approx \frac{\mu_R}{\mu_S} + \frac{1}{\mu_S}\mathbb{E}[R-\mu_R] - \frac{\mu_R}{\mu_S^2}\mathbb{E}[S-\mu_S]$$
$$-\frac{1}{2}\left[\frac{2}{\mu_S^3}\mu_R\mathbb{E}[(S-\mu_S)^2] + 2\frac{1}{\mu_S^2}\mathbb{E}[(R-\mu_R)(S-\mu_S)]\right].$$

Since $\mathbb{E}[R-\mu_R] = 0$ and $\mathbb{E}[S-\mu_S] = 0$, this simplifies to:

$$\mathbb{E}\left[\frac{R}{S}\right] \approx \frac{\mu_R}{\mu_S} - \frac{\mathrm{Cov}(R,S)}{\mu_S^2} + \frac{\mu_R \mathrm{Var}(S)}{\mu_S^3}. \tag{8.1}$$

For the calculation of the covariance, we can recall that

$$R = \sum_{v\in\mathcal{V}} Y_v^{(i)}, \quad S = \sum_{j\in\mathcal{I}}\sum_{w\in\mathcal{V}} Y_w^{(j)}.$$

The covariance is:

$$\mathrm{Cov}(R,S) = \mathbb{E}[RS] - \mathbb{E}[R]\mathbb{E}[S].$$

Recalling that in our case

$$\mathbb{E}[R|\mathcal{F}_t] = \mathbb{E}\left[\sum_{v\in\mathcal{V}} Y_v^{(i)}(t)|\mathcal{F}_t\right] = \sum_{v\in\mathcal{V}} \mathbb{E}[Y_v^{(i)}(t)|\mathcal{F}_t] = \sum_{v\in\mathcal{V}} x_v^{(i)}(t).$$

and

$$\mathbb{E}[S|\mathcal{F}_t] = \mathbb{E}\left[\sum_{i\in\mathcal{I}}\sum_{w\in\mathcal{V}} Y_w(i)(t)|\mathcal{F}_t\right] = \sum_{i\in\mathcal{I}}\sum_{w\in\mathcal{V}} \mathbb{E}[Y_w^{(i)}(t)|\mathcal{F}_t] = \sum_{i\in\mathcal{I}}\sum_{w\in\mathcal{V}} x_w^{(i)}(t).$$

Calculating $\mathbb{E}[RS]$

$$RS = \left(\sum_{v\in\mathcal{V}} Y_v^{(i)}(t)\right)\left(\sum_{i\in\mathcal{I}}\sum_{w\in\mathcal{V}} Y_w^{(i)}(t)\right) = \sum_{i\in\mathcal{I}}\sum_{v\in\mathcal{V}}\sum_{w\in\mathcal{V}} Y_v^{(i)}(t)Y_w^{(i)}(t).$$

We consider two cases:

- When $v = w$, we have:
$$\mathbb{E}[Y_v^{(i)}(t)Y_v^{(i)}(t)|\mathcal{F}_t] = \mathbb{E}[Y_v^{(i)}(t)|\mathcal{F}_t] = x_v^{(i)}(t).$$
- When $v \neq w$, since $Y_v^{(i)}$ and $Y_w^{(i)}$ are independent, we have:
$$\mathbb{E}[Y_v^{(i)}(t)Y_w^{(i)}(t)|\mathcal{F}_t] = \mathbb{E}[Y_v^{(i)}(t)|\mathcal{F}_t]\mathbb{E}[Y_w^{(i)}(t)|\mathcal{F}_t] = x_v^{(i)}(t)x_w^{(i)}(t).$$

Combining these two cases, we get:

$$\mathbb{E}[RS] = \sum_{i\in\mathcal{I}}\sum_{v\in\mathcal{V}} \mathbb{E}[Y_v^{(i)}(t)Y_v^{(i)}(t)] + \sum_{v\neq w}\sum_{i\in\mathcal{I}} \mathbb{E}[Y_v^{(i)}(t)Y_w^{(i)}(t)].$$

Thus:

$$\mathbb{E}[RS] = \sum_{i\in\mathcal{I}}\sum_{v\in\mathcal{V}} x_v^{(i)}(t) + \sum_{v\neq w}\sum_{i\in\mathcal{I}} x_v^{(i)}(t)x_w^{(i)}(t).$$

Now we can compute

$$\mathbb{E}[R]\mathbb{E}[S]$$

$$\mathbb{E}[R] = \sum_{v\in\mathcal{V}} x_v^{(i)}(t), \quad \mathbb{E}[S] = \sum_{i\in\mathcal{I}}\sum_{w\in\mathcal{V}} x_w^{(i)}(t).$$

Thus:

$$\mathbb{E}[R]\mathbb{E}[S] = \left(\sum_{v\in\mathcal{V}} x_v^{(i)}(t)\right)\left(\sum_{i\in\mathcal{I}}\sum_{w\in\mathcal{V}} x_w^{(i)}(t)\right).$$

Expanding:

$$\mathbb{E}[R]\mathbb{E}[S] = \sum_{i\in\mathcal{I}}\sum_{v\in\mathcal{V}} x_v^{(i)}(t) + \sum_{v\neq w}\sum_{i\in\mathcal{I}} x_v^{(i)}(t)x_w^{(i)}(t).$$

The two terms are the same so the covariance is 0.

The formula (8.1) reduces to

$$\mathbb{E}\left[\frac{R}{S}\right] \approx \frac{\mu_R}{\mu_S} + \frac{\mu_R \mathrm{Var}(S)}{\mu_S^3}.$$

Estimating the error committed by the approximation in our case is complex, given the structure of Binomial Poissons and requires finer tools. So we leave a more thorough analysis to future work.

Bibliography

[1] P. Covington, J. Adams, and E. Sargin, "Deep neural networks for youtube recommendations," *Proceedings of the 10th ACM Conference on Recommender Systems*, pp. 191–198, 2016.

[2] T. Vial, "Rapport de stage," *Stage effectué au sein du laboratoire du GIPSA-lab*, 2024.

[3] W. O. Kermack, A. G. McKendrick, and G. T. Walker, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. vol. 115, pp. 700–721, Aug. 1927.

[4] D. Daley and D. G. Kendall, "Epidemics and rumours," *Nature*, vol. vol. 204, 1964.

[5] L. M. Bettencourt, A. Cintrón-Arias, D. I. Kaiser, and C. Castillo-Chávez, "The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models," *Physica A: Statistical Mechanics and its Applications*, vol. 364, p. 513–536, May 2006.

[6] F. Jin, E. Dougherty, P. Saraf, Y. Cao, and N. Ramakrishnan, "Epidemiological modeling of news and rumors on twitter," vol. Proceedings of the 7th Workshop on Social Network Mining and Analysis, p. 9, 2013.

[7] F. Xiong, Y. Liu, Z. jiang Zhang, J. Zhu, and Y. Zhang, "An information diffusion model based on retweeting mechanism for online social media," *Physics Letters A*, vol. 376, no. 30, pp. 2103–2108, 2012.

[8] C. Richier, E. Altman, R. Elazouzi, T. Jimenez, G. Linares, and Y. Portilla, "Bio-inspired models for characterizing youtube viewcout," in *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, pp. 297–305, 2014.

[9] R. Sachak-Patwa, N. T. Fadai, and R. A. V. Gorder, "Understanding viral video dynamics through an epidemic modelling approach," *Physica A: Statistical Mechanics and its Applications*, vol. 502, pp. 416–435, 2018.

[10] R. Crane and D. Sornette, "Robust dynamic classes revealed by measuring the response function of a social system," *Proceedings of the National Academy of Sciences*, vol. 105, p. 15649–15653, Oct. 2008.

[11] F. M. Bass, "A new product growth for model consumer durables," *Management Science*, vol. 15, no. 5, pp. 215–227, 1969.

[12] X. Gao, Z. Zheng, Q. Chu, S. Tang, G. Chen, and Q. Deng, "Popularity prediction for single tweet based on heterogeneous bass model," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 5, pp. 2165–2178, 2021.

[13] T.-A. Hoang and E.-P. Lim, "Virality and susceptibility in information diffusions," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 6, pp. 146–153, Aug. 2021.

[14] D. Luu, E.-P. Lim, T. Hoàng, and F. Chua, "Modeling diffusion in social networks using network properties," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 6, 01 2012.

[15] M. Kim, D. Newth, and P. Christen, "Modeling dynamics of diffusion across heterogeneous social networks: News diffusion in social media," *entropy*, vol. 15, pp. 4215–4242, 10 2013.

[16] Q. Cao, H. Shen, J. Gao, B. Wei, and X. Cheng, "Popularity prediction on social platforms with coupled graph neural networks," 2019.

[17] R. Jin, T. Murata, and X. Liu, "Predicting popularity trend in social media networks with multi-layer temporal graph neural networks," *Complex & Intelligent Systems*, vol. 10, pp. 1–17, 04 2024.

[18] Y. Shang, B. Zhou, X. Zeng, Y. Wang, H. Yu, and Z. Zhang, "Predicting the popularity of online content by modeling the social influence and homophily features," *Frontiers in Physics*, vol. 10, 2022.

[19] M. Degroot, "Reaching a consensus," *J. Am. Stat. Assoc. 69 (345)*, vol. 16, pp. 118–121, 1974.

[20] N. Friedkin and E. Johnsen, "Social influence networks and opinion change," *Advances in Group Processes*, vol. 16, 01 1999.

[21] R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, vol. 5, 07 2002.

[22] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, "Mixing beliefs among interacting agents," *Advances in Complex Systems*, vol. 3, pp. 87–98, 01 2000.

[23] R. A. Holley and T. M. Liggett, "Ergodic theorems for weakly interacting infinite systems and the voter model," *Ann. Probab.*, vol. 3, no. 6, pp. 643–663, 1975.

[24] K. Sznajd-Weron, "Sznajd model and its applications," *Acta Physica Polonica B*, vol. 36, 05 2005.

[25] G. M. Insight, "Youtube statistics 2024 (demographics, users by country and more)." https://www.globalmediainsight.com/blog/youtube-users-statistics/#Monthly_Active_Users_on_YouTube, 2024. Accessed: 29-12-2024.

[26] G. Stocking, P. van Kessel, M. Barthel, K. Eva Matsa, and M. Khuzam, "Many americans get news on youtube, where news organizations and independent producers thrive side by side," *Pew Research Center*, pp. 31–97, 2020.

[27] M. Adavelli, "How many videos are uploaded on youtube every day?." https://techjury.net/video/how-many-videos-are-uploaded-to-youtube-a-day/#:~:text=YouTube%20stands%20as%20the%20top,engaging%20videos%20for%20their%20subscribers., 2023. Accessed: 29-12-2024.

[28] M. Castaldo, *Attention dynamics on YouTube: conceptual models, temporal analysis of engagement metrics, fake views.* Automatic. Université Grenoble Alpes, 2022.

[29] Z. Zhao, L. Hong, L. Wei, J. Chen, A. Nath, S. Andrews, A. Kumthekar, M. Sathiamoorthy, X. Yi, and E. Chi, "Recommending what video to watch next: A multitask ranking system," *Proceedings of the 13th ACM Conference on Recommender Systems*, pp. 43–51, 2019.

[30] S. Suresh, M. Srivastava, and Mohana, "Recommender systems: An overview research trends and future directions," *International Journal of Business and Systems Research*, pp. 14–52, 2021.

[31] D. Massimo and F. Ricci, "Popularity, novelty and relevance in point of interest recommendation: an experimental analysis," *Inf Technol Tourism*, vol. 23, pp. 473–508, 10 2021.

[32] K. Srikala, "Popularity based recommendation system," *International Journal of Engineering and Advanced Technology*, vol. 9, pp. 1561–1566, 02 2020.

[33] S. Lin, C. Gao, J. Chen, S. Zhou, B. Hu, Y. Feng, C. Chen, and C. Wang, "How do recommendation models amplify popularity bias? an analysis from the spectral perspective," *Proceedings of the Eighteenth ACM International Conference on Web Search and Data Mining (WSDM '25)*, p. 14 pages, 11 2024.

[34] Y. Bajdadi, "Rapport de stage: Temporal analysis of publication practices on youtube," *Stage effectué au sein du laboratoire du GIPSA-lab*, 2023.
[35] G. Chatzopoulou, C. Sheng, and M. Faloutsos, "A first step towards understanding popularity in youtube," *2010 INFOCOM IEEE Conference on Computer Communications Workshops*, pp. 1–6, 2010.
[36] A. M. Möller, S. E. B. R. Kühne, and J. Peter, "Exploring user responses to entertainment and political videos: An automated content analysis of youtube," *Social Science Computer Review*, pp. 510–528, 2018.
[37] F. de Figueiredo, F. Benevenuto, and J. M. Almeida, "The tube over time: characterizing popularity growth of youtube videos," *Web Search and Data Mining*, 2011.
[38] L. A. Liikkanen, "Three metrics for measuring user engagement with online media and a youtube case study," 2014.
[39] D. J. Welbourne and W. J. Grant, "Science communication on youtube: Factors that affect channel and video popularity," *Public Understanding of Science*, vol. 25, no. 6, pp. 706–718, 2016. PMID: 25698225.
[40] S. Yang, D. Brossard, D. Scheufele, and M. Xenos, "The science of youtube: What factors influence user engagement with online science videos?," *PLOS ONE*, vol. 17, p. e0267697, 05 2022.
[41] M. Bärtl, "Youtube channels, uploads and views: A statistical analysis of the past 10 years," *Convergence*, vol. 24, no. 1, pp. 16–32, 2018.
[42] F. Gantmacher, *The theory of matrices*. AMS Chelsea Publishing, 2000.
[43] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. part I," *CoRR*, vol. abs/1701.06307, 2017.
[44] M. Artzrouni, "On the convergence of infinite products of matrices," *Linear Algebra and its Applications*, vol. 74, pp. 11–21, 1986.
[45] G. Stewart, "On the powers of a matrix with perturbations," *Numerische Mathematik*, vol. 96, 2003.
[46] R. Durrett, *Probability: Theory and Examples*. Cambridge University Press, 5th ed., 2019.
[47] E. Bibbona, A. Masi, and M. Zanella, *Markov Chains and Stochastic Processes*. Springer, 2019.
[48] R. J. Brodie, A. Ilic, B. Juric, and L. Hollebeek, "Consumer engagement in a virtual brand community: An exploratory analysis," *Journal of Business Research*, vol. 66, no. 1, pp. 105–114, 2013. (1)Thought leadership in brand management(2)Health Marketing.