Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

SYMMETRIZING DYNAMICS: FROM CLASSICAL TO QUANTUM APPLICATIONS

Ph.D. candidate
Luca Mazzarella

Advisor
Prof. Francesco Ticozzi

School Director
Prof. Matteo Bertocco

Coordinator
Prof. Carlo Ferrari

XXVI Series
Ph.D. School in Information Engineering
Department of Information Engineering
University of Padova, 2014

Abstract

Over the last few decades a deep and systematic understanding of the dynamical behaviour of networked systems have been in the spotlight of a wide range of disciplines, including biology, applied mathematics, social science, electrical engineering and physics. One of the most intriguing features, both from the theoretical and applicative point of view, is the emergence of collective phenomena from topology-constrained interactions among the constituent parts of the networks.

Driven by the potential advantages promised by quantum computers for certain computational problems, the recent advances in the capability of modeling and control of complex quantum systems, such as improvements in the generation and coherent manipulation of individual systems, have opened new research directions towards "distributed" quantum information applications.

A variety of mathematical approaches has been employed to describe complex systems in different fields. For information processing applications a particularly rewarding one consists in modelling a networked system in the so-called multi-agent prospective. In this picture, the system is modelled as an ensemble of components, named agents, each one being assigned with an internal state, whose value evolves as result of the information exchange among the agents. As a result of the network topology, their capability of interaction is usually constrained to be local: for example, we could consider exchange of energy or information among neighbouring agents. The evolution law of the whole system inherits such locality constraints. For this class of models, a typical problem of interest is the characterization of the asymptotic behaviour of the system.

Among the issues regarding networked systems, the "consensus problem" and the related algorithms have received a significant share of attention during the last ten years. In this problem the network agents asymptotically have to attain agreement on the value of some objective variable under local communication constraints. A number of algorithms have been developed to address this problem, among which the celebrated gossip algorithm. The latter relies on switching dynamics and, under rather weak assumptions, exhibits robust convergence under variations in the interaction constraints, i.e. the network topology.

In this dissertation we reinterpret the goal of the consensus problem as a symmetrization problem, and we address it by a switching-type dynamics based on convex combinations of actions of a finite group. In order to study the convergence of our class of algorithms we lift the dynamics to an abstract, group-theoretic level that allows us to derive general conditions for convergence. Such conditions, in fact, are independent of the particular group action, and focus only on the group itself and the way the iterations are selected. Convergence is guaranteed provided that some mild assumptions on the selection rule for the iterations are fulfilled. Furthermore, this class of algorithms retains the robustness features and unsupervised character of the consensus algorithm.

Our reformulation allows to devise algorithms for applications as diverse as randomized discrete Fourier transform and random state generation. We pose a special emphasis on the extension of the consensus problem to the quantum domain. In this setting we highlight how, due to the richer mathematical structure over which the internal state is encoded, the definition of the consensus goal admits various extensions, each of them exhibiting different features. We also propose a suitable dissipative dynamics enacting the symmetrizing gossip interactions and then use our general result on convergence to prove it ensures asymptotic convergence.

Beside the technical results, one of the main contributions of our work is a new, generalized viewpoint on consensus, which allows us to extend the robustness of consensus-inspired algorithms to new problems in apparently unrelated fields. This reinforces the role of consensus algorithms as fundamental tools for distributed computing, both in the classical and the quantum setting.

Sommario

Negli ultimi decenni l'approfondito studio delle dinamiche su network è stato al centro dell'indagine in molti discipline tra cui, biologia, matematica applicata, scienze sociali, ingegneria dei sistemi e non ultima la fisica. Una delle caratteristiche più interessanti, sia da un punto di vista teorico che applicativo, è l'emergere di dinamiche collettive indotte dalle interazioni tra i costituenti del network nonostante i vincoli imposti dalla struttura topologica dello stesso.

Spinti dai potenziali vantaggi promessi dalla quantum computation nella soluzione di certi problemi, i recenti avanzamenti nella capacitá di descrivere e controllare sistemi complessi, come per esempio i miglioramenti nella generazione e manipolazione correnti di sistemi singoli, hanno aperto nuove direzioni di ricerca verso applicazioni di quantum information "distribuita".

In campi differenti sono stati sviluppati molteplici approcci per descrivere sistemi complessi. Tuttavia, per il campo dell'information processing è particolarmente adeguata la cosiddetta prospettiva multi-agente. In questa visuale, il sistema è descritto come un insieme di componenti, dette agenti, ognuna delle quali possiede uno stato interno, il cui valore evolve come risultato dello scambio di informazione ed energia tra gli agenti stessi. A causa della topologia del network, la capacità di interazione degli agenti è vincolata ad essere locale: per esempio, assumiamo che la comunicazione possa avvenire solo tra agenti prossimi. Le leggi di evoluzione dell'intero sistema ereditano tale vincolo di località. Per queste classi di modelli un tipico problema di interesse è caratterizzarne il comportamento asintotico.

Tra gli aspetti di interesse riguardanti i network, il "problema di consensus" a i relativi algoritmi hanno ricevuto grande attenzione nell'ultimo decennio. In

questo problema gli agenti del network devono asintoticamente accordarsi sul valore di una qualche variabile di interesse sotto il vincolo di comunicazioni locali. Un grande numero di algoritmi sono stati sviluppati per affrontare tale problema, tra i quali il celebrato algoritmo di gossip. Quest'ultimo si basa su una dinamica di tipo switching ed esibisce una convergenza robusta rispetto alla variazione dei vincoli di interazione quali ad esempio la topologia del network.

In questa tesi reinterpretiamo gli obiettivi del problema di consensus come un problema di simmetrizzazione che affrontiamo mediante dinamiche di tipo switching basate sulla combinazione convessa di azioni di un gruppo finito.

Per descrivere la convergenza di tali algoritmi proponiamo una descrizione più astratta di stampo gruppale. Tale descrizione ci permette di formulare criteri generali per la convergenza, indipendenti dalla particolare azione, che si focalizzano solo sul gruppo in quesitone e sul modo in cui le iterazioni sono generate. La convergenza viene garantita a patto che i meccanismi di selezione delle iterazioni rispettino alcuni criteri poco stringenti. Inoltre, la nostra classe di algoritmi mantiene le caratteristiche di robustezza degli algoritmi di gossip.

La nostra riformulazione ci permette di considerare algoritmi per diverse applicazioni quali ad esempio la trasformata di Fourier distribuita e la generazione di stati casuali. Inoltre, descriviamo approfonditamente l'estensione quantistica del problema del consensus. In quest'ambito mostriamo come, a causa della ricchezza delle strutture matematiche con cui gli stati interni sono descritti, la definizione dell'obiettivo di consensus ammetta diverse estensioni ognuna recanti diverse caratteristiche. Proponiamo, inoltre una dinamica dissipativa che asintoticamente realizza tale simmetrizzazione.

Oltre a risultati di tipo tecnico, uno dei contributi centrali del lavoro è un nuovo e generalizzato punto di vista sul consensus che permette di estendere la robustezza di tali algoritmi a problemi a prima vista scollegati, rinforzando così il ruolo di tali algoritmi come strumento per il calcolo distribuito sia in ambito classico che quantistico.

Citations to Previously Published Work

Chapters 2, 3 and 4 are largely based on the article preprint:

[1] L. Mazzarella, F. Ticozzi & A. Sarlette. From Consensus to Robust Randomized Algorithms: A Symmetrization Approach. arXiv:1311.3364v2.

Chapters 6, 7 and 8 are largely based on the journal paper:

[2] L. Mazzarella, F. Ticozzi and A. Sarlette. Consensus for Quantum Networks: From Symmetry to Gossip Interactions, accepted for publication in *IEEE Transactions on Automatic Control*. arXiv:1303.4077

and the conference proceeding:

[3] L. Mazzarella, A. Sarlette & F. Ticozzi. A new perspective on gossip iterations: From Symmetrization to quantum consensus, *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference*, 250-255, 10-13 Dec. 2013.

Acknowledgments

My doctoral studies have been a great opportunity of growth both academically and as a person. Therefore, I feel the nice obligation to thank anyone who has played whatever role in it.

First and foremost, I want to express my sincerest gratitude and estimation to my advisor Prof. Francesco Ticozzi for guiding me through the winding path of research with endless patience, always encouraging me to give my best in every situation. His advices and contagious enthusiasm have been fundamental during this journey.

Next I want thank Prof. Alain Sarlette. His precious contributions and invaluable though-stimulating discussions helped me to improve my understanding of the consensus problem.

I also want like to thank all the "Automatica" group for their warm welcome and especially Prof. Augusto Ferrante for his constant encouragement and care.

I thank Prof. Gerard Milburn for giving me the unique opportunity to visit the EQuS center in Birsbane and Dr. Casey Meyer for all our stimulating discussions on quantum transport.

I thank all the Quantum Future staff, it has been an honor for me being part of this project. A special thank goes to Prof. Paolo Villoresi, Prof. Giuseppe Vallone, Prof. Alexander Sergienko and Prof. Nicola Laurenti for their collaboration, all the valuable conversations (not only about research) and for all the nice time at the various "quantum lunch" and "quantum fish".

I also want to thank my present and former mates at the DEI: Chiara, Francesco, Mattia Z., Saverio, Damiano, Giulio, Gian Antonio, Fabio, Mattia B., Matteo, Alberto, Nicola and Davide for sharing with me the "glory and

(the very few) misery" of the graduate student life.

Thanks Ginaluca, Giovanni, Maurizio, Mattia, Marina, Paolo R., Stefano, Serena, Claudio, Minos, Simone, Francesco, Giulia, Paolo T. and Giorgia. I hope you will always be part of my life.

Thanks to Cristina for constantly being on my side and making my life a deal better.

Last but definitely not least, I want to thank my parents, Muarizio and Nives, for their unconditional love and support.

Structure of the Thesis

The topics covered in this dissertation are at the boundary between quantum theory and automatic & control theory. The dissertation is divided into two parts. In the first part we reformulate the classic consensus problem as a switching dynamics leading to symmetrization with respect to the action of a suitable finite group. We show how this symmetrization framework, in addition to being a new view point for the gossip problem, directly extends the desirable features of consensus-like algorithms, such as robustness, to new control problems. More precisely, in Chapter 1 we introduce the consensus problem in the operational multi-agent picture and then we focus on the gossip algorithm recalling the result that characterize its convergence features. In Chapter 2 we begin to establish the symmetrizing framework by showing that the gossip interaction can be written as a suitable convex combination of action of the permutation group and by characterizing the consensus situation as invariance with respect the action of the group. In Chapter 3 we develop a superior point of view for the symmetrizing picture that allows us to study the convergence of our class of algorithms only in terms of the group at hand and of the switching signals. Through these Chapters the gossip algorithm it used as running example. Chapter 4 is devoted to the applications of our framework to various problem such as, the consensus for probability distribution, the generation of random state from a set and the randomized computation of the discrete Fourier transform. The robustness features of the algorithms are highlighted throughout this chapter.

In the second part we develop the symmetrizing picture in the quantum domain, here we give independent proofs of many of the results established in the first part using the language and tools of quantum theory. In Chapter 5 we

review the framework that is needed to model a finite dimensional quantum systems. In view of the structure of quantum states, in Chapter 6 we characterize a hierarchy of different consensus situations for a network of quantum systems while in Chapter 7 we built the evolution for the quantum network keeping into account the locality constraints and prove the convergence properties. Finally, Chapter 8 covers a wide range of applications, some of them being purely quantum such as the dynamical decoupling and the purification and cooling of a sample, some being the extension of those presented in 4 such as the symmetrization of probability distribution of non commuting random variables and the generation of random states for a quantum system. For this latter application, we present a proof of principle of an almost quadratic speed-up with respect to the classical case present in Chapter 4. In Appendix A we fix the notation and recover some fundamental results about theory and representation for finite groups and about graphs theory while Appendix B is devoted to the proofs of some results presented in Chapter 7.

Contents

1 Classical conensus algorithm

1.1 Consensus Problem

1.1.1 Introduction

Among the recent trends in control and systems theory, the field of distributed control, estimation and optimization on networks has stimulated an impressive amount of research. Situations where the agents of a networked system have to autonomously reach agreement on some global piece of information by communicating with only a limited number of peers are ubiquitous in a variety of contexts, such as: sensor network [1, 2, 3], control of mobile agents [4, 5, 6, 7] and load balancing problems [8, 9, 10].

The information on which we want to reach agreement or *consensus* is linked, depending on the particular application, to some global property of the network and is achieved through a distributed algorithm (or dynamics) that usually has to fulfill the following requirements [11, 12, 13, 14]:

- there is no centralized entity or super agent that coordinates the computation, for example it is usually unreasonable to assume that a single nodes can get access instantaneously to the global information of the network. In this sense the network is said to be *unsupervised*.
- The topology of the network might depend on time, in the sense that even the set of agent and their communication capability might be time-dependent.
- Finally the agents are capable to share data with few "neighbors" among the network. The notion of proximity is characterized by the (instantaneous) network topology.

Furthermore, additional restrictions might apply in situations, for example in application involving wireless networks an agent cannot simultaneously communicates with more than one agent while in other such as process networks, data cannot be simultaneously transmitted to more than an agent.

Motivated by this set of constraints, several consensus algorithms have been formulated in order to distribute, in an unsupervised fashion, the computational burden across the network. They differentiate on the interaction scheme they use, on their robustness against the change in the network topology and on their evolution that can be deterministic or random, synchronous or asynchronous.

In this section we begin formalizing the theoretical framework for the consensus problem by analyzing a particular instance, the average consensus problem, in which the goal is to estimate the average of some global quantity across the network. We then consider a particular algorithm, the celebrated *gossip algorithm* [12, 13], and recall under which conditions it converges asymptotically to the average of global quantity of interest.

In this section we are going to employing concepts and notations from graph theory that are reviewed in Appendix A.1.

1.1.2 Average Consensus Problem

Consensus problems for a networked system, in the so called *operational multi-agent picture*, are typically formulated along the following lines.

- In order to model the network and the communication capability of its agents it is introduced a graph $G(V, E)$ whose set of nodes V is in one to one correspondence with the agents and the set of link $E(t)$ (possibly time-dependent) precise which agent can exchange information at each instant time, fig. 1.1 . Furthermore to each agent is associated an internal state $x_k \in \mathbb{R}^n$. In what follows we are always going to assume the graph to be undirected (see section A.1) and the set of vertices to be time independent.

- An *interaction protocol* that specifies how the information shared by the is processed at each instant of time, usually the most significant constraints for this ingredient are the *locality constraints*, i.e. the capability of the agent of communicating with only a restricted number of *neigbours*. A

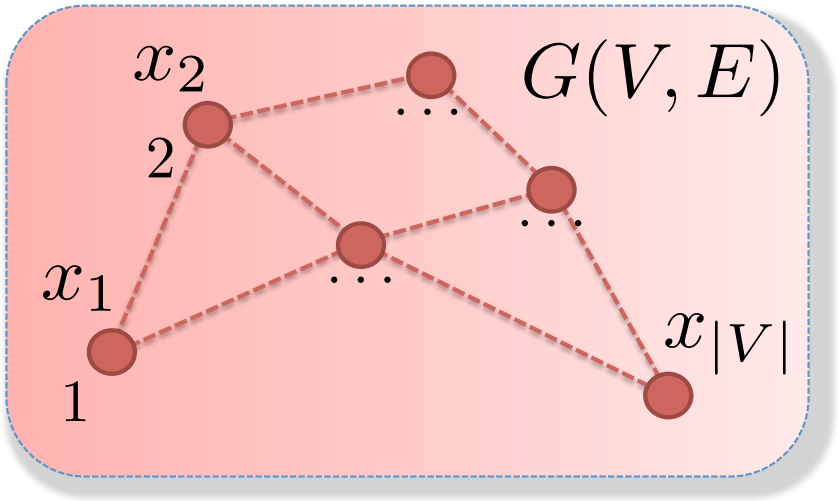

Figure 1.1: A graph comprising of a set of vertices V and a set of edges E.

review of interaction protocol and their feature is postponed to the next section.

- A *consensus situation* that specifies the computation goal. More precisely, let us collect all the internal states in a single vector representing the *global state* $x = (x_1(0), \ldots, x_{|V|}(0))^T$. We are going to consider consensus situation of the following type:

  Definition 1 (Consenus)**.** We say that a state is in *consensus* if it belongs to the set:

  $$\mathcal{C} = \{(\bar{x}_1, \ldots, \bar{x}_{|V|})^T \in \mathbb{R}^{n|V|} \ : \ \bar{x}_j = \bar{x}_k \, \forall \, j, k \ \in \ V\}, \tag{1.1}$$

  Where the components of $\bar{x}$ reflect some function of the initial condition $x(0) = (x_1(0), \ldots, x_{|V|}(0))^T$.

  We will refer to $\mathcal{C}$ as the *consenus set*. Note that, at the end of the computation, each agent has an estimation of the global quantity of interest. This feature is particularly desirable if only the agents of a subset of the network are accessible for the read-out.

For example, if we are interested in the computation of the average we have that the consensus set becomes:

$$\mathcal{C}^{av} = \{\bar{x} \in \mathbb{R}^{nm} \ : \ \bar{x}_k = \tfrac{1}{|V|} \sum_{j \in V} x_j(0) \quad \forall \ k \ \in \ V\} \tag{1.2}$$

this particular instance of the consensus problem is called *average consensus problem.*

1.1.3 The Interaction Protocol

I this section we present the design of dynamical systems or algorithms that by iterating (in a deterministic of random fashion) some local interaction protocols drives the systems toward a consensus state.

As we have anticipated many consensus algorithms have been proposed in the literature each one with different feature motivated by the particular constraints of the problem at hand.

In most of the literature are considerate algorithms were each node runs a first order (linear) dynamical system to update its estimation of the target variable. The systems are coupled according to the available communication capability. For sake of simplicity let us consider an unweighted and undirected graph whose set of edges is time independent. A single iteration of a consensus algorithm take the name of *consensus interaction* and can be written as follows:

$$x_i(t+1) := x_i(t) + u_i(t). \tag{1.3}$$

The function $u_i(t)$ is called *input function* and depend on the particular interaction. Has we have anticipated we aim to focus on interaction that are locally constrained hence the input function is built with the information coming form a subset of neighbor of i. More precisely, let $\mathcal{N}_i$ be the set of neighbors of i and $\mathcal{B}_i(t) \subseteq \mathcal{N}_i$, $u_i(t)$ can be written as:

$$u_i(t) = \sum_{j \in \mathcal{B}_i(t) \subseteq \mathcal{N}_i} f(x_i, x_j) \tag{1.4}$$

with $f(x,y) = -f(y,x)$. More generally, locality can be defined with *quasi-local operators*. Instead of considering a graph, we define a set of *neighborhoods* $\mathcal{N}_j \subseteq \{1,\dots,m\}$ for $j = 1,\dots,M$, and a quasi-local operator is one that leaves all subsystems unchanged except those of one neighborhood $\mathcal{N}_j$. This neighborhood notation is not customary for the consensus problem. In section 7.1 we will see how it will come at hand in characterizing consensus in the quantum domain. A dynamics with local coupling would write

$$x(t+1) = \textstyle\sum_{j=1}^{M} V_j(t)\,(x(t)) = P(t)x(t), \tag{1.5}$$

where the $V_j(t)$ are quasi-local operator acting on the neighborhoods $\mathcal{N}_j$.

In what follows we are going to assume $P(t)$ to be either a deterministic sequence or a sequence of i.i.d. matrix-valued random variables. In the latter case, $x(t)$ is a stochastic process itself. We have the following definitions for a deterministic algorithm that asymptotically realizes consensus.

Definition 2 (Deterministic case)**.** The algorithm generated by the deterministic sequence $P(t)$ asymptotically realizes consensus if for every $x(0) \in \mathbb{R}^{nm}$ the following conditions are fulfilled:

1. if $x(0) \in \mathcal{C}$ then $x(t) \in \mathcal{C}$,

2. There exist a real constant α for which:

$$\lim_{t\to\infty} x(t) = \alpha\mathbf{1} \tag{1.6}$$

where $\mathbf{1}$ is nm-dimensional row vectors $\mathbf{1} = (\underbrace{1,\ldots,1}_{n|m|})^T$.

If $P(t)$ is a sequence of random variables we speak of *probabilistic consensus.*

Definition 3 (Random case)**.** The algorithm generated by the sequence of stochastic matrices $P(t)$ asymptotically realizes probabilistic consensus if for every $x(0) \in \mathbb{R}^{n|m|}$ the following conditions are fulfilled:

1′. if $x(0) \in \mathcal{C}$ then $x(t) \in \mathcal{C}$,

2′. There exist a real constant α for which almost surely:

$$\lim_{t\to\infty} x(t) = \alpha\mathbf{1} \tag{1.7}$$

where $\mathbf{1}$ is nm-dimensional row vectors $\mathbf{1} = (\underbrace{1,\ldots,1}_{n|m|})^T$.

It is worth noting that both the definitions require that the consensus situation is defined through an equilibrium for the dynamical systems, in Definition 8 of section 2.1 we will generalize the definition of asymptotic consensus in a way that does not imply convergence to an equilibrium but requires that the dynamics is confined in the generalized consensus set.

It is easy to see that in order to achieve (probabilistic) average consensus case the constant α must be (almost surely) equal to:

$$\alpha = |V|^{-1}\mathbf{1}^T x(0). \tag{1.8}$$

Let us now draw the readers attention on the fact that if $P(t)$ represents a sequence of stochastic matrices, i.e. if for every t:

$$P_{i,j}(t) \geqslant 0 \;\forall i,j \in V \quad \sum_{j=1}^{|V|} P_{i,j}(t) = 1 \quad \forall i \in V, \tag{1.9}$$

it is apparent to see that we have that:

$$P(t)\mathbf{1} = \mathbf{1} \quad \forall\, t. \tag{1.10}$$

In this case the first condition for consensus is satisfied because a consensus state is an equilibrium for a stochastic matrix for every values of α.

Furthermore, if the sequence $P(t)$ is also doubly stochastic, i.e. it is stochastic and in addition:

$$\sum_{i=1}^{|V|} P_{i,j}(t) = 1 \quad \forall\, j \in V, \tag{1.11}$$

we have also that:

$$\mathbf{1}^T = \mathbf{1}^T P(t) \quad \forall\, t, \tag{1.12}$$

this implies that the average is left invariant by the dynamics for every t, in fact:

$$|V|^{-1}\mathbf{1}^T x(t+1) = |V|^{-1}\mathbf{1}^T P(t)x(t) = |V|^{-1}\mathbf{1}^T x(t) = \cdots = |V|^{-1}\mathbf{1}^T x(0). \tag{1.13}$$

Hence, if either condition (2) or condition (2′) are fulfilled we have that α:

$$\alpha = |V|^{-1}\mathbf{1}^T x(0), \tag{1.14}$$

thus ensuring average consensus.

It is now apparent that designing a sequence of interactions such that the dynamics converge toward a consensus state is the key problem in this framework. This issue has been widely studied in the literature [15, 12, 13].

1.1.4 Gossip Consensus

We now present the celebrated *gossip algorithm* an algorithm designed to asymptotically compute the average of the internal states across the network. Let us consider an *undirected graph* and let us assume that at every instant an edge of the graph is selected according to some mechanism that can be deterministic or random, then the linked agents share their information and updated their estimation of the average by computing the weighted average, without loss of generality we consider scalar internal variables, i.e. $x_k(t) \in \mathbb{R} \ \forall\ k \in V$. More precisely:

Definition 4 (Gossip interaction)**.** Let assume that at time t the edge (j,k) is selected from the set of available edges at that time $E(t)$; the agents then

update their internal variable according to:

$$\begin{aligned} x_j(t+1) &= (1-\alpha)x_j(t) + \alpha x_k(t) \\ x_k(t+1) &= (1-\alpha)x_k(t) + \alpha x_j(t) \\ x_l(t+1) &= x_l(t) \ \ \forall l \in V \backslash \{j,k\}. \end{aligned} \tag{1.15}$$

with $\alpha \in (0,1)$.

Let us now write the previous expression in a more compact way. Let us denote with $e_i \in \mathbb{R}^{|V|}$ the vector with all null component but the i-th that is set equal to 1 i.e, $e_i = (0,\dots,0,1,0,\dots,0)$. Consider now the $\mathbb{R}^{|V|} \times \mathbb{R}^{|V|}$ matrix:

$$W_{j,k} = \mathbb{I} - \alpha(e_j - e_k)(e_k - e_j)^T. \tag{1.16}$$

If a time t the edge (j,k) is selected then $P(t) = W_{j,k}$ and the gossip interaction can be written as:

$$x(t+1) = W_{j,k}x(t) \tag{1.17}$$

It is now easy to see that the gossip interaction preserve at each step the total average across the network, the matrix $W_{j,k}$ is in fact doubly stochastic. Hence an algorithm designed whose steps are interactions of the form (1.15) is in principle a good candidate for reach asymptotically average consensus.

In order to achieve convergence to consensus it is crucial the design of the selection mechanism of the edges. Let us consider two mechanism a deterministic one and a random one:

- *Cyclic interaction:* at each time t one link $(j(t),k(t))$ is selected deterministically by cycling through the elements of a time-invariant edge set E.
- *Random interaction:* at each time t one link $(j(t),k(t))$ is selected at random, $(j(t),k(t))$ being a single-valued random variable onto the edge set $E(t)$.

These selection mechanisms allow the information flow only between a single edge of the communication graph, namely the selected one; for this reason this type of algorithms is said *asynchronous*. On the other hand, if the selection mechanism is such that all the agent communicates in a single iteration the algorithm is said *synchronous*. In section 7.1 we are going to see how by

using the neighborhoods we can define an intermediate situation in which a small number of edges are selected for a single interaction. In what follows we are going to consider mostly asynchronous algorithms.

The synchronous scheme rely on a global timing structure across the network, such requirement, in the limit of a large network might be unreasonable even in a scenario where the graph is time independent [12]. The random interaction can be realized assuming that each agent posses a clock ticking at some probabilist rate [12] (e.g. Poissonian) in way to ensure that the probability that two agents are activated at the same instant is negligible, once activated the agent choose among its neighbors according some probability distribution (e.g. uniform).

1.1.5 Asymptotic Behavior

The asymptotic behavior of iteration of the gossip interaction according to the the cyclic interaction and the random interaction have been widely studied across the literature and are summarized in the following propositions that precise the requirement on the connectivity of the underlying graph. The following result (see e.g. [12]) characterizes convergence to consensus with the gossip algorithm.

Proposition 1. Consider $G(V, E)$ an undirected graph that is connected hence for any pair of vertices $a, b \in V$, there exists a sequence of vertices $v_0 = a, v_1, v_2, ..., v_{n-1}, v_n = b$ such that $(v_{k-1}, v_k) \in E$ for all $k = 0, 1, ..., n$. If one step of the classical gossip algorithm (1.15) is applied at each time, selecting the updated edge by cyclically running through all the edges of $G(V, E)$, then the system exponentially converges to average consensus. Moreover, if the updated edge (j, k) is selected randomly according to a fixed probability distribution $\{q_{j,k}\}$, with all $q_{j,k} > 0$, then asymptotic average consensus is ensured with probability one, in the sense that: for any $\delta, \varepsilon > 0$, there exists a time $T > 0$ such that

$$\mathbb{P}\left[\|x_k(T) - \bar{x})\|^2 > \varepsilon \|x_k(0) - \bar{x}\|^2\right] < \delta\,,$$

where $\mathbb{P}$ denotes the probability measure induced by the randomization, $\|x\|^2 = \sum_{k=1}^m x_k^T x_k$ and $\bar{x} = \frac{1}{m}\sum_{k=1}^m x_k(t) = \sum_{k=1}^m x_k(0)$ for any choice of the edges.

Proof. We denote $x^T x = \|x\|^2$ for short and $|E|$ the number of edges in $G(V, E)$. At any step of the gossip algorithm, $W := \frac{1}{2m}\sum_{k,j=1}^m \|x_k - x_j\|^2 = \sum_{k=1}^m \|x_k - \bar{x}\|^2$ can only remain unchanged (if the two nodes of the selected

edge have the same value) or decrease (as soon as an edge with different node values is selected). Therefore W is a (non-strict) Lyapunov function for the system dynamics, and when the edges of a connected graph are selected in a cyclic way, a direct application of the LaSalle invariance theorem (see e.g. [16]) shows that the system asymptotically converges to the consensus set. Since the map associated to one full cycle of edge selections is linear and time-invariant, this convergence is exponential. For such convergence to be possible, there must exist some $\lambda > 0$ and integer $M > 0$ such that $W(T) \leq W(0)\lambda$ if the edge choice between $t = 0$ and $t = T = M|E|$ corresponds to M cycles of gossip. When edges are selected randomly, any particular sequence of b consecutive edge selections has a probability greater than $\bar{q}^b > 0$ to appear at least once during any time interval of length at least b, where $\bar{q} = \min_{(j,k)\in E} q_{j,k}$. In particular, if we target $W(T) < \varepsilon W(0) = \lambda^r W(0)$, we can say that there is a probability at least $\bar{q}^{rM|E|}$ to select r times a succession of M cyclic interactions between $t = t_0$ and $t = t_0 + rM|E|$. If this happens once, any preceding or following edge choice can only improve W (because of our first statement in this proof). We conclude by noting that over a time interval $brM|E|$, there is then a probability $< (1 - \bar{q}^{rM|E|})^b$ to have never selected r times a succession of M cyclic interactions, and thus potentially miss $W(T) < \varepsilon W(0)$; the probability that this happens can be made arbitrarily small by taking b (thus T) sufficiently large. □

Let us now briefly recall the convergence condition in the case of a time-dependent set of edges for the determinist and the random selection [17, 12].

Proposition 2. If there exists some finite and fixed $T > 0$ such that the union of edges selected during $[t, t+T]$ form a connected graph for all t, then iteration of the gossip interaction (1.15) asymptotically leads to average consensus.

Proposition 3. If there exists some finite and fixed $T > 0$ and $\delta > 0$ such that the union of edges selected during $[t, t + T]$ form a connected graph for all t with probability greater of equal than δ, then iteration of of the gossip interaction (1.15) asymptotically leads to average consensus with probability 1.

Summing up, iterations of the gossip interaction thus perform a distributed asynchronous computation of the average, in a robust way with respect to the network size and topology and to parameter α, as long as the graph is not completely disconnected.

Before closing this section we would like to make few remarks.

- We have formulate the gossip interaction in the case that the internal variable of the agent is a scalar quantity, anyway the generalization to the vector case is straightforward, (1.15) is left unchanged while (1.17) is generalize by using in (1.16) $e_i \in \mathbb{R}^{n|V|}$

$$e_i = (\mathbf{0}, \dots, \mathbf{0}, \mathbf{1}, \mathbf{0}, \dots, \mathbf{0}) \tag{1.18}$$

  with $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^n$ and $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^n$. Furthermore convergence is guaranteed with the same assumptions.

- The gossip algorithm is defined over a directed graph, in fact it relies on symmetric communication. Algorithms to solve the consensus problem in the context of directed graphs have been widely studied in the literature and the condition for convergence have been characterized, see for example [11].

In the next section we begin to present our contributions. We will reformulate the gossip interaction as a convex combination of suitable permutations and we will characterize consensus as invariance with respect to the action of the permutation group of the asymptotic global state.

2

From gossip interaction to group action

In the previous section we have briefly reviewed the consensus problem and the algorithms to tackle it focusing on the particular instance of the average consensus problem and on the gossip algorithm.

Here we show how the gossip algorithm can be reinterpreted as a convex combination of permutations, see Appendix A.2. This allows us to redefine the consensus situation as invariance with respect to the action of the permutation group of the asymptotic global state. For sake of simplicity we will again assume without loss of generality that the internal state is a real scalar variable. The evolution associated to a single gossip interaction (1.15) can be interpreted as a convex combination of two permutations, namely the trivial one (identity) and the transposition that swaps the internal states of the agents linked by the selected edge. This can be seen by looking at the matrix that defines the iteration. More precisely, let us suppose that at time t the edge (j,k) is selected, then $P(t) = W_{j,k}$ can be rewritten as:

$$\begin{aligned} W_{j,k} &= \mathbb{I} - \alpha(e_j - e_k)(e_j - e_k)^T \\ &= \sum_{i \in V} e_i e_i^T - \alpha(e_j e_j^T - e_k e_k^T) + \alpha(e_j e_k^T - e_k e_j^T). \end{aligned} \tag{2.1}$$

Up to a reordering of the basis $\{e_i\}_{i=1}^{|V|}$ the operator $W_{j,k}$ can be decomposed as:

$$W_{j,k} = (1-\alpha)\mathbb{I}_{j,k} \oplus \mathbb{I}_{\overline{j,k}} + \alpha A_{j,k} \oplus \mathbb{I}_{\overline{j,k}} \tag{2.2}$$

where $\mathbb{I}_{j,k}$ is the identity operator on the subspace spanned by $\{e_j, e_k\}$, $\mathbb{I}_{\overline{j,k}}$ the identity operator on the orthogonal complement of the subspace spanned by $\{e_j, e_k\}$ in $\mathbb{R}^{|V|}$ and $A_{j,k}$ is define as:

$$A_{j,k} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.3}$$

It is easy to see that $S_{j,k} = A_{j,k} \oplus \mathbb{I}_{\overline{j,k}}$ it the transposition that swaps of the j-th with the k-th element of x. Note that transpositions act nontrivially only on two agent at a time, furthermore these agents are "close" in the graph topology.

The action of $W_{j,k}$ can be thus rewritten as follows:

$$\begin{aligned} (x_j(t+1),\ x_k(t+1)) &= (1-\alpha)\ (x_j(t),\ x_k(t)) + \alpha\ (x_k(t),\ x_j(t)) \\ x_\ell(t+1) &= x_\ell(t) \quad \text{for all } \ell \notin \{j,k\} \end{aligned} \tag{2.4}$$

We have shown how the gossip interaction can be seen as a convex combination of permutations. Now we show that this feature can be extended also to the propagator an how this finding reflects on the definition of consensus situation.

Let $\mathfrak{P}_V$ denote the group of all permutations of the set V and for $\pi \in \mathfrak{P}_V$ let P_π be the unique linear operator such that:

$$P_\pi\,(x_1, x_2, ..., x_{|V|}) = (x_{\pi(1)}, x_{\pi(2)}, ..., x_{\pi(|V|)}) \qquad \forall\ (x_1, x_2, ..., x_{|V|})\ \in\ \mathbb{R}^{|V|}. \tag{2.5}$$

By using linearity of the gossip interaction (1.15) and basic group properties, it is also possible to show that the evolution up to time t of the full state vector $x(t) = (x_1(t), \ldots, x_m(t))$ can always be written — although maybe not uniquely — as a convex combination of permutation operators on the initial states [1]:

$$x(t) = \sum_{\pi \in \mathfrak{P}_V} w_\pi(t) P_\pi\, x(0),$$

with $w_\pi(t) \geq 0$ and $\sum_\pi w_\pi(t) = 1$ for every t.

Note that every map of this form preserves the average $\bar{x}(t)$ because any convex combinations of permutations is a doubly stochastic matrix.

An important result is that the connectedness of a graph is equivalent to the property that the transpositions associated to its edges generate the whole permutation group, see Appendix A.2 or [18, 19]. It is easy to see that each permutation can be decomposed as a sequence of transposition, yet this decomposition is not unique and the set of all the transpositions is not minimal generating set. Let us introduce the concept of *transposition graph*.

Definition 5 (transposition graph)**.** Let S be a set of transpositions in the permutation group of a set of n objects $\Sigma := \{1, \ldots, n\}$,. A transposition

[1] This basic result will be proved in a more general setting later.

graph $T(S)$ is defined as the graph:

$$T(S) = G(\Sigma, \Lambda), \tag{2.6}$$

with Λ given by:

$$\Lambda = \{(i,j) \in \Sigma \times \Sigma : \ S_{i,j} \in S\}. \tag{2.7}$$

In the context of gossip model we associate to each edge the correspondent transposition, hence our graph $G(V,E)$ can be seen as a transposition graph with $\Sigma = V$ and with the $\Lambda = E$.

Proposition 4 ([19])**.** Let S be a set of transpositions in the permutation group $\mathfrak{P}_n$ of n objects. Then S is a generating set for $\mathfrak{P}_n$ if and only if its transposition graph $T(S)$ is connected.

A necessary condition for a state x for being a consensus state is to be an equilibrium of for the gossip algorithm and this in turn imply that:

$$W_{j,k}x = x \quad \forall\,(j,k) \in E \implies S_{j,k}x = x \quad \forall\,(j,k) \in E \tag{2.8}$$

Furthermore, if the graph is connected, the set of available transposition is a generating set of for $\mathfrak{P}_V$.

The above facts allow us to reformulate the consenus situation as being any state in the set of permutation invariant states:

$$\mathcal{C} = \{x \in \mathcal{X} = \mathbb{R}^{|V|n} : \ P_\pi\, x = x \ \text{ for all } \pi \in \mathfrak{P}_V\}\,. \tag{2.9}$$

Hence, consensus can be equivalently described as reaching a state that is invariant under (the action P_π on $\mathcal{X}$ of) any element of the permutation group.

We call this *symmetrization with respect to the permutation group.* In the next sections we develop a general framework to tackle symmetrization tasks by iterative, distributed algorithms. This allows for direct extension of the gossip consensus example to different state spaces, to networks that are more general than graphs, and to computational or control tasks not directly related to networks and consensus.

2.1 Symmetrization from Group Actions

In this section we presents the key definitions and algorithmic elements of finite-group symmetrization on vector spaces. In particular, linear gossip can be seen as a particular case of this class of symmetrizing iterations. Further examples are developed in Section 4.

2.1.1 Notation and Symmetrization Task

Let $\mathcal{G}$ be a finite group, with number of elements $|\mathcal{G}|$. Let $\mathcal{X}$ be a vector space over a field $\mathbb{R}$ or $\mathbb{C}$, endowed with an inner product: $\langle\ ,\ \rangle : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{C}$ and consider the induced norm $||\cdot|| : \mathcal{X} \longrightarrow \mathbb{R}$.

We now introduce the concept of *linear action* of a group on a vector space, this mathematical tool allows us to formalize the symmetrization framework.

Definition 6 (linear action)**.** The *linear action* of $\mathcal{G}$ on $\mathcal{X}$, that is a linear map:

$$a : \mathcal{G} \times \mathcal{X} \to \mathcal{X}, \tag{2.10}$$

such that, for all $x, y \in \mathcal{X}$ and all $g, h \in \mathcal{G}$:

$$\begin{aligned} a(g, \alpha x + \beta y) &= \alpha a(g, x) + \beta a(g, y) \quad \alpha, \beta \in \mathbb{F}, && (2.11)\\ a(hg, x) &= a(h, a(g, x)), && (2.12)\\ a(\mathrm{e}_{\mathcal{G}}, x) &= x. && (2.13) \end{aligned}$$

where $\mathrm{e}_{\mathcal{G}}$ is the identity of $\mathcal{G}$.

Note that for a fixed $g \in \mathcal{G}$ we have that $a(g, \cdot) : \mathcal{X} \to \mathcal{X}$ is a linear map on $\mathcal{X}$. More precisely every linear action is associated to a *representation* of $\mathcal{G}$ on $\mathcal{X}$ (see section A.2). By using the inner product of $\mathcal{X}$, we can define the *adjoint* of $a(g, \cdot)$ as the unique linear map $a^\dagger(g, \cdot)$ that satisfies:

$$\langle y\,,\, a(g, x)\rangle = \langle a^\dagger(g, y)\,,\, x\rangle \quad \forall\, x, y\, \in \mathcal{X}\,. \tag{2.14}$$

We now reformulate the consensus situation within this new framework.

Definition 7 (fixed point of the action)**.** An element $\bar{x} \in \mathcal{X}$ is a fixed point of the action of $\mathcal{G}$ if

$$a(g, \bar{x}) = \bar{x} \quad \forall\, g \in \mathcal{G}. \tag{2.15}$$

We denote the set of such fixed points as $\mathcal{C}^{\mathcal{G}} \subseteq \mathcal{X}$.

Due to the linearity of the action, $\mathcal{C}^{\mathcal{G}}$ is a vector space. We are now ready to set our goal:

Goal Our main goal is the *symmetrization* of any initial condition $x \in \mathcal{X}$ with respect to the action of $\mathcal{G}$, that is, construct an algorithm or a dynamical system that (asymptotically, with probability 1) drives any $x \in \mathcal{X}$ to some related $\bar{x} \in \mathcal{C}^{\mathcal{G}}$.

Consider any time-varying discrete-time dynamics $x(t+1) = \mathcal{E}_t(\, x(t)\,)$ on $\mathcal{X}$. We denote $\mathcal{E}_{t,0}(\cdot)$ the map associated to the evolution from time 0 up to time t, such that $x(t) = \mathcal{E}_{t,0}(\, x(0)\,)$ we call this map *propagator*. Let $\|\cdot\|$ be a norm associated to the inner product in $\mathcal{X}$.

Definition 8 (Asymptotic symmetrization A.S.)**.** The algorithm associated to iterations $\{\mathcal{E}_t\}_{t\geq 0}$ attains asymptotic symmetrization if for all $x \in \mathcal{X}$ it holds:

$$\lim_{t\to\infty} \|a(g, \mathcal{E}_{t,0}(x)) - \mathcal{E}_{t,0}(x)\| = 0 \quad \forall\, g \in \mathcal{G}. \tag{2.16}$$

We will also consider sequences of maps $\{\mathcal{E}_t\}_{t\geq 0}$ that can be randomized; in this case, the above definition applies but convergence with probability one is understood.

Our definition of asymptotic symmetrization generalize thus the definition of asymptotic consensus in our group theoretic framework. Note that, at this stage the definition of A.S. does not required that the algorithm drive any initial condition to a single state (or alternative that the propagator converge to a fixed map) but only that asymptotical the state trajectory is confined in the set of fixed point of the action. Note that the "classical" consensus definition prescribe convergence toward an equilibrium.

In the next section we study a particular class of algorithm, those whose step are, following the analogy with the gossip interaction, convex combinations of group action selected according to some mechanism for a subset of a given group.

2.2 A Class of Algorithms

In this section we concentrate on those particular dynamics that can be obtain with interactions given by convex combinations of actions. We first consider one step evolution, that are the analogous of the gossip interaction and then we study the dynamics obtained by the concatenation of such maps.

For a given group $\mathcal{G}$, vector space $\mathcal{X}$ and linear action $a : \mathcal{G} \times \mathcal{X} \to \mathcal{X}$, we will be interested in linear maps $\mathcal{F}$ of the form:

$$\mathcal{F}(x) \quad = \quad \sum_{g \in \mathcal{G}} \mathsf{s}_g \, a(g, x) \tag{2.17}$$

with $\mathsf{s}_g \geq 0$ and $\sum_{g \in \mathcal{G}} \mathsf{s}_g = 1$ for every g. Such a map is completely specified by the choice of convex weights s_g. From here on, we shall call a vector whose elements are nonnegative and sum to 1 a *vector of convex weights*.

We construct discrete-time dynamics on $\mathcal{X}$ by selecting at each time step t a vector of convex weights, more precisely we consider interaction of the type:

Interaction: Let assume that at time t is selected the vector

$$\mathsf{s}(t) = (\mathsf{s}_{g_1}(t), \mathsf{s}_{g_2}(t), \ldots, \mathsf{s}_{g_{|\mathcal{G}|}}(t)) \in \mathbb{R}^{|\mathcal{G}|},$$

the state $x(t)$ is update through the corresponding map of type $\mathcal{F}(x)$, i.e.

$$x(t+1) = \mathcal{E}_t(x(t)) := \sum_{g \in \mathcal{G}} \mathsf{s}_g(t) \, a(g, x(t)) \, . \tag{2.18}$$

We assume that $\mathsf{s}(t)$ is selected deterministically or randomly from some possibly infinite set $\mathbf{S}$. Inspired by the gossip interaction, where each step of the dynamics is induced by a doubly stochastic matrix that is a convex combination of the identity and a swap operation, we consider scenarios where any $\mathsf{s} \in \mathbf{S}$ assigns nonzero weights only to a restricted set of $g \in \mathcal{G}$. This remains a sensible requirement also in applications unrelated to the gossip algorithm. It might be unfeasible, for a given set of control capabilities:

- to implement a large number of group elements in a single instant,
- act non trivially on a large set of agents simultaneously.

From a dynamical systems perspective, we can interpret (2.18) as a discrete-time *switching* system, whose generator is chosen at each time between a set of maps of the form (2.17), according to the switching signal $\mathsf{s}(t)$. The resulting propagator $\mathcal{E}_{t,0}(\cdot)$ is also a convex combination of group actions, i.e. of the form $\mathcal{F}(\cdot)$ given in (2.17).

Lemma 1. If the iterations have the form (2.18), then there exists a vector $\mathsf{p}(t) = (\mathsf{p}_{g_1}(t), \mathsf{p}_{g_2}(t), \ldots, \mathsf{p}_{g_{|\mathcal{G}|}}(t)) \in \mathbb{R}^{|\mathcal{G}|}$ such that for any t we can write:

$$x(t) = \mathcal{E}_{t,0}(x(0)) = \sum_{g \in \mathcal{G}} \mathsf{p}_g(t)\, a(g, x(0)) \tag{2.19}$$

for any $x(0) \in \mathcal{X}$, with $\mathsf{p}_g(t) \geq 0\ \forall g$ and $\sum_{g\in\mathcal{G}} \mathsf{p}_g(t) = 1$.

Proof. Proceed by inductive reasoning on t. For $t = 1$, (2.19) trivially holds because $\mathcal{E}_{1,0}(x) = \mathcal{E}_0(x)$ is given by (2.18). Now assume (2.19) holds for some t. Then

$$\begin{aligned}
\mathcal{E}_{t+1,0}(x) &= \mathcal{E}_t \circ \mathcal{E}_{t,0}(x) \\
\text{(def. } \mathcal{E}) &= \sum_{h\in\mathcal{G}} \mathsf{s}_h(t) a(h, \sum_{g\in\mathcal{G}} \mathsf{p}_g(t) a(g, x)) \\
\text{(linearity)} &= \sum_{h,g\in\mathcal{G}} \mathsf{s}_h(t)\mathsf{p}_g(t)\, a(h, a(g, x)) \\
\text{(def.action)} &= \sum_{h,g\in\mathcal{G}} \mathsf{s}_h(t)\mathsf{p}_g(t)\, a(hg, x)) \\
\text{(var.change)} &= \sum_{h,g'\in\mathcal{G}} \mathsf{s}_h(t)\mathsf{p}_{h^{-1}g'}(t)\, a(g', x)) \\
&= \sum_{g'\in\mathcal{G}} \mathsf{p}_{g'}(t+1)\, a(g', x))\,,
\end{aligned}$$

where we have defined:

$$\mathsf{p}_{g'}(t+1) = \sum_{h\in\mathcal{G}} \mathsf{s}_h(t)\mathsf{p}_{h^{-1}g'}(t). \tag{2.20}$$

Noting that $g' \mapsto h^{-1}g'$ is a group automorphism such that $\sum_{g'\in\mathcal{G}}\ \mathsf{p}_{h^{-1}g'}(t) = 1$ for each fixed h, one easily checks that $\mathsf{p}(t+1)$ satisfies the requirements of a vector of convex weights. Hence the statement holds for $t+1$ and we get the conclusion by induction. □

Note that (2.20) is equivalent to the definition of group convolution [19] and therefore can be written in the more compact way:

$$\mathsf{p}(t+1) = \mathsf{s}(t) \star \mathsf{p}(t), \tag{2.21}$$

as the convolution of $\mathsf{s}(t)$ with $\mathsf{p}(t)$, i.e. as the convolution of the v.c.w. the define the interaction step with the v.c.w. that keep track of the group elements that have been so far implemented.

In this section we have begun the characterization of the evolution resulting from interactions in the form of convex condition of action highlighting how

this represent a reasonable control assumption in relevant scenarios. We have also proved that this feature is reflected in the form of the propagator. In the next section we are going to study the asymptotic behavior of this class of algorithms.

2.3 The Symmetrizing Map

In Definition 8 of Section 2.1 we have redefined the consensus situation as *asymptotic symmetrization A.S.*. A general time-varying map might achieve symmetrization according to (2.16) without ever converging to a fixed point the Definition of A.S., in fact, only requires that the state trajectory is confined into the set of fixed point $\mathcal{C}^{\mathcal{G}}$. However, for dynamics of the form of the *group intercation*, i.e.:

$$x(t+1) = \mathcal{E}_t(x(t)) := \sum_{g\in\mathcal{G}} \mathsf{s}_g(t)\, a(g, x(t))\,. \tag{2.22}$$

with $\mathsf{s}(t)$ v.c.w., we have the following result.

Proposition 5. An evolution defined by $\mathcal{E}_t$ of the form (2.22) attains asymptotic symmetrization if and only if the propagator:

$$\mathcal{E}_{t,0}(\cdot) = \sum_{g\in\mathcal{G}} \mathsf{p}_g(t)\, a(g,\cdot), \tag{2.23}$$

converges to the fixed map

$$\bar{\mathcal{F}}(\cdot) \;:=\; \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} a(g,\cdot)\,. \tag{2.24}$$

Proof. Assume asymptotic symmetrization is attained. Taking the (finite) sum over all $g \in \mathcal{G}$ of:

$$\lim_{t\to\infty} \|a(g, \mathcal{E}_{t,0}(x)) - \mathcal{E}_{t,0}(x)\| = 0 \quad \forall\, g \in \mathcal{G}, \tag{2.25}$$

then dividing by $|\mathcal{G}|$ and using the triangle inequality gives:

$$\begin{aligned}
0 &= \lim_{t\to+\infty} \left\| \tfrac{1}{|\mathcal{G}|} \sum_{g\in\mathcal{G}} a\left(g, \sum_{h\in\mathcal{G}} \mathsf{p}_h(t) a(h,x)\right) \;-\; \mathcal{E}_{t,0}(x) \right\| \\
&= \lim_{t\to+\infty} \left\| \tfrac{1}{|\mathcal{G}|} \sum_{g,h\in\mathcal{G}} \mathsf{p}_h(t)\, a\left(\, g,\, a(h,x)\,\right) \;-\; \mathcal{E}_{t,0}(x) \right\| \\
&= \lim_{t\to+\infty} \left\| \tfrac{1}{|\mathcal{G}|} \sum_{g,h\in\mathcal{G}} \mathsf{p}_h(t)\, a(gh,x) \;-\; \mathcal{E}_{t,0}(x) \right\| \\
&= \lim_{t\to+\infty} \left\| \tfrac{1}{|\mathcal{G}|} \sum_{g,h'\in\mathcal{G}} \mathsf{p}_{g^{-1}h'}(t)\, a(h',x) \;-\; \mathcal{E}_{t,0}(x) \right\| \qquad (2.26) \\
&= \lim_{t\to+\infty} \left\| \tfrac{1}{|\mathcal{G}|} \sum_{h'\in\mathcal{G}} a(h',x) \;-\; \mathcal{E}_{t,0}(x) \right\| \qquad (2.27)
\end{aligned}$$

for all $x \in \mathcal{X}$, which would imply that $\mathcal{E}_{t,0}$ converges to $\bar{\mathcal{F}}$. To go from (2.26) to (2.27), we sum on g for each fixed h': that yields $\sum_{g\in\mathcal{G}} \mathsf{p}_{g^{-1}h'}(t) = \sum_{g'\in\mathcal{G}} \mathsf{p}_{g'}(t) = 1$ for all h', thanks to the facts that $g \mapsto g^{-1}$, and $g \mapsto gh$ (for fixed h), are group automorphisms. The proof of the converse is trivial. □

The proof builds on the finite cardinality of $\mathcal{G}$ and remains valid if $\mathcal{X}$ is infinite-dimensional. From the proof follows an immediate corollary:

Corollary 1. This establishes that $\bar{\mathcal{F}}$ is the unique projector onto $\mathcal{C}^{\mathcal{G}}$ that can be obtained as a convex combination of the group actions.

Notice however that if the actions associated to different $g \in \mathcal{G}$ are not all linearly independent, there will be more than one vector p corresponding to the same map $\mathcal{F}$, this fact will be further discuss in the next section.

We now present some results that holds if, for each element, the adjoint of the action associated to the group element is equal to the action associated to some group element.

Lemma 2. If there exists a group automorphism $g \mapsto h(g)$ such that

$$a^{\dagger}(g,\cdot) = a(h(g),\cdot) \quad \forall g \in \mathcal{G}\,, \tag{2.28}$$

then $\bar{\mathcal{F}}$ is an orthogonal projection.

Proof. Eq. (2.24) readily yields that $\bar{\mathcal{F}} = \bar{\mathcal{F}}^2$ and that (2.28) ensures $\bar{\mathcal{F}} = \bar{\mathcal{F}}^{\dagger}$. □

This indeed holds for example if the group action is a unitary representation of $\mathcal{G}$. Note that if the group is finite it alway admits such a representation. Unitary actions are also part of a class of actions that allow us to easily determine a set of conserved quantities, depending only on the initial $x(0)$, as is the case for the mean in the gossip example.

Lemma 3. If there exists an endomorphism (not necessarily an automorphism) $g \in \mathcal{G} \mapsto h(g) \in \mathcal{G}$ such that (2.28) holds, then for any $\bar{z} \in \mathcal{C}^{\mathcal{G}}$ we have

$$\langle \bar{z}, x(t)\rangle = \langle \bar{z}, x(0)\rangle \quad \forall\, t\,. \tag{2.29}$$

Proof. For any t it holds that:

$$\begin{aligned}\langle \bar{z}, x(t)\rangle &= \langle \bar{z}\,, \sum_{g\in\mathcal{G}} \mathsf{p}_g(t)\, a(g, x_0)\rangle \\ &= \sum_{g\in\mathcal{G}} \mathsf{p}_g(t)\, \langle a^\dagger(g, \bar{z})\,,\, x_0\rangle \\ &= \sum_{g\in\mathcal{G}} \mathsf{p}_g(t)\, \langle a(h(g), \bar{z})\,,\, x_0\rangle \\ &= \sum_{g\in\mathcal{G}} \mathsf{p}_g(t)\, \langle \bar{z}, x_0\rangle \;=\; \langle \bar{z}, x_0\rangle\,.\end{aligned}$$

□

Note that if we are interested in the average $\bar{z} = \mathbf{1}/|\mathcal{G}|$. The above proof works as soon as $h(g) \in \mathcal{G}$ for every $g \in \mathcal{G}$ but does not requires the map $h(\cdot)$ to be either injunctive or surjective.

Before close this section let us apply the symmetrization framework to the gossip algorithm.

2.3.1 Example: Linear Gossip

Consider the gossip algorithm described in Section 1.1.3, assume that each agent is assigned an internal state $x_k(t) \in \mathbb{R}^n$. To recast it in our framework, we choose:

- $\mathcal{X} = \mathbb{R}^{|V|n}$.
- $\mathcal{G} = \mathfrak{P}_V$ the group of all permutations of the set V.

We can think of any $x \in \mathcal{X}$ as a column vector that stacks the n-dimensional state vectors of the $|V|$ subsystems. With the linear permutation operator P_π defined Section 2, the action of the group is simply:

$$a(\pi, x) = P_\pi x. \tag{2.30}$$

Notice that this action is self-adjoint every permutation can, in fact, be written as a finite product of tranpositions.

We have already established that consensus corresponds to the fixed points of this action, i.e. $\mathcal{C} = \mathcal{C}^{\mathfrak{P}_V}$.

From Proposition 5 and Lemma 2 (with the trivial automorphism $h(g) = g$), the map:

$$\bar{\mathcal{F}} = \frac{1}{m!} \sum_{\pi\in\mathfrak{P}_V} P_\pi\,, \tag{2.31}$$

is the orthogonal projection onto the consensus set.

Next we turn to the evolution model. For the linear gossip, at any time the $|V|!$-dimensional vector $\mathsf{s}(t)$ has only two nonzero entries:

- $(1-\alpha(t))$ on the component corresponding to the group identity,
- $\alpha(t)$ associated to swapping j and k.

If α and the graph with $|E|$ edges are constant, then $\mathsf{s}(t)$ can switch between $|E|$ values. Let P_{e} and $P_{(j,k)}$ denote the linear operators P_π that respectively implement the identity and the swapping of subsystems j and k. These can be represented as $n|V| \times n|V|$ matrices: $P_{\mathrm{e}} = I_{n|V|}$, the identity, and $P_{(j,k)} = Q_{(j,k)} \otimes I_n$, the Kronecker product between the identity on $\mathbb{R}^n$ and $Q_{(j,k)}$ the $|V| \times |V|$ matrix that swaps the coordinates j and k of a vector of length $|V|$. Then the elementary evolution step associated to the selection of edge (j,k) at time t writes:

$$\begin{aligned} x(t+1) &= \sum_\pi \mathsf{s}_\pi(t)\, a(\pi, x(t)) \\ &= (1-\alpha(t))\, P_{\mathrm{e}} x(t) + \alpha(t)\, P_{(j,k)} x(t). \end{aligned}$$

Finally, let us look at conserved quantities. Denoting z_c the value on row c of vector $z \in \mathcal{X} = \mathbb{R}^{n|V|}$, the set $\mathcal{C} = \mathcal{C}^{\mathfrak{P}_V}$ consists of all $z \in \mathcal{X}$ such that $z_{jn-d+1} = z_{kn-d+1}$ for all subsystems $j, k \in \{1, 2, ..., |V|\}$ and all components $d \in \{1, 2, ..., n\}$. This vector space is spanned in particular by the vectors $z^d \in X$, $d = 1, 2, ..., n$, defined by:

$$z^d_{jn-d+1} = 1/m \text{ for all } j\,, \text{ other components } 0\,.$$

Hence by Lemma 3, we get as conserved quantities all the linear functionals of the form

$$\langle \bar{z}, x \rangle \;=\; \sum_{d=1}^{n} f_d \,\langle z^d, x \rangle \;=\; \sum_{d=1}^{n} f_d \;\mathrm{avg}(x)_d,$$

with arbitrary $f_1, f_2, ..., f_n \in \mathbb{R}$, where $\mathrm{avg}(x)_d$ denotes the average of the d^{th} component of the subsystem states.

3

Action-independent Dynamics

In the last section we have shown that in order to have asymptotic symmetrization according to Definition 8 when the gossip interactions are convex combinations of group actions, asymptotically the propagator have to converge to the fixed map given that is a convex combination of group actions with uniforms weights over the group.

In this section we begin to investigate under which conditions we can ensure convergence to this fixed map. More precisely, we discusses *sufficient conditions* for obtaining symmetrization, that are *independent of the actions* but depend only on $\mathcal{G}$ and on the selected sequence of convex weights $\mathsf{s}(t)$ at each step. These conditions are also *necessary* if the particular actions associated to all elements of $\mathcal{G}$ are linearly independent. Since such actions exist for any finite group $\mathcal{G}$, the following conditions can be viewed as *necessary and sufficient for obtaining symmetrization on all possible actions associated to a given group dynamics*

One representation with linearly independent elements is the *regular representation* see Appendix A.2 for more details. In other words:

Goal we ensure asymptotic symmetrization for a general group-based algorithm in the form (2.18) based only on:

- the group properties,
- the selection rules for the convex vectors $\mathsf{s}(t)$,

for *any* underlying vector spaces and action.

This frees us from the need to prove convergence for each specific application. Section 4 provides a series of examples obtained by extending in this way

the gossip-type algorithm.

More explicitly, Lemma 1 suggests that for studying the dynamics on $\mathcal{X}$ according to (2.18), it is sufficient to look at the evolution of the convex weights $\mathsf{p}(t)$, for the reasons explained above we call the evolution of $\mathsf{p}(t)$ *lifted evolution*. The proof of the Lemma proposes the dynamics

$$\mathsf{p}_g(t+1) \;=\; \sum_{h\in\mathcal{G}} \mathsf{s}_h(t)\mathsf{p}_{h^{-1}g}(t) \tag{3.1}$$

for all $g \in \mathcal{G}$. If the group actions are linearly dependent, then several weights $\mathsf{s}(t)$ or $\mathsf{p}(t)$ can be associated to any map of the form $\mathcal{F}$ and clearly (3.1) is not the unique dynamics corresponding to (2.18). However, if we want to study (2.18) by focusing on the group properties, and prove convergence in a way that is valid for *all possible actions associated to the group*, then (3.1) is the unique lift of (2.18) that achieves this goal. In the current section we hence study the behavior of (3.1).

Again, let us choose an ordering of $\mathcal{G}$ and let us introduce a basis for the v.c.w. $\{\mathsf{e}_{g_i}\}_{i=1}^{|\mathcal{G}|}$ composed by those vectors with all the entries equal to zero but for the i-th, i.e. $e_{g_i} = (0,\dots,0,1,0,\dots,0)$. Furthermore let us donate with $(\cdot\,,\cdot) : \mathbb{R}^{|\mathcal{G}|}\times\mathbb{R}^{|\mathcal{G}|}\to\mathbb{R}$ the scalar product in $\mathbb{R}^{|\mathcal{G}|}$. Hence we have the following orthogonality relations:

$$(\mathsf{e}_g,\mathsf{e}_h) = \delta_{g,h} \qquad \forall\, g,h \,\in\, \mathcal{G}, \tag{3.2}$$

and we have that every vector (v.c.w.) can be written as:

$$\mathsf{p}(t) = \sum_{g\in\mathcal{G}} \mathsf{p}_g(t)\mathsf{e}_g \qquad \text{with} \qquad \mathsf{p}_g(t) := (\mathsf{p}(t),\mathsf{e}_g). \tag{3.3}$$

In view of this we can write (3.1) in a more compact form:

Proposition 6. Let P_h the permutation matrix over $|V|$ objects such that:

$$P_h\mathsf{e}_g = \mathsf{e}_{hg} \qquad \forall\, g,h \,\in\, \mathcal{G}, \tag{3.4}$$

namely P_h denotes the permutation matrix that moves the component from row g of a vector in $\mathbb{R}^{|\mathcal{G}|}$ towards the new row hg.Then we have that:

$$\mathsf{p}(t+1) = \left(\sum_{h\in\mathcal{G}} \mathsf{s}_h(t)P_h\right)\mathsf{p}(t) \tag{3.5}$$

Proof. By using (3.3) we have that:

$$\begin{aligned}
\mathsf{p}(t+1) &= \sum_g \mathsf{p}_g(t+1)\mathsf{e}_g && (3.6)\\
&= \sum_g \left(\sum_h \mathsf{s}_h(t)\mathsf{p}_{h^{-1}g}(t)\right)\mathsf{e}_g && (3.7)\\
&= \sum_g \left(\sum_h \mathsf{s}_h(t)(\mathsf{p}(t),\mathsf{e}_{h^{-1}g})\right)\mathsf{e}_g && (3.8)\\
&= \sum_g \left(\sum_h \mathsf{s}_h(t)(\mathsf{p}(t),P_{h^{-1}}\mathsf{e}_g)\right)\mathsf{e}_g && (3.9)\\
&= \sum_g \left(\sum_h \mathsf{s}_h(t)(P_h\mathsf{p}(t),\mathsf{e}_g)\right)\mathsf{e}_g &&\\
&= \sum_h \mathsf{s}_h(t)\sum_g (P_h\mathsf{p}(t),\mathsf{e}_g)\,\mathsf{e}_g && (3.10)\\
&= \left(\sum_{h\in\mathcal{G}} \mathsf{s}_h(t)P_h\right)\mathsf{p}(t). && (3.11)
\end{aligned}$$

Where from (3.6) to (3.7) we have used (3.1) and from (3.8) to (3.9) the fact that a permutation matrix is an isometry, i.e.:

$$P_h P_h^T = \mathbb{I} \qquad \forall\, h \in \mathcal{G}. \tag{3.12}$$

Hence, we have that $P_{h^{-1}}^T = P_h$ that in turn implies that:

$$(\mathsf{p}(t), P_{h^{-1}}\mathsf{e}_g) = (P_{h^{-1}}^T\mathsf{p}(t),\mathsf{e}_g) = (P_h\mathsf{p}(t),\mathsf{e}_g). \tag{3.13}$$

Finally from (3.10) to (3.11) we have used the fact that:

$$P_h\mathsf{p}(t) = \sum_g (P_h\mathsf{p}(t),\mathsf{e}_g)\,\mathsf{e}_g. \tag{3.14}$$

□

Now if we set $\tilde{M}(t) := \sum_{h\in\mathcal{G}} \mathsf{s}_h(t)P_h$ we can write for the propagator:

$$\begin{aligned}
\mathsf{p}(t+1) &= \left(\sum_{h\in\mathcal{G}} \mathsf{s}_h(t)P_h\right)\mathsf{p}(t) = \tilde{M}(t)\mathsf{p}(t) && (3.15)\\
&= \left(\prod_{i=0}^{t} \tilde{M}(i)\right)\mathsf{p}(0)\,, && (3.16)
\end{aligned}$$

Since for every t the $\tilde{M}(t)$ are convex combination of permutations the Birkhoff theorem ensure that for every t the $\tilde{M}(t)$ are doubly stochastic matrices. For

each given sequence $\mathsf{s}(0), \mathsf{s}(1), ...$, this looks like the transition dynamics of a (time-inhomogeneous) *Markov chain* on the distribution $\mathsf{p}(t)$ over $\mathcal{G}$.

Definition 8 is satisfied independently of the particular actions associated to $\mathcal{G}$ if we can ensure convergence to a vector p such that:

$$\mathsf{p}_g = \mathsf{p}_{h^{-1}g} \quad \forall\, g, h \in \mathcal{G}\,. \tag{3.17}$$

Since for g fixed $\{h^{-1}g : h \in \mathcal{G}\} = \mathcal{G}$, this is equivalent to

$$\mathsf{p}_g = 1/|\mathcal{G}| =: \hat{\mathsf{p}}_g \;\; \forall\, g \in \mathcal{G}\,, \tag{3.18}$$

in accordance with Proposition 5. To attain symmetrization, we thus require that the dynamics of p converges to the unique value $\mathsf{p} = \hat{\mathsf{p}}$ given by (3.18).

The targeted convergence to a uniform distribution $\hat{\mathsf{p}}$ under switched dynamics (3.15) with doubly stochastic transition matrix $\tilde{M}$, is reminiscent of the standard average consensus problem between $|\mathcal{G}|$ agents in $\mathbb{R}$. There are however at least two major differences between these frameworks.

1. The state $\mathsf{p}(t)$ models $\mathcal{E}_{t,0}$ from the original problem. In particular, $\mathsf{p}(0)$ models $\mathcal{E}_{0,0}$ which is the identity. Hence, in principle, we would only need to study the evolution from this *known initial state.*

2. The transition matrix has a different structure inherited from its constituents. For average consensus the transition matrix is essentially the identity plus a sum of symmetric edge-interaction-matrices, with 4 nonzero entries of equal magnitude per edge of the graph. For p, it is a sum of permutation matrices, each of them with $|\mathcal{G}|$ nonzero entries.

The second point actually alleviates the first one: by group translation, convergence to $\hat{\mathsf{p}}$ from the particular initial condition $\mathsf{p}(0)$ corresponding to identity $\mathcal{E}_{0,0}$, implies convergence to $\hat{\mathsf{p}}$ from *any* initial convex weights vector $\mathsf{p}(0)$.

In this section we have shown how to study of the asymptotic behavior of our algorithms is quite simplified by using considering the *lifted dynamics*, whose evolution is driven by a time-inhomogeneous Markov chain.

The following section investigates when the system defined by (3.15) converges to symmetrization. The resemblance with classical consensus will guide us to derive convergence conditions, although they will have to be translated to match the $\mathsf{p}(t)$ and $\mathsf{s}(t)$ structure (see second point).

3.0.2 Example: $\mathsf{p}(t)$ for Gossip Consensus

Let us quickly formulate the gossip algorithm in the action-independent form. In Section 2.3.1, we illustrated how $x(t+1) = A(t)x(t)$, with

$$A(t) \quad = \quad (1-\alpha)\, I_{|V|n} + \alpha(t)\, P_{(j,k)}$$

when edge (j,k) is selected at time t. The doubly-stochastic $\tilde{M}(t) = (1-\alpha)I + \alpha\Pi_{(j,k)}$ describing the $\mathsf{p}(t)$ dynamics has dimensions $|V|! \times |V|!$ (independently of n), with two nonzero entries *on each row and column*: $\tilde{M}_{g,g} = (1-\alpha)$ and $\tilde{M}_{g,\pi_{(j,k)}g} = \alpha$ for all $g \in \mathcal{G}$.

Convergence of the p-dynamics is not necessary for convergence of the linear gossip algorithm. Indeed, a dimension counting argument suffices to show that the corresponding actions of $\mathfrak{P}_V$ are not linearly independent for $|V| \geq 4$: the space of possible actions has dimension $|V|^2$ (consider $A(t) = I_n \otimes A_V(t)$ and count the number of entries in matrix $A_V(t)$), while there are $|V|!$ permutations and $|V|! > |V|^2$ for $|V| \geq 4$. This means that ensuring convergence of the switched $\tilde{M}$ dynamics for p is in principle more demanding than for the switched A for x. However, as we prove in the next section, convergence on p follows from the typical assumptions of consensus, and allows us to draw conclusions that are valid for all possible $\mathcal{X}$ and actions of $\mathfrak{P}_V$.

3.1 Convergence Analysis

We now examine the convergence properties of (3.15) with a switching signal $\mathsf{s}(t)$. This reduces to analyzing an infinite product of doubly stochastic matrices $\tilde{M}(t)$. This problem has been investigated in much detail in other contexts, including standard linear consensus [20, 21, 5, 17]. Among others, [5] proposes a common quadratic Lyapunov function for all possible switchings, which shows that instability is not possible.

The question is then, under which conditions is $\hat{\mathsf{p}}$ *asymptotically* stable. We first give convergence results for deterministic $\mathsf{s}(t)$. Their adaptation to a stochastically selected $\mathsf{s}(t)$ is explained at the end of the section.

3.1.1 Formal Conditions and Convergence Proof

In the context of consensus on graphs, a sufficient condition for convergence is given in terms of a requirement that the union of all edges that appear during a uniformly bounded time interval, must form a connected graph at all times (see e.g. [17]). This result could be applied to (3.15), if we view each group element as a node of a *Cayley graph* (see Appendix A.2) and draw the directed edges that correspond to the group elements h with $\mathsf{s}_h(t) \geq \underline{\alpha} > 0$ at time t.

The problem at hand however has more structure: an arbitrary adjacency matrix for a graph on N nodes has order N^2 parameters, while (3.15) shows that $\tilde{M}(t)$ is defined by $|\mathcal{G}| = N$ elements only — namely the vector $\mathsf{s}(t)$.

In fact, for the sequence of s from time t to time $t+T$ writes let us define a vector of convex weights:

$$\mathsf{q}_g(t+T, T) := s(t+T) \star \cdots \star s(t). \tag{3.19}$$

For the evolution from time t to time $t+T$ we can thus write:

$$\prod_{i=0}^{T-1} \tilde{M}(t+i) = \sum_{g \in \mathcal{G}} \mathsf{q}_g(t+T, T)\, P_g\,. \tag{3.20}$$

This again involves only $|\mathcal{G}| = N$ elements $\mathsf{q}_g(t+T, T)$. We therefore give independent convergence proofs, in the hope to highlight the role of the assumptions in a way that is more natural in the group-theoretic framework. We next formulate a condition that essentially translates the connected-graph requirement (in fact rather its essential consequence, i.e. that the transition matrix from t to $t+T$ is primitive) into our framework.

Assumption 1 (primitivity assumption)**.** Assume the sequence $\mathsf{s}(t)$ to be such that there exist some finite $T, \delta > 0$, such that for each time t:

$$\mathsf{q}_g(t, t+T) > \delta \quad \forall g \in \mathcal{G}\,. \tag{3.21}$$

This assumption can be translated into properties of the transition matrices in (3.15). If $M(t) = M$ for each t, then the assumption is equivalent to M being primitive. In the general case, we request that each $\prod_{i=0}^{T-1} \tilde{M}(t+i)$ is primitive, with all entries at least δ.

Notice how Assumption 1 does not require that $\{g \in \mathcal{G} : \mathsf{s}_g(i) > \delta \text{ for some } i \in [t, t+T]\} = \mathcal{G}$. Thus a priori, the (combination of) available actions for all t may be restricted to a subset $\mathcal{S}$ of $\mathcal{G}$; a necessary condition for Assumption 1 to hold is then that $\mathcal{S}$ generates $\mathcal{G}$. This is equivalent to requiring that the union of edges appearing during a time interval T in the corresponding Cayley graph form a connected graph, but not necessarily the complete graph. The equivalence can be seen as follows, consider the family of Cayley graphs:

$$\Gamma(\mathcal{G}, E_{\mathcal{S}_t}), \ldots, \Gamma(\mathcal{G}, E_{\mathcal{S}_{T+t}}), \tag{3.22}$$

with $\mathcal{S}_k := \{h \in \mathcal{G} :\ s_h(k) \geqslant \delta > 0\}$, their union is still a Cayley graph given by:

$$\cup_{k=t}^{T+t} \Gamma(\mathcal{G}, E_{\mathcal{S}_k}) = \Gamma(\mathcal{G}, \cup_{k=t}^{T+t} E_{\mathcal{S}_k}), \tag{3.23}$$

and it is connected if and only if $\cup_{k=t}^{T+t} \mathcal{S}_k$ forms a set of generator for $\mathcal{G}$. We will further examine Assumption 1 in Section 3.1.2.

Now let us formally establish that Assumption 1 is a sufficient condition to ensure convergence to $\hat{\mathsf{p}}$.

Theorem 1. For any switching sequence $\mathsf{s}(t)$ satisfying Assumption 1, the algorithm (3.15) makes any initial condition $\mathsf{p}(0)$ converge to the uniform vector $\hat{\mathsf{p}}$.

Before giving the proof, let us recall some basic facts about *relative entropy* and the *log sum inequality*. The relative entropy, or Kullback-Leibler pseudo-distance [22] of a vector of convex weights $\{\mathsf{u}_g\}_{g\in\mathcal{G}}$ with respect to another one $\{\mathsf{v}_g\}_{g\in\mathcal{G}}$ is given by:

$$K(\mathsf{u}\|\mathsf{v}) = \sum_{g\in\mathcal{G}} \mathsf{u}_g \left(\log \mathsf{u}_g - \log \mathsf{v}_g\right). \tag{3.24}$$

This expression is not symmetric in u, v, but $K(\mathsf{u}\|\mathsf{v}) \geqslant 0$ and the equality holds if and only if $\mathsf{u} = \mathsf{v}$. We shall also use the following [22].

Proposition 7 (Log Sum Inequality)**.** Let $\{a_i\}_{i=1}^n$ and $\{b_i\}_{i=1}^n$ be nonnegative numbers. Then it holds:

$$\sum_{i=1}^{n} a_i \log \frac{a_i}{b_i} \geqslant \left(\sum_{i=1}^{n} a_i\right) \log \frac{\sum_i a_i}{\sum_i b_i}. \tag{3.25}$$

Furthermore, excluding the singular cases where $\sum_i a_i = 0$ or $\sum_i b_i = 0$, the equality holds if and only if $\frac{a_i}{b_i} = \alpha$ is constant over $i = 1, \ldots, n$.

proof of Theorem 1. $K(\mathsf{p}(t)\|\hat{\mathsf{p}})$ is nonnegative and it equals zero if and only if $\mathsf{p}(t) = \hat{\mathsf{p}}$. To use it as a strict Lyapunov function, it remains to prove that, under Assumption 1, the relative entropy of $\mathsf{p}(t)$ with respect to $\hat{\mathsf{p}}$ strictly decreases after (any) T steps. For every t we have that:

$$\begin{aligned} K(\mathsf{p}(t+T)\|\hat{\mathsf{p}}) &= \sum_{g\in\mathcal{G}} \mathsf{p}_g(t+T) \log \frac{\mathsf{p}_g(t+T)}{\hat{\mathsf{p}}_g} \\ &= \sum_{g\in\mathcal{G}} \left(\sum_{h\in\mathcal{G}} \mathsf{q}_h(t,t+T)\mathsf{p}_{h^{-1}g}(t)\right) \log \frac{\sum_h \mathsf{q}_h(t,t+T)\mathsf{p}_{h^{-1}g}(t)}{\sum_h \mathsf{q}_h(t,t+T)\hat{\mathsf{p}}_g}. \end{aligned}$$

Now by applying the log sum inequality over h for each fixed g we get:

$$\begin{aligned} &\left(\sum_{h\in\mathcal{G}} \mathsf{q}_h(t,t+T)\mathsf{p}_{h^{-1}g}(t)\right) \log \frac{\sum_h \mathsf{q}_h(t,t+T)\mathsf{p}_{h^{-1}g}(t)}{\sum_h \mathsf{q}_h(t,t+T)\hat{\mathsf{p}}_g} \\ &\quad\leq \sum_{h\in\mathcal{G}} \left(\mathsf{q}_h(t,t+T)\mathsf{p}_{h^{-1}g}(t) \log \frac{\mathsf{q}_h(t,t+T)\mathsf{p}_{h^{-1}g}(t)}{\mathsf{q}_h(t,t+T)\hat{\mathsf{p}}_{h^{-1}g}}\right). \end{aligned} \tag{3.26}$$

Furthermore, Assumption 1 allows us: (i) to divide by $\mathsf{q}_h(t,t+T)$; and (ii) in conjunction with the fact that $\sum_g \mathsf{p}_g(t) = 1$ for all t, to exclude the singular cases in Proposition 7. Therefore the equality in (3.26) holds if and only if

$$\frac{\mathsf{q}_h(t,t+T)\mathsf{p}_{h^{-1}g}(t)}{\mathsf{q}_h(t,t+T)\hat{\mathsf{p}}_{h^{-1}g}} = \frac{\mathsf{p}_{h^{-1}g}(t)}{\hat{\mathsf{p}}_{h^{-1}g}}$$

is constant over all $g' = h^{-1}g \in \mathcal{G}$. Since $\sum_{g'\in\mathcal{G}} \mathsf{p}_{g'}(t) = \sum_{g'} \hat{\mathsf{p}}_{g'} = 1$ for every t, the equality holds if and only if $\mathsf{p}(t) = \hat{\mathsf{p}}$. Returning to the sum over g, we thus get

$$0 \leqslant K(\mathsf{p}(t+T)\|\hat{\mathsf{p}}) \leqslant K(\mathsf{p}(t)\|\hat{\mathsf{p}}) \tag{3.27}$$

and each equality holds if and only if $\mathsf{p}(t) = \hat{\mathsf{p}}$. Henceforth the Lyapunov function $K(\mathsf{p}(t)\|\hat{\mathsf{p}})$ strictly decreases after any T steps, as the requirement $q_h(t,t+T) > \delta$ ensures that for any given $\mathsf{p}(t) \neq \hat{\mathsf{p}}$, we get in (3.26) a strict contraction factor independent of $\mathsf{s}(t)$. This ensures, by Lyapunov arguments, that the system asymptotically converges to $\mathsf{p} = \hat{\mathsf{p}}$. □

As an immediate corollary, we have symmetrization on $\mathcal{X}$ with the associated actions, for *any* $\mathcal{X}$*, any linear group action and any* $\mathsf{s}(t)$ *satisfying Assumption 1.*

Corollary 2. Any algorithm of the form (2.18) on a vector space $\mathcal{X}$ with $\mathsf{s}(t)$ satisfying Assumption 1, asymptotically converges to $\lim_{t\to+\infty} x(t) = \bar{\mathcal{F}}(x(0))$.

3.1.2 Examining Switching Signals

Let us now provide some typical examples of switching signals $\mathsf{s}(t)$ and check if they satisfy Assumption 1. It is actually instructive to start by listing some cases that lead to a violation of the assumption.

- If (possibly after some initial transient) the vector $\mathsf{s}(t)$ contains a single nonzero entry at any time, then $\mathsf{q}(t,T)$ will also contain a single element.
- Consider that (after some initial transient) $\mathsf{s}_g(t)$ can be nonzero at any time only for $g \in \mathcal{S}$, a *subgroup* of $\mathcal{G}$. Then each $\tilde{M}(t)$ is a weighted sum of P_g with $g \in \mathcal{S}$, and by subgroup properties the propagator $\prod_{i=0}^{t-1} \tilde{M}(i)$ is also a weighted sum of P_g with g restricted to $\mathcal{S}$, such that we can have $q_g(t,T) \neq 0$ for at most all $g \in \mathcal{S}$.
- More generally, if $\mathsf{s}_g(t)$ can be nonzero at any time only for $g \in \mathcal{S}$, now being some subset of $\mathcal{G}$, and the elements of $\mathcal{S}$ do not generate the whole group, then Assumption 1 cannot hold.

Conversely, sufficient conditions for Assumption 1 to hold include the following.

- If there exists a set $\mathcal{J} \subset \mathcal{G}$ that generates $\mathcal{G}$ and such that for each t, there exists $i \in [t, t+T]$ such that $\mathcal{S}_i = \{g \in \mathcal{G} : \mathsf{s}_g(i) > \delta\}$ contains $\mathcal{J} \cup \{\mathrm{e}_\mathcal{G}\}$, then Assumption 1 is satisfied.
- If $\mathcal{G}$ is Abelian, then the order in which the group elements are selected has no importance, but it is still relevant to know which ones are selected at the same time or not. Then we can use a reduced Cayley graph to investigate Assumption 1 as follows. For each time t, take the set $\mathcal{S}_t = \{g \in \mathcal{G} : \mathsf{s}_g(t) > \delta\}$, choose one $\bar{g}_t \in \mathcal{S}_t$ and let $\bar{\mathcal{S}}(t) = \{\bar{g}_t^{-1} g : g \in \mathcal{S} \setminus \{\bar{g}_t\}\}$. Then consider a starting time t_0 and recursively construct a graph as follows. Start with a single node $\mathrm{e}_\mathcal{G}$. At each step $i = 1, 2, ..., T$, add edges (and potentially vertices) to connect every vertex $h \in \mathcal{G}$ that is already present in the graph at step $i-1$, with the set of nodes $\{s\,h : s \in$

$\bar{\mathcal{S}}_{t_0+i}\}$. If for all t we have $\mathsf{s}_{\mathsf{e}_{\mathcal{G}}}(t) > \delta$, and for all t_0 the graph obtained at $i = T$ contains all the $g \in \mathcal{G}$, then Assumption 1 is satisfied.

3.2 Stochastic Convergence

So far we have always formulated convergence properties for a given switching signal $\mathsf{s}(t)$. We now briefly indicate how they can be adapted when $\mathsf{s}(t)$ is selected at random. We thus consider that at each time t, $\mathsf{s}(t)$ is selected from a set $\mathfrak{S}$ according to some given probability distribution, independently of the $\mathsf{s}(i)$ for $i \neq t$. In other words, the $\mathsf{s}(t)$ are independent, not necessarily identically distributed, random variables over a set of vectors of convex weights. Then we get the following convergence result.

Theorem 2. Assume that there exist some fixed values of T,δ, and $\varepsilon > 0$ for which the statement of Assumption 1 holds with probability at least ε at each time t. Then for any $\gamma > 0$, the probability of having an Euclidean distance $\|\mathsf{p}(t) - \hat{\mathsf{p}}\| < \gamma$ converges to 1 as t converges to $+\infty$.

Proof. On finite-dimensional space $\mathbb{R}^{|\mathcal{G}|}$, the Euclidean distance can be bounded by a monotone function of the Kullback-Leibler pseudo-distance, so it is in fact sufficient to show that $K(\mathsf{p}(t)\|\hat{\mathsf{p}}) < \gamma$ with probability 1. Consider any particular sequence $\mathsf{s}(t)$ that satisfies Assumption 1 for some given T, δ. By Theorem 1 and since $\mathsf{p}(t)$ evolves in a compact set (it has a finite number $|\mathcal{G}|$ of elements, each belonging to $[0, 1]$), we know that for any $\gamma > 0$, there must exist some finite integer N such that

$$K(\mathsf{p}(N \cdot T)\|\hat{\mathsf{p}}) < \gamma \tag{3.28}$$

for all $\mathsf{p}(0)$. Moreover, since the set of possible $\mathsf{s}(t)$ for each t is compact, one can find for any γ a value of N such that (3.28) holds over all sequences that satisfy Assumption 1 with fixed T, δ. Note that, since $K(\mathsf{p}(t)\|\hat{\mathsf{p}})$ is a Lyapunov function, (3.28) guarantees that $K(\mathsf{p}(t)\|\hat{\mathsf{p}}) < \gamma$ for all $t \geq N \cdot T$.

Now consider a randomly chosen sequence $\mathsf{s}(t)$ of $B \cdot N \cdot T$ elements, with $B > 1$. By hypothesis, the probability of $\mathsf{s}(t)$ satisfying Assumption 1 on each interval of length T is greater than ε. Therefore, noting that the sequences $(\mathsf{s}(t))_{t\in[kT,\,(k+1)T)}$ for $k = 0, 1, ..., N-1$ are mutually independent, the probability that we have selected a sequence $\mathsf{s}(t)$ which satisfies Assumption 1 over the $N \cdot T$ first elements — i.e. guaranteeing that (3.28) is satisfied — is at least ε^N. The probability that a sequence of $B \cdot N \cdot T$ elements contains no subsequence of $N \cdot T$ consecutive elements satisfying Assumption 1, is at most $(1-\varepsilon^N)^B$; the latter converges to 0 as B goes to ∞. But if a sequence satisfies Assumption 1 over some interval $[T_0,\ T_0+N\cdot T]$, then (3.28) applies to the system starting at time T_0 with initial state $\mathsf{p}(T_0)$, thus ensuring $K(\mathsf{p}(t)\|\hat{\mathsf{p}}) < \gamma$

for all $t \geq T_0 + N \cdot T$. Altogether, we get that as B goes to ∞, there is a probability 1 that a selected sequence of $B \cdot N \cdot T$ elements satisfies $K(\mathsf{p}(t)\|\hat{\mathsf{p}}) < \gamma$ for all $t \geq B \cdot N \cdot T$ and all $\mathsf{p}(0)$, which is our claim. □

Let us briefly discuss some examples of stochastic evolutions.

- If at each time, we randomly select a single element $h(t)$ from $\mathcal{G}$ with probability of $h(t) = g$ being greater than zero for all g, and take

$$\begin{aligned} \mathsf{s}_{h(t)}(t) = \alpha\,, \quad \mathsf{s}_{\mathrm{e}_\mathcal{G}}(t) = (1-\alpha)\,, \\ \mathsf{s}_g(t) = 0 \quad \text{for } g \notin \{h(t), \mathrm{e}_\mathcal{G}\}\,, \end{aligned} \tag{3.29}$$

  then the requirements of Theorem 2 are clearly satisfied. Of course this situation directly generalizes to cases where more than one $h(t) \in \mathcal{G}$ is applied at each time.

- Like in the deterministic case, a similar result is obtained if in (3.29) we randomly select $h(t)$ from some subset $\mathcal{S}$ of $\mathcal{G}$, and this subset generates the whole group. The subset may also vary (e.g. cyclically) with time, as long as it allows with nonzero probability to construct one sequence satisfying assumption 1. The linear gossip algorithm fits in this category, as the connected graph condition in Propositions 2 and 3 ensure that swaps of adjacent agents can be selected in a way that generates the whole group of permutations.

A few remarks are in order.

Remark 4 (Time-varying possibilities)**.** Theorem 2 only requires some uniform upper bound T on a time interval that guarantees that all group elements are associated with weights of at least $\delta > 0$. It thus allows for dynamics where $\mathsf{p}(t)$ does not evolve towards $\hat{\mathsf{p}}$ for shorter time intervals, as long as there is a nonzero probability to reduce the distance from $\hat{\mathsf{p}}$ in finite time. Therefore, we can ensure convergence if, for example, one strictly contractive evolution is applied only every T_0 steps, while we do not know how s_g is selected in between.

Remark 5 (Explicit robustness to α)**.** A major contribution of Theorem 2 is to establish the *robustness* of consensus-like algorithms with respect to uncertainties in the values of $\mathsf{s}_g(t)$ for a wide variety of applications (see Section 4). Indeed, if we consider that the $h \in \mathcal{S}$ for which $\mathsf{s}_h \neq 0$ are chosen deterministically, but the values $\mathsf{s}_h(t)$ are randomly chosen in some compact set strictly

inside $[0,1]$ for all t, then Assumption 1 holds with given T either for all such sequences or for none; in the former case, compactness ensures that δ is bounded from below, and Theorem 2 holds. This shows that it is not important to control the exact proportions in which the chosen actions are applied. Typically in a gossip algorithm [12], one uses the maximally mixing value $\alpha = 1/2$. Nonetheless, convergence holds provided that $\alpha(t) \in [\underline{\alpha}, \overline{\alpha}] \subset (0,1)$ for all t. Of course, the choice of $\mathsf{s}(t)$ can severely affect convergence *speed*, but this discussion goes beyond the scope of the present paper.

Remark 6. In relation with Assumption 1, it is useful to work with sequences satisfying (with a given non-zero probability) $\mathsf{s}_{\mathrm{e}_\mathcal{G}}(t) \geq \beta$ at any t for some constant $\beta > 0$. Indeed, this ensures that once $q_g(t, t+t_1) \geq \delta' > 0$ for some $t_1 \leq T$, we have $q_g(t, t+T) \geq \delta = \delta'\beta^{T-t_1}$. Most results in linear consensus [20, 23, 17] explicitly make this assumption. Not assuming $\mathsf{s}_{\mathrm{e}_\mathcal{G}}(t) \geq \beta > 0$ for all t generally makes it necessary to perform a detailed analysis of the successions in $\mathsf{s}(t)$ in order to ensure Assumption 1.

4 Applications

We now illustrate a variety of applications covered by our framework by starting with consensus-type problems and next considering more general symmetrization problems. This list is by no means assumed to be exhaustive, and we are confident that more areas of application will be identified.

4.1 Linear Consensus

The gossip algorithm of Sfection 1.1.4 is one basic application of our framework. The group-theoretic language also encompasses other basic linear algorithms for average consensus of m subsystems in $\mathbb{R}^n$.

The most standard consensus algorithm implements, at each time, a motion of each subsystem towards the average of its neighbors in an *undirected graph* $G = (V, E(t))$. Thus the edges of G model a set of interactions that are all simultaneously active. This corresponds to setting $\mathsf{s}_g(t) \neq 0$ for $g = e$ and for all $g \in \mathfrak{P}_V$ that model a pairwise permutation of two agents linked by an edge in $E(t)$; gossip, with a single edge active at a time and hence only two nonzero elements in $\mathsf{s}_g(t)$, is just a particular case[1].

In the group-theoretic formulation, there seems no reason to limit our algorithmic building blocks to pairwise permutations. Including more general permutations allows one to cover situations with explicit multipartite interactions, *e.g.* where subsystem 1 forwards its value to 2, who simultaneously transmits its value to 3, and so on. Selecting $\mathsf{s}_g \neq 0$ specifically for g corresponding to such situations, allows to model linear consensus with *non-symmetric* state transi-

[1] We recall that, since the actions associated to $\mathfrak{P}_V$ in standard consensus are not linearly independent, this is not the only way to lift the consensus dynamics to the permutation group.

tion matrix $A(t)$. The resulting $A(t)$ however will still be doubly-stochastic for any s. As proved by Birkhoff [24], any doubly stochastic matrix can be decomposed as a convex sum of permutations. The corresponding network structure is called a *balanced directed graph* [5], and one could argue that the interpretation as a sum of general permutations gives a sensible rationale as why a graph might be ensured to be balanced in the consensus context. In this sense, any consensus algorithm on a balanced directed graph can be seen as a generalization of a gossip-type algorithm. Convergence, independently of the particular application, is guaranteed if Assumption 1 is satisfied.

4.2 Gossip Symmetrizing Probability Distributions

Consider a collection of m subsystems, each one possessing a random variable y_j on the same outcome set Y, for $j = 1, 2, \ldots, m$. We denote $\mathbb{P}$ the joint probability distribution of the y_j. In order to maintain a compact notation we will consider Y countable, but the uncountable case does not present additional technical difficulties. We are interested in symmetrizing the joint probability distribution, i.e. attaining a distribution $\hat{\mathbb{P}}$ such that

$$\begin{aligned} &\hat{\mathbb{P}}[y_1 = a_1, ..., y_j = a_j, ..., y_k = a_k, ..., y_m = a_m] \\ &= \hat{\mathbb{P}}[y_1 = a_1, ..., y_j = a_k, ..., y_k = a_j, ..., y_m = a_m] \end{aligned} \tag{4.1}$$

for all choices of j, k and of the considered outcomes $\{a_i\}$. The invariance then also holds for general permutations in $\mathfrak{P}_m$. We want to achieve this in a distributed way, where at each time t a reduced set $E(t)$ of pairwise interactions are available.

Our framework suggests the following randomized way to perform this task. At each time t a pair (j, k) is selected from $E(t)$, the random variables at these locations are swapped with probability α, and remain in place with probability $1-\alpha$. This random action still leaves $y_j(t+1), y_k(t+1)$ two random variables on Y, but their probability distributions have changed: e.g. the new random variable $y_j(t+1)$ at location j follows the marginal distribution of $y_j(t)$ with probability $1-\alpha$, or it follows the marginal distribution of $y_k(t)$, with probability α. Overall, *not knowing whether the random variables have been exchanged or not*, the resulting probability distribution for the $y_i(t+1)$, $i = 1, 2, ..., m$

writes:

$$\begin{aligned}\mathbb{P}_{t+1}[y_1 = a_1, ..., y_j = a_j, ..., y_k = a_k, ..., y_m = a_m] \;=\\ (1-\alpha)\,\mathbb{P}_t[y_1 = a_1, ..., y_j = a_j, ..., y_k = a_k, ..., y_m = a_m]\\ +\alpha\,\mathbb{P}_t[y_1 = a_1, ..., y_j = a_k, ..., y_k = a_j, ..., y_m = a_m]\end{aligned} \tag{4.2}$$

In the group symmetrization picture, this framework (goal (4.1) and dynamics (4.2)) corresponds to the exact same setting as standard gossip consensus, with $\mathcal{G} = \mathfrak{P}_m$ the group of permutations on m objects. Only the action is different, now implementing a swap on probability distributions (*including all correlations with other random variables* than the ones involved in the swap), instead of a swap of real numbers.

4.3 Randomized Discrete Fourier Transform

The above applications all involve permutations as the underlying group, in the context of a network of subsystems. We now show how a different group structure can cover the discrete Fourier transform, although we do not directly see a practical use for the resulting algorithm.

The discrete Fourier transform of a (column) vector $x = (x_0, x_1, ..., x_{N-1}) \in \mathbb{C}^N$ is the (column) vector $\chi = (\chi_0, \chi_1, ..., \chi_{N-1})$ with

$$\chi_k = \frac{1}{N}\sum_{n=0}^{N-1} e^{-i\,\frac{k\,n\,2\pi}{N}}\,x_n \quad \text{for } k = 0, 1, ..., N-1\,, \tag{4.3}$$

up to normalization[2]. The complex numbers $\{e^{i\,k\,2\pi/N} : k = 0, 1, ..., N-1\}$ characterizing the Fourier transform form a faithful representation of the cyclic group of order N, that is the Abelian group generated by a single element $\bar{g}$,

$$\mathcal{G}_{c,N} = \{\,\mathrm{e} = \bar{g}^0 = \bar{g}^N, \bar{g}, \bar{g}^2, \bar{g}^3, ..., \bar{g}^{N-1}\,\}\,.$$

We next show how the computation of (4.3) can be obtained as a byproduct of a symmetrization task with respect to an action of $\mathcal{G}_{c,N}$.

It is convenient to consider the vector space $\mathbb{R}^{N\times N}$ and associate to the (column) vector $x \in \mathbb{R}^N$ the square matrix $X = x\,\mathbf{1}^T$, where $\mathbf{1}^T$ is the row

[2]Our developments can be extended to functions on finite Abelian groups, with the Fourier transform defined on characters.

vector of ones. To $\bar{g} \simeq e^{i\,2\pi/N}$ we associate the group action $a(\bar{g}, \cdot) = Q(\cdot)$ defined by:

$$X \;\mapsto\; Q(X) = \sigma\, X\, D^{-1} \tag{4.4}$$

$$\text{with} \qquad D = \mathrm{diag}(1,\, e^{i\,2\pi/N},\, e^{i\,4\pi/N}, ..., e^{i\,(N-1)2\pi/N})$$

$$\sigma = \begin{pmatrix} 0 & 1 & 0 & 0 & ... & 0 \\ 0 & 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 0 & 1 & ... & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & ... & 1 \\ 1 & 0 & 0 & 0 & ... & 0 \end{pmatrix} .$$

The action corresponding to a general group element is obtained by composition. Direct computation shows that the m, n element of $\hat{X} = \frac{1}{N} \sum_{k=0}^{N-1} Q^k(X)$, resulting from the symmetrization of X under the action Q, equals

$$\hat{X}_{[m,n]} \;=\; \frac{1}{N} \sum_{k=0}^{N} x_{(m+k \bmod (N-1))}\; e^{-i\frac{2\pi k}{N} n} .$$

Hence symmetrization under this action of $\mathcal{G}_{c,N}$ gives the Fourier transform of x as:

$$\chi^T = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \end{bmatrix} \hat{X} .$$

The robust convergence of algorithm (3.15) thus indicates that the Fourier transform does not necessarily have to be computed in an orderly fashion, but can asymptotically result from rather arbitrary convex combinations of the actions Q^k with different k, as long as the $\mathsf{s}(t)$ ensure sufficient mixing. Note that the actions $\{Q^0, Q^1, ..., Q^{N-1}\}$ are all linearly independent, so the map from dynamics on group actions to dynamics on p is one-to-one.

4.4 Random State Generation

A variety of applications require to generate random numbers, codewords or, more generally, *states* with a target probability distribution. This includes among others the Markov chain Montecarlo methods [25] as well as classical and quantum cryptography protocols [26]. A basic probability distribution is the uniform or Haar measure on compact sets. Random sample generators must hence be able to transform some *generic* source of randomness – i.e. not necessarily uniform nor in fact exactly known – into a (almost) *uniform* probability distribution. There are various ways of doing this, and our framework points to a particular class of so-called random circuits [27, 28]. Indeed, group

symmetrization provides a robust way to obtain a uniform distribution on a finite set of states $\mathcal{Y}$ that are linked by a group of transformations $\mathcal{G}$, if we can pick elements of $\mathcal{G}$ with some generic probability distribution.

More precisely, consider a finite group $\mathcal{G}$, and its linear action $a(g,\cdot)$ on a vector space $\mathcal{X}$. For some fixed $y_{\mathrm{e}} \in \mathcal{X}$, consider its *orbit*, i.e. the set $\mathrm{Orb}_{\mathcal{G}}(y_{\mathrm{e}}) = \{y_g = a(g, y_{\mathrm{e}}),\, g \in \mathcal{G}\}$. We want to generate a state $y(T)$ that is uniformly (pseudo-)randomly distributed over $\mathrm{Orb}_{\mathcal{G}}(y_{\mathrm{e}})$, by passing a deterministic $y(0) \in \mathrm{Orb}_{\mathcal{G}}(y_{\mathrm{e}})$ through a sequence of (pseudo-)random operations, labeled for convenience by time $t = 0, 1, ..., T-1$. Each operation is associated to a $g(t) \in \mathcal{G}$, drawn according to some possibly unknown probability distributions $\mathsf{s}_g(t)$, mutually independent at each time. We make the technical assumption that $g \neq h \Rightarrow a(g, y(0)) \neq a(h, y(0))$ i.e. $|\mathrm{Orb}_{\mathcal{G}}(y_{\mathrm{e}})| = |\mathcal{G}|$.

As y propagates through the sequence according to $y(t+1) = a(g(t), y(t))$, the probability $\mathsf{p}_h(t)$ to have $y(t+1) = a(h, y(0))$ follows dynamics (3.15). Hence according to Theorem 1, it is sufficient that $\mathsf{s}(t)$ allows to satisfy Assumption 1 to ensure that the distribution of $y(T)$ converges to the *uniform* distribution over $\mathrm{Orb}_{\mathcal{G}}(y_{\mathrm{e}})$ as $T \to \infty$. Note that for a fixed circuit distribution $\mathsf{s}_g(t)$, we indeed apply Theorem 1 as we are modeling the *deterministic* evolution (as t increases) of a probability distribution.

We will come back on the generation of random state in Section 8.6 where we will illustrate a proof of principle on how to achieve an almost quadratic speed-up by perform a suitable quantum walk on the set of unitary transformations.

Remark 7. In addition to finite groups, the case in which $\mathcal{G}$ becomes a continuous Lie group is of great interest for practical applications, including quantum information and more specifically random quantum circuit theory [27, 28]. In that framework, the space of interest is associated to a register of N quantum bits, so that $\mathcal{X} \cong \mathbb{C}^{2^N}$; the group of physically relevant unitary evolutions for the register, or *gates*, is $\mathcal{G} = SU(2^N)$. The finite group setting can effectively approximate such continuous distribution by considering a sufficiently dense subset of the Lie group. It is well known [26] that there exist finite *universal sets* of gates which generate a mathematically dense subset of $SU(2^N)$; ensuring $\mathsf{s}_g(t) > 0$ on such a universal set, is sufficient to satisfy Assumption 1 for any finite subset of a dense subset of $SU(2^N)$.

Introduction to Quantum Consensus

Exploring the links between information processing tasks and stochastic dynamics on networks has recently opened new research directions towards "distributed" quantum information applications. In essence, these involve an interplay between symmetry, locality constraints and engineered dissipation. These are the key ingredients in many quantum information applications, among which noise protection and dynamical error-correction [29, 30, 31, 32], open-system quantum simulators [33, 34] and quantum computers [35], entanglement generation through stabilizing dissipative dynamics [36, 37] as well as most tasks in the stabilization of open multipartite quantum systems [38, 39, 40].

In this second part we are going to present a detailed application of our symmetrizing framework to the consensus problem in the quantum domain focusing in particular on the gossip algorithm. For most of the results regarding the convergence of our algorithms we provide independent proofs using the tools of quantum mechanics. We begin identifying and characterizing a hierarchy of quantum consensus situations and study how these can be reached by suitable dissipative quantum dynamics, while preserving some global information on the network state. Our results tie the structure of symmetric states and correlations [41, 42] to their potential generation via locality constrained resources, similarly in spirit to what has been recently done by characterizing another relevant class of states, namely frustration-free ground states of quasi-local Hamiltonians with dissipative generators [37]. In addition, the ideas and methods we present can be directly employed in order to symmetrize the state of a large system towards permutation-invariant statistics, guarantee effective sampling from large networks of quantum systems, achieve robust broadcast

of information, or realize purification and cooling with limited resources. More details about three possible direct applications of our gossip-type algorithm are presented in Chapter 8.

An attempt to lift the consensus problem to the quantum domain has been presented in [43]. It is based on a "cone geometry" approach, viewing quantum Kraus maps as the non-commutative generalization of Markov chains that model consensus algorithms. The authors show how Birkhoff's Theorem and Hilbert's projective metric lead to a general convergence result and contraction ratio estimation. However, by describing the dynamics of the whole system of interest as governed by a single Markov transition mechanism, this formulation does not account for subsystem structure or network connections in the quantum setting. It therefore defines consensus as asymptotic convergence to a scalar multiple of the identity: for quantum states this corresponds to a fully mixed, most uncertain state which is rarely the desired target for applications.

5

Quantum Essentials

In the second part of this dissertation we will apply the framework developed in the first part to the quantum domain. We begin by reviewing how to model a quantum systems, see e.g. [38, 44, 39, 45, 26]. Despite most of the results we are going to show still holds in the case of a finite network of infinite dimensional quantum systems we will focus on the finite dimensional. As we are going to show, a finite dimensional system present the advantage that can be tackled with linear-algebric tools the same employed in classic system-theory.

In the standard formulation of quantum mechanics [26], a quantum system is described in a separable complex Hilbert $\mathcal{H}$. If the system variables we want to model can assume only a finite set of values it is sufficient to consider a finite dimensional Hilbert space. We denote the complex conjugate of $w \in \mathbb{C}$ as $\bar{w}$.

An Hilbert space is equipped by definition with an hermitian inner product $\langle\ ,\ \rangle_{\mathcal{H}} : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ such that:

- $\langle x\,, y\rangle_{\mathcal{H}} = \overline{\langle y\,, x\rangle}_{\mathcal{H}}\ \forall x, y \in \mathcal{H}$.
- $\langle x\,, \alpha y + \beta z\rangle_{\mathcal{H}} = \alpha\langle x\,, y\rangle_{\mathcal{H}} + \beta\langle x\,, z\rangle_{\mathcal{H}} \quad \forall\, \alpha, \beta\ \in \mathbb{C}$ and $\forall\, x, y, z\ \in \mathcal{H}$.
- $\langle x\,, x\rangle_{\mathcal{H}} \geqslant 0\ \forall\, x\ \in \mathcal{H}$ and the equality holds if and only if x is the null vector.

In case it is clear which Hilbert space we are considering we are going to omit the subscript $\mathcal{H}$. Form the first requirement and the linearity in the second argument ti follows that

$$\langle \alpha y + \beta z\,, x\rangle_{\mathcal{H}} = \bar{\alpha}\langle y\,, x\rangle_{\mathcal{H}} + \bar{\beta}\langle z\,, x\rangle_{\mathcal{H}} \qquad \forall\, \alpha, \beta\ \in \mathbb{C} \ \text{ and } \ \forall\, x, y, z\ \in \mathcal{H}. \tag{5.1}$$

The set of linear operators on $\mathcal{H}$ is a complex vector space and is naturally equipped with an hermitian inner product, the so called Hilbert-Schmidt defined as:

$$\langle X, Y \rangle_{\mathfrak{B}(\mathcal{H})} = \mathrm{Tr}[X^\dagger Y], \tag{5.2}$$

the latter is well defined in finite dimension, hence $\mathfrak{B}(\mathcal{H})$ is an Hilbert space itself.

The *adjoint operator* of $X \in \mathfrak{B}(\mathcal{H})$ is that unique operator $X^\dagger \in \mathfrak{B}(\mathcal{H})$ that satisfies:

$$\langle Xy, z \rangle_{\mathcal{H}} = \langle y, X^\dagger z \rangle_{\mathcal{H}} \quad \forall\, y, z \in \mathcal{H}. \tag{5.3}$$

If an operator X is such that $X = X^\dagger$ it is called *self-adjoint* (or *hermitian*). We denote the subset of self adjoint operators with $\mathfrak{H}(\mathcal{H})$. We denote the norm induced by the inner product as $||x||_{\mathcal{H}} = \sqrt{\langle x\,, x \rangle}_{\mathcal{H}}$.

The *outer product* of $x, y \in \mathcal{H}$ is defined as $xy^\dagger$ denoting the linear operator over $\mathcal{H}$ such as:

$$(xy^\dagger)z := x\langle y, z \rangle_{\mathcal{H}} \quad \forall\, z \in \mathcal{H}. \tag{5.4}$$

Let us now briefly recover some well known results concerning the projections on an Hilbert space space.

Definition 9 (Projector)**.** A linear operator Π, is called projector if:

$$\Pi^2 = \Pi \qquad \Pi^\dagger = \Pi \tag{5.5}$$

i.e. it is self adjoint and idempotent.

From the definition it follows that a projector is a positive-semidefinite operator, in fact:

$$\langle x, \Pi x \rangle_{\mathcal{H}} = \langle x, \Pi^2 x \rangle_{\mathcal{H}} = \langle \Pi^\dagger x, \Pi x \rangle_{\mathcal{H}} = ||\Pi x||^2_{\mathcal{H}}. \tag{5.6}$$

Furthermore, it can be easily shown that the eigenvalues of a projector are either equal to 0 or to 1. We say that Π is a *one-dimensional projector* if $\mathrm{rank}[\Pi] = 1$. A set of projectors $\{\Pi\}$ is said to be orthogonal if:

$$\Pi_i \Pi_j = \delta_{i,j} \Pi_j \qquad \Pi_i,\, \Pi_j \in \{\Pi\}. \tag{5.7}$$

a set of projectors is said *complete* if:

$$\sum_i \Pi_i = \mathbb{I}. \tag{5.8}$$

Diract notation: Following a common convention both in the physics literature and in the control literature we adopt the so called Dirac notation. Given an Hilbert space $\mathcal{H}$ the symbol $|\psi\rangle$ is used to denote e vector of $\mathcal{H}$ and is called *ket*, while its conjugate is denoted by $\langle\psi|$ and is called *bra*. In this notation the inner product of $|\phi\rangle$ and $|\psi\rangle$ becomes $\langle\phi|\psi\rangle$, for the outer we have that $|\phi\rangle\langle\psi|$ stands for the linear operator such that:

$$(|\phi\rangle\langle\psi|)|\eta\rangle := |\phi\rangle\langle\psi|\eta\rangle \qquad \forall \ \ |\eta\rangle \in \mathcal{H}. \tag{5.9}$$

Furthermore if $|\langle\psi|\psi\rangle|^2 = 1$ then $|\psi\rangle\langle\psi|$ is a one-dimensional projector.

Matrix represention: For an N-dimesional Hilbert space it holds that $\mathcal{H} \simeq \mathbb{C}^N$, vectors in $\mathcal{H}$ can thus be represented with N-entries complex column vectors and their adjoint with N-entries row vectors with complex conjugate entries. An operator on $\mathcal{H}$ is thus represented by a $N \times N$ complex matrix and the adjoint of an operator corresponds to the conjugate transpose of its matrix representation. Given two operators X, Y we define their commutator as $[X, Y] = XY - YX$. Hence in finite dimension operator and matrix are essentially equivalent.

5.1 Observables

A physically measurable quantity of the system is called *Observable* and it is associated to a self-adjoint operator for which the *spectral theorem* holds. Consider a N-dimensional Hilbert space:

Theorem 3 (Spectral Theorem [45])**.** For every $X \in \mathfrak{H}(\mathcal{H})$ there exist a set of real distinct eigenvalues $\{\lambda_i\}_{i=1}^{K\leqslant N}$ and a set of complete orthogonal projectors $\{\Pi_i\}_{i=1}^{K\leqslant N}$ such that:

$$X = \sum_{i=1}^{K\leqslant N} \lambda_i \Pi_i \tag{5.10}$$

The eigenvalues constitute the possible *outcomes* of a measurement of the observable X, while the Π_i represents the *quantum events*. We will see in a moment how, given the state of the system, quantum events allow for the computation of the probabilities for the various outcomes.

5.2 States

In general the state of a quantum system is represented a *density matrix*, i.e. an operator such that:

- $\rho^\dagger = \rho$,
- $\rho \geqslant 0$, i.e. $\langle \rho x, x \rangle \geqslant 0$ for every $x \in \mathcal{H}$,
- $\mathrm{Tr}(\rho) = 1$.

From the previous requirements it follows that $0 < \mathrm{Tr}[\rho^2] \leqslant 1$. We call the subset of $\mathfrak{B}(\mathcal{H})$ that satisfies the previous constraints $\mathfrak{D}(\mathcal{H})$, i.e.:

$$\mathfrak{D}(\mathcal{H}) = \{\rho \in \mathfrak{B}(\mathcal{H}) : \ \rho^\dagger = \rho,\ \rho \geqslant 0,\ \mathrm{Tr}[\rho] = 1\}. \tag{5.11}$$

The set $\mathfrak{D}(\mathcal{H})$ is convex and compact in $\mathfrak{B}(\mathcal{H})$ and has in general a complicated structure, anyway in the case of a two dimensional quantum system we can formulate a simple parametrization that goes under the name of Bloch sphere, see Section 5.4. A one-dimensional projector is an example of state. In fact, we have seen in the previous section that one-dimensional projectors are self-adjoin, positive-semidefinite and trace one. Furthermore, they are of the form $|\psi\rangle\langle\psi|$ for some unit-norm vector $\psi \in \mathcal{H}$ (up to an irrelevant global phase factor).

Since the set $\mathfrak{D}(\mathcal{H})$ is a convex one, we have that every convex combinations of one dimensional projector still represents a quantum states. Actually, by using the spectral theorem, we can prove the converse result, i.e. that every state admits a convex decomposition in terms of one dimensional projectors. More precisely:

Proposition 8 ([45])**.** For every density matrix ρ there exist a set of real eignvalues $\{\mu_i\}_{i=1}^N$, such that:

$$0 \leqslant \mu_i \leqslant 1 \ \ \forall\, i \qquad \text{and} \quad \sum_{i=1}^{N} \mu_i = 1, \tag{5.12}$$

and a set of complete orthogonal one-dimensional projector $\{\Pi_i\}_{i=1}^N$ such that:

$$\rho = \sum_{i=1}^{N} \mu_i \Pi_i. \tag{5.13}$$

The latter is called *spectral decomposition.*

This decomposition in general is not unique [45]. Recall that an *extreme element* of a convex set is an element that admit only a trivial convex decomposition in terms of the set elements, i.e. a decomposition with a single weight

equal to one. The extreme points of the set $\mathfrak{D}(\mathcal{H})$ are called *pure states*, if a state is not pure is called *mixed state*. It can be proved, see for example [45], that there is a one to one correspondence between pure states and one dimensional projectors:

Proposition 9. Consider $\rho \in \mathfrak{D}(\mathcal{H})$, then the following statements are equivalent:

- ρ is a pure state,
- ρ is a one-dimensional projector,
- $\mathrm{Tr}[\rho^2] = 1$.

Summing up, every quantum state can be written as a convex decomposition of pure states. Pure states, on the other hand, admit only a trivial decomposition and they represent situations in which we have a full knowledge of the system. If the decomposition is not trivial it describes scenarios in which our knowledge of the system is affected by classical uncertainty, i.e. we only have a probabilistic knowledge of which pure state the system is prepared in. The state that maximize this uncertainty is given by $\rho = \frac{\mathbb{I}}{N}$, i.e. all the weight of convex decomposition are equal to $\frac{1}{N}$, the state is called accordingly *maximally mixed state*.

Before closing this section we note that a density operator can be employed also to describes *statistical ensembles* of identical systems.

5.3 Measurements

A *projective* (or von Neumann) observation, or measurement, of a quantum system is characterized by am *observable*, $\sigma \in \mathfrak{H}(\mathcal{H})$. Its spectral decomposition $\sigma = \sum_{j=1}^{d} \lambda_j \Pi_j$, with $d \leq N$ distinct eigenvalues $\{\lambda_j\}$ and projectors onto associated eigenspaces $\{\Pi_j\}$, governs the stochastic outcome of the measurement and the possibly modified state of the system after measurement: having state ρ before the measurement, the latter's outcome will be λ_j with probability $\mathbb{P}_\rho(\lambda_j) = \mathrm{Tr}(\Pi_j \rho) =: p_j$; and if outcome λ_j is obtained, then the state after the measurement is $\rho|_j = \Pi_j \rho \Pi_j \,/\, p_j$. The probability to observe λ'_k in a subsequent measurement of $\sigma' = \sum_{k=1}^{d'} \lambda'_k \Pi'_k$, with eigenvalues $\{\lambda'_k\}$ and projectors onto associated eigenspaces $\{\Pi'_k\}$ that *do not necessarily commute with the* Π_j, is then:

$$\mathbb{P}_{\rho|_j}(\Pi'_k) = \mathrm{Tr}(\Pi'_k \Pi_j \rho \Pi_j) / \, p_j.$$

It follows that the probability of observing the *ordered* sequence of two events first λ_j, then λ'_k, given the initial ρ, is

$$\mathbb{P}_\rho(\Pi_j, \Pi'_k) = \mathrm{Tr}(\Pi'_k \Pi_j \rho \Pi_j).$$

If Π_j and Π'_k do not commute, a different ordering in a sequence of measurements can change the resulting probability. If Π_j and Π'_k do commute, and only then, the joint probability of observing λ_j, λ_k is independent of the measurement order for all ρ, and simplifies to

$$\mathbb{P}_\rho(\Pi_j, \Pi'_k) = \mathrm{Tr}(\Pi'_k \Pi_j \rho).$$

5.4 Qubits and Bloch Representation

A *qubit* is a (generic, abstract) quantum system associated to a two-dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^2$; a standard basis for the latter is conventionally given by two vectors denoted $|0\rangle \simeq [\ 1 \quad 0\ ]^T$ and $|1\rangle \simeq [\ 0 \quad 1\ ]^T$.

The traceless unitary hermitian Pauli operators $\sigma_x, \sigma_y, \sigma_z$ and the identity operator I together form an orthonormal basis for all hermitians operators on $\mathcal{H}$. Explicitly, $\sigma_x = |1\rangle\langle 0| + |0\rangle\langle 1|, \sigma_y = i|1\rangle\langle 0| - i|0\rangle\langle 1|, \sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$. With the standard basis, these are associated to the matrices:

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

These all have eigenvalues $1, -1$.

We briefly illustrate quantum projective measurement for this example.

Assume for instance that the initial state is $\rho = \frac{1}{3}|0\rangle\langle 0| + \frac{2}{3}|1\rangle\langle 1|$ and we perform a measurement of $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1| = \lambda_1\Pi_1 + \lambda_2\Pi_2$, where $\Pi_1 = |0\rangle\langle 0|$ and $\Pi_2 = |1\rangle\langle 1|$. Then we get outcome $\lambda_1 = 1$ with probability $\mathrm{Tr}(\rho\Pi_1) = \frac{1}{3}$, and if that is recorded, we update the state to $\frac{\Pi_1\rho\Pi_1}{\mathrm{Tr}(\Pi_1\rho\Pi_1)} = |0\rangle\langle 0|$. Outcome $\lambda_2 = -1$ will appear with probability $\frac{2}{3}$, and in that case we shall transform the state to $\frac{\Pi_2\rho\Pi_2}{\mathrm{Tr}(\Pi_2\rho\Pi_2)} = |1\rangle\langle 1|$. Consider now the same initial state ρ but we perform a measurement of σ_x. Then similar calculations yield that we get outcome $\lambda_1 = 1$ with probability $\mathrm{Tr}(\rho\Pi_1) = \frac{1}{2}$, with now $\Pi_1 = \frac{1}{2}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|)$, while the post-measurement state gets updated to Π_1; and with probability $\mathrm{Tr}(\rho\Pi_2) = \frac{1}{2}$ we get outcome $\lambda_2 = -1$, with $\Pi_2 = \frac{1}{2}(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)$, and update the state to Π_2. If we perform a measurement of I instead, then we always get the unique result 1 and the

state ρ does not change. For systems on higher-dimensional Hilbert spaces, measurements associated to degenerate operators can project the state to a subspace of dimension > 1, leading after measurement to a modified state which depends on the initial state. This is always the case when carrying out a measurement on one part of a multipartite quantum system.

Keeping into account that $\mathrm{Tr}[\mathbb{I}_2\rho] = 1$ and that $\{\mathbb{I}_2, \sigma_x, \sigma_y, \sigma_z\}$ form an orthonormal basis for the Hermitian operators on $\mathcal{H}$ we have that each qubit state can be written in the so called *Bloch representation*, i.e.:

$$\rho = \frac{1}{2}(\mathbb{I} + x\sigma_x + y\sigma_y + z\sigma_z), \tag{5.14}$$

with $x, y, z \in \mathbb{R}^3$. The eigenvalues are given by:

$$\lambda_\pm = \frac{1}{2}(1 \pm \sqrt{x^2 + y^2 + z^2}). \tag{5.15}$$

Hence, in order for ρ to be a positive-semidefinite operator it must be $\sqrt{x^2 + y^2 + z^2} \leqslant 1$. Let us define the *Bloch vector* as $r = (x, y, z) \in \mathbb{R}^3$ we have that there is an one to one correspondence between the vectors of the three dimensional unit sphere and the set of states of a two dimensional quantum system, i.e. that the Bloch vector completely characterize the state. If the Bloch vector has unit norm it represent a pure state, the state has in fact a trivial canonical convex decomposition. Hence, the pure states of a qubit occupy the surface of the there dimensional unit sphere. We just mention that in larger dimensions such characterization is more complicated [45], for example it is no longer true that the boundary of $\mathfrak{D}(\mathcal{H})$ is composed only by pure states.

5.5 Multipartite Systems and Partial Trace

For simplicity, we present the interaction of two quantum systems; the case of $n > 2$ systems is easily obtained by iteration. If two quantum systems, with associated Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, are taken together to form a larger bipartite quantum system, the Hilbert space $\mathcal{H}_{1,2}$ associated to the composite quantum system is the tensor product of the individual quantum subsystem Hilbert spaces, $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Let $\{|\psi_k\rangle\}_{k=1}^{n_1}$ and $\{|\phi_l\rangle\}_{l=1}^{n_2}$ be orthonormal bases for $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then an orthonormal basis for $\mathcal{H}_{1,2}$ is

$$\{|\psi_k\rangle \otimes |\phi_l\rangle\}_{k,l=1}^{n_1,n_2}\,, \tag{5.16}$$

from which we get that $\dim(\mathcal{H}_{1,2}) = \dim(\mathcal{H}_1)\dim(\mathcal{H}_2) = n_1 n_2$. We use the short notation $|\psi, \phi\rangle := |\psi\rangle \otimes |\phi\rangle$ for any $|\psi\rangle \in \mathcal{H}_1$ and $|\phi\rangle \in \mathcal{H}_2$. The composite Hilbert space is naturally endowed with the inner-product:

$$\langle u_1, u_2 | v_1, v_2 \rangle_{\mathcal{H}_{1,2}} := \langle u_1 | v_1 \rangle_{\mathcal{H}_1} \langle u_2 | v_2 \rangle_{\mathcal{H}_2} . \tag{5.17}$$

A representation and basis for operators in $\mathfrak{B}(\mathcal{H}_{1,2})$ is derived from its vector counterpart in the standard way. In particular, given two operators $X_1 \in \mathfrak{B}(\mathcal{H}_1)$ and $X_2 \in \mathfrak{B}(\mathcal{H}_2)$, one can define $X_1 \otimes X_2 \in \mathfrak{B}(\mathcal{H}_{1,2})$ as the linear operator such that $\forall |u_1\rangle \in \mathcal{H}_1 \, , \; |u_2\rangle \in \mathcal{H}_2$:

$$X_1 \otimes X_2(|u_1\rangle \otimes |u_2\rangle) = X_1|u_1\rangle \otimes X_2|u_2\rangle \quad . \tag{5.18}$$

If two operators are in the form $X_1 \otimes \mathcal{I}_2$ and $\mathcal{I}_1 \otimes X_2$, i.e. they act non-trivially only on different parts of the multipartite system, then they commute for any X_1 and X_2. It is worth noting that in matrix representation, the tensor product corresponds to the Kronecker product.

The partial trace over $\mathcal{H}_1$ is the unique linear map

$$\mathrm{Tr}_{\mathcal{H}_1} : \mathfrak{B}(\mathcal{H}_1 \otimes \mathcal{H}_2) \longrightarrow \mathfrak{B}(\mathcal{H}_2),$$

such that, for any $X_{1,2} \in \mathfrak{B}(\mathcal{H}_{1,2})$ and any $X_2 \in \mathfrak{B}(\mathcal{H}_2)$,

$$\mathrm{Tr}[\mathrm{Tr}_{\mathcal{H}_1}[X_{1,2}]X_2] = \mathrm{Tr}[X_{1,2}(\mathcal{I}_1 \otimes X_2)].$$

If $\{|\psi_k\rangle\}_{k=1}^{n_1}$ and $\{|\phi_l\rangle\}_{l=1}^{n_2}$ are orthonormal bases for $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, the partial trace over $\mathcal{H}_1$ can be written as:

$$\mathrm{Tr}_{\mathcal{H}_1}[X_{1,2}] = \sum_{k=1}^{n_1} \sum_{l,i=1}^{n_2} \langle \psi_k \otimes \phi_l | X_{1,2} | \psi_k \otimes \phi_i \rangle |\phi_l\rangle\langle\phi_i|. \tag{5.19}$$

The partial trace over $\mathcal{H}_2$ can be written in a similar fashion.

5.6 Unitary Quantum Dynamics

Quantum theory postulates that the evolution of a quantum state in an isolated systems is govern by the Liouville-von Neumann equation:

$$\hbar \frac{\partial}{\partial t} \rho_t = [H, \rho_t], \tag{5.20}$$

where H is a self-adjoint operator called *Hamiltonian* and $\hbar$ is the reduced Plank's constant that we set equal to one from now on. Note that the Liouville-von Neumann equation does not allows for measurement operations [26]. This

because such operation would require the coupling with an extenal measurement apparatus violating thus the assumption that the system is isolated.

Suppose that at time $t = 0$ the state of the systems is given by ρ_0 the Liouville-von Neumann equation is formally solved as:

$$\rho_t = \mathrm{e}^{-iHt}\rho_0\mathrm{e}^{+iHt} = U(t)\rho_0 U^\dagger(t), \tag{5.21}$$

where $U(t)$ is a unitary operator $U(t) \in \mathfrak{U}(\mathcal{H})$. The dynamics of a quantum state is thus given by a unitary conjugation. If in particular the initial state is a pure one, say $\rho_0 = |\psi_0\rangle\langle\psi_0|$, it remains pure through the whole evolution and the dynamics can be described by the so called *Schrödinger equation*:

$$\frac{\partial}{\partial t}|\psi_t\rangle = H|\psi_t\rangle, \tag{5.22}$$

In case the Hamiltonian of the system is time dependent the evolution is still unitary and the corresponding operator is computed by mean of the Dyson series [46]:

$$U(t) = \mathcal{T}\exp\left(-i\int_0^t duH(u)\right), \tag{5.23}$$

where $\mathcal{T}$ is the so-called time ordering operator.

So far we have considered the dynamical evolution of the state of the system, this consists in the so called *Schrödinger picture*. In classical mechanics it is usually considered the dynamical evolution of some quantity of the system, such for example the position or the momentum of a classical particle. In quantum mechanics we can establishes a similar point of view given by the so-called *Heisenberg picture*. Consider an observable $X \in \mathfrak{B}(\mathcal{H})$ and assume that $\rho_t = U(t)\rho_0 U^\dagger(t)$, the expectation value of X in the state ρ_t is given by:

$$\mathrm{Tr}[X\rho_t] = \mathrm{Tr}[XU(t)\rho_0 U^\dagger(t)] = \mathrm{Tr}[U^\dagger(t)XU(t)\rho_0], \tag{5.24}$$

where we have used the ciclic property of the trace. Let us define

$$X_t := U^\dagger(t)XU(t), \tag{5.25}$$

the dynamical evolution of X is given by the dual evolution with respect to the Hilbert-Schmidt inner product and it is called *Heisenberg picture*. Equation (5.25) can be viewed as conjugate action of the adjoint unitary operator. Note that the two pictures are equivalent because yield the same expectation values, we have in fact:

$$\mathrm{Tr}[X\rho_t] = \mathrm{Tr}[X_t\rho_0]. \tag{5.26}$$

5.7 General Quantum Dynamics

So fare we have presented how to model the isolated quantum systems. Anyway, in realistic scenarios our system of interest, $\mathcal{H}_{\mathcal{S}}$, is coupled to an environment or reservoir, $\mathcal{H}_{\mathcal{R}}$, whose degree of freedom we may not able to model (or we may not be interested). For this cases it has been devised a reduced description that considers only the degree of freedom of the system by averaging the reservoir using the partial trace. The whole system-reservoir is represented in the Hilbert space $\mathcal{H}_{\mathcal{S},\mathcal{R}} \simeq \mathcal{H}_{\mathcal{S}} \otimes \mathcal{H}_{\mathcal{R}}$, and its evolution is still unitary and given by the operator $U_{\mathcal{S},\mathcal{R}}(t) = \mathrm{e}^{-iH_{\mathcal{S},\mathcal{R}}t}$ where $H_{\mathcal{S},\mathcal{R}}$ is an Hamiltonian that belongs to $\mathfrak{H}(\mathcal{H}_{\mathcal{S},\mathcal{R}})$. Assume now the initial state to be factorized $\rho_0 \otimes \xi \in \mathfrak{D}(\mathcal{H}_{\mathcal{S},\mathcal{R}})$, we consider:

$$\rho(t) := \mathcal{E}_{(0,t)}(\rho_0) = \mathrm{Tr}_{\mathcal{R}}[U_{\mathcal{S},\mathcal{R}}(t)(\rho_0 \otimes \xi)U^{\dagger}_{\mathcal{S},\mathcal{R}}(t)] \tag{5.27}$$

In general this procedure yields a Non-Markovian dynamics for $\rho(t) \in \mathfrak{D}(\mathcal{H}_{\mathcal{S}})$ that involves the computation non-trivial memory kernels [39]. Nonetheless, if some relevant assumptions on the memory time-scale of the reservoir are fulfilled, a set of suitable approximations can be evoked such that the resulting reduced dynamics is described by a Markovian *quantum dynamical semigroup*, see e.g. [39, 47]. More precisely, a *quantum channels* [48, 26] is a linear, completely positive (CP) and trace preserving (TP) maps from density operators to density operators $\mathcal{E} : \mathfrak{D}(\mathcal{H}^m) \to \mathfrak{D}(\mathcal{H}^m)$. It can be shown that such maps admit an *operator sum representation* (OSR), also known as *Kraus decomposition*:

$$\mathcal{E}(\rho) = \sum_{k=1}^{K} A_k \rho A_k^{\dagger} \quad \text{with} \quad \sum_{k=1}^{K} A_k^{\dagger} A_k = I \tag{5.28}$$

where $K \leqslant (\dim(\mathcal{H}))^2$. The representation is not unique, however the relation between all the possible different representations is well known (see [26, Theorem 8.2]).

A *quantum dynamical semigroup* [47] is a family of quantum channels $\{\mathcal{E}_t\}_{t\geqslant 0}$ from $\mathfrak{D}(\mathcal{H})$, such that:

- (Markov property) $\mathcal{E}_t\mathcal{E}_s = \mathcal{E}_{s+t}\ \forall\, t, s \geqslant 0$, i.e. the family is a semigroup.
- $\mathrm{Tr}[\mathcal{E}_t(\rho)X]$ is a continuous function in t for every $\rho \in \mathfrak{D}(\mathcal{H}_{\mathcal{S}})$ and $X \in \mathfrak{H}(\mathcal{H}_{\mathcal{S}})$.

5.8 Notation

In this chapter we have illustrated how to model a finite dimensional quantum and highlighting how this can be done with algebraic tools. In essence, everything can be read more or less verbatim with the following translation table:

$\mathcal{H}$, n-dimensional Hilbert space	$\rightarrow$	$\mathbb{C}^n$
$\|x\rangle \in \mathcal{H}$	$\rightarrow$	column vector, $x \in \mathbb{C}^n$
$\langle x\| \in \mathcal{H}^\dagger$	$\rightarrow$	row vector, $x^\dagger$
X, operator	$\rightarrow$	X, complex matrix

Correspondingly, the (adjoint) $\dagger$ symbol indicates the transpose-conjugate in matrix representation, and the tensor product $\otimes$ is associated to the Kronecker matrix product.

6

Quantum Consensus Definitions and their Relationships

Defining what a consensus situation ought to be in a quantum "network" is not a straightforward task. More than one definition may be appropriate depending on the type of *symmetry* we are seeking. Following the analogy with the classical case can help, but quantum measurement outcomes are intrinsically stochastic, so we must consider *probabilistic* consensus situations from the beginning. Let us explore different options by first discussing a simple case.

Example 1: When is a quantum network in consensus? Consider a multipartite quantum system composed of three qubits, with associated Hilbert space $\mathcal{H}^3 = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2$ spanned by 2^3 basis vectors denoted by $\{|a,b,c\rangle = |a\rangle \otimes |b\rangle \otimes |c\rangle \,:\, a,b,c \in \{0,1\}\}$, and three observables of the form $\sigma^{(1)} = \sigma_z \otimes I \otimes I$, $\sigma^{(2)} = I \otimes \sigma_z \otimes I$, $\sigma^{(3)} = I \otimes I \otimes \sigma_z$, where the Pauli matrix $\sigma_z = \mathrm{diag}(1,-1)$ with respect to the ordered basis $\{|0\rangle, |1\rangle\}$. These correspond to observables of the quantity associated to σ_z for each of the subsystems, i.e. measuring $\sigma^{(3)}$ gives result $+1$ (resp.-1) if the third qubit is in state $|0\rangle$ (resp.$|1\rangle$). It seems natural to say that the system is in consensus with respect to the expectation of σ_z if

$$\mathrm{Tr}(\rho\sigma^{(1)}) = \mathrm{Tr}(\rho\sigma^{(2)}) = \mathrm{Tr}(\rho\sigma^{(3)}). \tag{6.1}$$

The conditions for this to happen can be worked out explicitly in terms of the diagonal elements of the state ρ. In particular it is easy to check that all the

following states satisfy (6.1):

$$\begin{aligned}
\rho^A &= \frac{1}{8} I \otimes (|0\rangle + |1\rangle)(\langle 0| + \langle 1|) \otimes (|0\rangle + |1\rangle)(\langle 0| + \langle 1|); \\
\rho^B &= \frac{1}{2}(|0,0,1\rangle\langle 0,0,1| + |1,1,0\rangle\langle 1,1,0|); \\
\rho^C &= \frac{1}{8} I \otimes I \otimes I; \\
\rho^D &= \frac{1}{2}(|0,0,0\rangle\langle 0,0,0| + |1,1,1\rangle\langle 1,1,1|); \\
\rho^E &= |0,0,0\rangle\langle 0,0,0|; \\
\rho^F &= \frac{1}{2}(|0,0,0\rangle + |1,1,1\rangle)(\langle 0,0,0| + \langle 1,1,1|)\,.
\end{aligned}$$

All these states, except ρ^E, have $\mathrm{Tr}(\rho\sigma^{(i)}) = 0$ for $i = 1, 2, 3$. The states $\rho^B, \rho^C, \rho^D, \rho^E$ are diagonal in the canonical basis and hence can be interpreted as classical probabilities on the set $\{-1, +1\} \times \{-1, +1\} \times \{-1, +1\}$ of possible outcomes for the joint measurements of $\sigma^{(j)}$, $j = 1, 2, 3$.

The requirement (6.1) can be strengthened by requesting it to hold when σ_z is replaced by *any* observable $\sigma \in \mathfrak{B}(\mathbb{C}^2)$ in the definition of $\sigma^{(1)}, \sigma^{(2)}, \sigma^{(3)}$. This is equivalent to imposing that the reduced states for the three subsystems are the same. It is then easy to check that $\rho^B, \rho^C, \rho^D, \rho^E, \rho^F$ satisfy this requirement, while ρ^A does not. In fact the reduced states for ρ^A are:

$$\rho_1^A = \frac{1}{2} I, \quad \rho_2^A = \rho_3^A = \frac{1}{2}(|0\rangle + |1\rangle)(\langle 0| + \langle 1|).$$

In the light of (2.9), another potential definition of quantum consensus would require the *state to be symmetric*, i.e. invariant under any permutations of the subsystems. This choice can be motivated by the classical case, where the consensus state is indeed permutation invariant. Among the states defined in Example 1, only $\rho^C, \rho^D, \rho^E, \rho^F$ are permutation invariant.

Lastly, one might want subsystem agreement not only on the observable averages, but *on each realization of a stochastic measurement (see Appendix 5.4)*; namely, that each projective measurement of the (commuting and hence compatible) observables $\sigma_1, \sigma_2, \sigma_3$ gives *perfectly correlated results* for the three subsystems. Thus among all possible measurement results $\{-1, +1\}^{\times 3}$, one wants that only $(-1, -1, -1)$ and $(+1, +1, +1)$ have a nonzero probability to occur[1]. The states ρ^A, ρ^B and ρ^C do not satisfy this definition of consensus; indeed, for these three states, the distribution of measurement results for qubit 1 is either independent (ρ^A, ρ^C) or anti-correlated (ρ^B) to the measurements

[1]The set $\{c_1, c_2, c_3, ...\}^{\times n}$ is the cartesian product of $\{c_1, c_2, c_3, ...\}$ by itself n times, i.e. the set of n-tuples with components taken from $\{c_1, c_2, c_3, ...\}$.

of at least another qubit. On the other hand, ρ^D, ρ^E, ρ^F always yield perfectly correlated results. Note that mixed states can lead to correlated results, when they express perfect classical correlations as ρ^D does. □

Let us formalize the ideas emerging from the former example. Consider a multipartite system composed of m isomorphic subsystems, labeled with indices $i = 1, \ldots, m$, with associated Hilbert space $\mathcal{H}^m := \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_m \simeq \mathcal{H}^{\otimes m}$, with $\dim(\mathcal{H}_i) = \dim(\mathcal{H}) = n$ and $n \geqslant 2$. We shall refer to this multipartite system as to our *quantum network*. For any operator $X \in \mathfrak{B}(\mathcal{H})$, we will denote by $X^{\otimes m}$ the tensor product $X \otimes X \otimes ... \otimes X$ with m factors. Given an operator $\sigma \in \mathfrak{B}(\mathcal{H})$, we denote by $\sigma^{(i)}$ the local operator:

$$\sigma^{(i)} := I^{\otimes(i-1)} \otimes \sigma \otimes I^{\otimes(m-i)}.$$

Permutations of quantum subsystems are expressed by a unitary operator $U_\pi \in \mathfrak{U}(\mathcal{H})$, which is uniquely defined by

$$U_\pi^\dagger (X_1 \otimes \ldots \otimes X_m) U_\pi = X_{\pi(1)} \otimes \ldots \otimes X_{\pi(m)} \tag{6.2}$$

for any operators $X_1, \ldots X_m$ in $\mathfrak{B}(\mathcal{H})$, where π is a permutation of the set $\{1, \ldots, m\}$ integers, the action of U_π is extended by linearity to the whole $\mathfrak{B}(\mathcal{H}^m)$. Note that expression (6.2) is equivalent to consider the action of the permutation π on the observable $X_1 \otimes \ldots \otimes X_m$ given by:

$$a(\pi, X_1 \otimes \ldots \otimes X_m) = U_\pi^\dagger (X_1 \otimes \ldots \otimes X_m) U_\pi. \tag{6.3}$$

A state or observable is said to be *permutation invariant* if it commutes with all the subsystem permutations. It is worth noting that given any observable $Q \in \mathfrak{H}(\mathcal{H}^m)$, we can define a permutation invariant observable X by considering:

$$X = \frac{1}{m!} \sum_{\pi \in \mathfrak{P}} U_\pi^\dagger Q U_\pi \,. \tag{6.4}$$

Definition 10 (σEC)**.** Given $\sigma \in \mathfrak{B}(\mathcal{H})$, a state $\rho \in \mathfrak{D}(\mathcal{H}^m)$ is in σ*-Expectation Consensus* (σEC) if:

$$\mathrm{Tr}(\sigma^{(1)}\rho) = \ldots = \mathrm{Tr}(\sigma^{(k)}\rho).$$

The reduced state (analog of a marginal distribution) of subsystem k for an overall system state ρ is defined by $\bar{\rho}_k = \mathrm{Tr}_{\left(\bigotimes_{j \neq k} \mathcal{H}_j\right)}(\rho)$.

Definition 11 (RSC)**.** A state $\rho \in \mathfrak{D}(\mathcal{H}^m)$ is in *Reduced State Consensus* (RSC) if

$$\bar{\rho}_1 = \bar{\rho}_2 = \ldots = \bar{\rho}_m.$$

Definition 12 (SSC)**.** A state $\rho \in \mathfrak{D}(\mathcal{H}^m)$ is in *Symmetric State Consensus* (SSC) if, for each unitary permutation U_π,

$$U_\pi \, \rho \, U_\pi^\dagger = \rho \, .$$

Symmetric state consensus is thus equivalent to the symmetrization with respect to the permutation group of m objects with the action given by:

$$a(\pi, \rho) = U_\pi \rho U_\pi^\dagger. \tag{6.5}$$

Definition 13. [σSMC] Given an observable σ with spectral decomposition $\sigma = \sum_{j=1}^{d} s_j \Pi_j \in \mathfrak{H}(\mathcal{H})$,[2] a state $\rho \in \mathfrak{D}(\mathcal{H}^m)$ is in *Single σ-Measurement Consensus* (σSMC) if:

$$\mathrm{Tr}(\Pi_j^{(k)} \Pi_j^{(\ell)} \rho) = \mathrm{Tr}(\Pi_j^{(\ell)} \rho), \tag{6.6}$$

for all $k, \ell \in \{1, \ldots, m\}$, and for each j.

The definition of σSMC requires that the outcomes of σ measurements on different subsystems be exactly the same *for each trial*. Indeed, in this last definition, the right-hand side of (6.6) is the probability of obtaining s_j as a measurement result on both subsystems ℓ and k (note that $\Pi_j^{(k)}$ and $\Pi_j^{(\ell)}$ commute, so this joint measurement $\Pi_j^{(k)} \Pi_j^{(\ell)}$ is well-defined). Then if (6.6) holds, the probability of s_j on k conditional to observing s_j on ℓ is one (assuming that $\Pi_j^{(\ell)} \rho \Pi_j^{(\ell)} \neq 0$; that special case is trivial and can be treated separately).

All the states in our example satisfy σ_zEC, all but ρ^A satisfy RSC, ρ^C to ρ^F satisfy SSC, and ρ^D to ρ^F satisfy σ_zSMC. There obviously seems to be a hierarchy in these definitions, and the following properties are meant to better characterize them.

Theorem 4. The following chain of implications holds:

$$\text{SSC} \implies \text{RSC} \implies \sigma\text{EC},$$

while the converse implications are not true in general.

[2] We here assume that all s_j are different, so $d \leq n$. See Appendix 5.4 for more on observables and related stochastic measurement results.

Proof. SSC $\implies$ *RSC*: If $U_\pi \rho U_\pi^\dagger = \rho$ for each permutation, consider in particular $U_{(\ell,k)}$ that swaps subsystems ℓ and k. Then

$$\bar{\rho}_k = \mathrm{Tr}_{\bigotimes_{j\neq k}\mathcal{H}_j}(\rho) = \mathrm{Tr}_{\bigotimes_{j\neq k}\mathcal{H}_j}(U_{(\ell,k)}\rho U_{(\ell,k)}^\dagger) = \bar{\rho}_\ell,$$

and the reasoning can be repeated for any pair. *RSC* $\implies$ σEC is immediate by definition. States ρ^B and ρ^A from Example 1 provide counterexamples for the converse of the first and of the second implication, respectively.

□

In order to obtain converse relations one has to add some hypotheses:

Proposition 10. The following hold:

1. A state is RSC if and only if it is σEC *for all* $\sigma \in \mathfrak{H}(\mathcal{H})$;

2. If ρ is in RSC, with $\bar{\rho}_k$ a pure state for each k, then it is also in SSC.

The proof is given in Appendix B. We next characterize the notion of σSMC, and explore its relationship with the other notions. Consider the set of projections $\{\Pi_j\}_{j=1}^d$ as in Definition 13, and let us define $\Pi_{\mathrm{sym}} = \sum_{j=1}^d \Pi_j^{\otimes m}$.

Theorem 5. A state is in σSMC if and only if it holds

$$\mathrm{Tr}(\Pi_{\mathrm{sym}}\rho) = 1, \tag{6.7}$$

or equivalently

$$\Pi_{\mathrm{sym}}\rho\Pi_{\mathrm{sym}} = \Pi_{\mathrm{sym}}\rho = \rho. \tag{6.8}$$

Furthermore:

(a) σSMC implies σEC;

(b) σSMC for σ with non-degenerate spectrum implies RSC;

(c) σSMC for σ with non-degenerate spectrum implies SSC;

(d) The converse implications of (a),(b) or (c) do not hold;

(e) It is impossible for a state to be σSMC with respect to all $\sigma \in \mathfrak{H}(\mathcal{H})$.

The proof is given in Appendix B. We thus have, as could be expected, that σSMC is in general a stronger notion of consensus, as long as σ has non-degenerate spectrum.

Remark: It is worth remarking how all these definitions could be given for classical systems, in the context of consensus for random variables or for probability distributions of the state values. In this case, for example, σEC would require the expectation of a set of random variables, each one associated to a subsystem, to be the same in all subsystems; RSC would require the marginal distributions on each subsystem to be equal; and SSC would require that the joint probability distribution is invariant with respect to subsystem permutations.

7

Quantum Evolutions on Networks and their Asymptotic Properties

7.1 Quantum Dynamics and Locality

According to Schrödinger's equation, *isolated* quantum systems evolve unitarily, see Section 5.6 and [44, 26]. However, unitary dynamics are not enough when we are interested in studying or engineering convergence features for a quantum system. A more general framework that includes (Markovian) open-system evolutions is offered by *quantum channels*, see Section 5.7 or [48, 26], that is, linear, completely positive (CP) and trace preserving (TP) maps from density operators to density operators $\mathcal{E} : \mathfrak{D}(\mathcal{H}^m) \to \mathfrak{D}(\mathcal{H}^m)$. It can be shown that such maps admit an *operator sum representation* (OSR), also known as *Kraus decomposition*:

$$\mathcal{E}(\rho) = \sum_{k=1}^{K} A_k \rho A_k^\dagger \quad \text{with} \quad \sum_{k=1}^{K} A_k^\dagger A_k = I \tag{7.1}$$

where $K \leqslant (\dim(\mathcal{H}))^2$. The representation is not unique, however the relation between all the possible different representations is well known (see [26, Theorem 8.2]). A CPTP map is said *unital* if $\mathcal{E}(I) = I$. These maps represent the quantum equivalent of doubly-stochastic transition matrices for Markov processes. Since each step of the gossip algorithm was given by a doubly stochastic matrix, unital maps are thus good candidates as building blocks for a dynamics that asymptotically realizes consensus. A particular set of unital quantum channels is given by random unitaries [49]. A channel belongs to this class when it admits an OSR with K operators $A_k = \sqrt{p_k} U_k$, with $U_k \in \mathfrak{U}(\mathcal{H}^m)$ and

$p_k \geq 0$ such that $\sum_{k=1}^{K} p_k = 1$:

$$\mathcal{E}(\rho) = \sum_{k=1}^{K} p_k \, U_k \rho U_k^\dagger.$$

Such a map can be thought of as a probabilistic mixture of unitary evolutions.

Given a CPTP map $\mathcal{E}$, we can define its dual map with respect to the Hilbert-Schmidt inner product $\mathcal{E}^\dagger : \mathfrak{B}(\mathcal{H}) \to \mathfrak{B}(\mathcal{H})$ through the relation:

$$\mathrm{Tr}[A\,\mathcal{E}(\rho)] = \mathrm{Tr}[\mathcal{E}^\dagger(A)\,\rho]\,. \tag{7.2}$$

This dual map is still linear and completely positive, while the fact that $\mathcal{E}$ is trace preserving implies that $\mathcal{E}^\dagger$ is always unital. Considering the dynamics in the dual picture, i.e. with time-invariant states and maps acting on the observables, is called Heisenberg's picture in the physics literature and provides an equivalent description of quantum system evolution. Note that the adjoint of the unitary action considered in (6.5) is accordingly given by:

$$\mathrm{Tr}[A\,a(U,\rho)] = \mathrm{Tr}[A\,U\rho U^\dagger] = \mathrm{Tr}[U^\dagger A U\,\rho] = \mathrm{Tr}[a^\dagger(U,A)\,\rho] \tag{7.3}$$

We now introduce locality notions for the quantum network. Consider the multipartite system introduced in Section 6: following [37], we say that an operator in $\mathfrak{B}(\mathcal{H})$ is quasi-local if it acts non-trivially only on one neighborhood $\mathcal{N}_j \subseteq \{1, \ldots, m\}$:

Definition 14 (Quantum quasi-local operator)**.** An operator V is quasi-local with respect to a set of neighborhoods $\{\mathcal{N}_j\ ,\ j = 1, 2, ..., M\}$, if and only if there exists $j \in \{1, 2, ..., M\}$ such that:

$$V = V_{\mathcal{N}_j} \otimes I_{\overline{\mathcal{N}_j}} \tag{7.4}$$

where, with a slight abuse of notation, $V_{\mathcal{N}_j}$ accounts for the nontrivial action on $\mathcal{H}_{\mathcal{N}_j}$ and $I_{\overline{\mathcal{N}_j}} = \bigotimes_{k \notin \mathcal{N}_j} I_k$.

7.1.1 Timing of operations and evolution types

As we have seen in classical consensus, an important aspect is that the graph (and the related interaction law) can be time-varying. For instance one can assume that all edges are activated for the whole time (*synchronous update*), at the other extreme that they are activated one at a time, or some at each time (*asynchronous update*), according to some predefined time-varying sequence or

by random selection of edges. Again, convergence properties for all these cases can be linked to the connectedness of the "average graph" [17].

In the quantum case also this distinction can be made. The elementary dynamical interaction that we consider, replacing "one edge" of the classical case, is a CPTP map involving one neighborhood only:

$$\mathcal{E}_{\mathcal{N}_j}(\rho) = \sum_{k=1}^{K} p_k V_k(t) \rho V_k^\dagger(t), \tag{7.5}$$

where all the $V_k(t) \in \mathfrak{U}(\mathcal{H}^m)$ are quasi-local with respect to the neighborhood $\mathcal{N}_j$, $j \in \{1, 2, ..., M\}$. One of the reasons for focusing on this class of evolutions stems directly from applications: methods for implementing unitary evolutions, as well as related unital channels with the aid of some ancillary systems, are available in a number of diverse experimental settings. On the other hand, constructing arbitrary quantum channels is a more challenging task [50], and can be generally done with good approximation only in the limit of fast control and/or short time scales [33]. The building block (7.5) can lead to different evolutions for the whole system, depending on neighborhood selection:

- *Random single interactions:* at each time t one neighborhood $\mathcal{N}_{j(t)}$ is selected at random, $j(t)$ being a single-valued random variable onto the neighborhood index set.

- *Cyclic single interactions:* at each time t one neighborhood $\mathcal{N}_{j(t)}$ is selected deterministically, for example periodically cycling between the available j.

- *Random or cyclic asynchronous interactions:* similar to the previous options, but a subset of several neighborhoods is selected at each time t. We can request the selected neighborhoods to be disjoint or not. This choice may have consequences for the implementation and convergence *speed*, but not for the convergence property of our algorithm, so we will not consider it further.

- *Synchronous interactions:* all the available interactions are activated at each time, weighted by some $q_j \geq 0$ with $\sum_{j=1}^{M} q_j = 1$ to maintain a trace-preserving map:

$$\mathcal{E}(\rho) = \sum_{j=1}^{M} q_j \, \mathcal{E}_{\mathcal{N}_j}(\rho) \, . \tag{7.6}$$

- *Expected evolution:* we study the evolution *in expectation* of the random interaction protocol which selects neighborhood $\mathcal{N}_j$ with probability q_j at each t. Remarkably, the evolution to ρ_{t+1} given ρ_t then follows the same law (7.6) as the synchronous case. Note that convergence of the expected evolution to consensus does not guarantee (at all) that a(ny) single evolution, determined by a realization of the random process $\{j(t)\}_{t\geq 0}$, would converge to consensus. Nevertheless, the statistics of any measurements performed at any time on the system will be exactly the same for (7.6) as for the associated random evolution. On average, convergence in expectation is indistinguishable from trajectory-wise convergence.

The last two cases involve a time-independent map. Another time-independent map is obtained if we consider cyclic interactions of period T and we focus on the state at the end of every cycle:

$$\rho_{t+T} = \mathcal{E}_C(\rho_t) = \mathcal{E}_{\mathcal{N}_T} \circ \ldots \circ \mathcal{E}_{\mathcal{N}_1}(\rho_t). \tag{7.7}$$

The consensus goal can now be specified formally.

Let $d(\rho_a, \mathcal{C}) = \inf_{\rho\in\mathcal{C}} \|\rho_a - \rho\|$, where $\mathcal{C} \subset \mathfrak{D}(\mathcal{H})$ and $\|\cdot\|$ is any p-norm on $\mathfrak{B}(\mathcal{H})$. Given a sequence of channels $\{\mathcal{E}_t(\cdot)\}_{t=0}^{\infty}$, define $\hat{\mathcal{E}}_t(\rho_0) = \rho_t = \mathcal{E}_t \circ \mathcal{E}_{t-1} \circ \cdots \circ \mathcal{E}_1(\rho_0)$, and $\mathcal{C}_{\sigma\text{EC}}$ to be the set of states in σEC consensus.

Definition 15 (Asymptotic Consensus)**.** A sequence of channels $\{\mathcal{E}_t(\cdot)\}_{t=0}^{\infty}$, is said to *asymptotically achieve σEC* if

$$\lim_{t\to\infty} d(\hat{\mathcal{E}}_t(\rho_0), \mathcal{C}_{\sigma\text{EC}}) = 0, \tag{7.8}$$

for all initial states ρ_0.

The same definition holds for RSC, SSC, and σSMC by substituting the corresponding state sets in (7.8). Note that for the SSC, if we are considering a unital map such that the elements of its Kraus decomposition forms a finite group, the previous definition is equivalent to Definition 8. Note that the quantum gossip interaction is composed by a convex combination of actions. Hence, if consensus is asymptotically realized, the algorithm has to converge toward an equilibrium, i.e. the propagator converges to the fixed map specified in Proposition 5.

Definition 16 (Asymptotic Average Consensus)**.** We say that the sequence of channels $\{\mathcal{E}_t(\cdot)\}_{t=0}^{\infty}$ asymptotically achieves *S-average σEC* for some $S \in$

$\mathfrak{H}(\mathcal{H}^m)$ if it asymptotically achieves σEC and for all ρ_0, it holds:

$$\begin{aligned} \lim_{t\to\infty} \mathrm{Tr}(\sigma\bar{\rho}_\ell(t)) &= \lim_{t\to\infty} \mathrm{Tr}(\sigma^{(\ell)}\rho(t)) = \lim_{t\to\infty} \mathrm{Tr}(S\rho(t)) \\ &= \mathrm{Tr}(S\rho_0) \end{aligned} \tag{7.9}$$

for all $\ell \in \{1,\dots,m\}$. The same definition holds for σSMC.
We say that the sequence of channels $\{\mathcal{E}_t(\cdot)\}_{t=0}^{\infty}$ asymptotically achieves *S-average RSC (resp. SSC)* if it asymptotically achieves RSC (resp. SSC) and for $S \in \mathfrak{H}(\mathcal{H}^m)$ *there exists* a $\sigma \in \mathfrak{H}(\mathcal{H})$ such that (7.9) holds for all ρ_0.

By expressing the action of quantum channels in the dual (Heisenberg) picture, it is possible to obtain a clear characterization of the dynamics that satisfy (7.9).

Proposition 11. Consider a sequence of CPTP channels $\{\mathcal{E}_t(\cdot)\}_{t=0}^{\infty}$, and call $\hat{\mathcal{E}}_t = \mathcal{E}_t \circ \mathcal{E}_{t-1} \circ \ldots \circ \mathcal{E}_1$. The associated dynamics satisfies (7.9) if and only if

$$S = \lim_{t\to\infty} \hat{\mathcal{E}}_t^\dagger(S) \quad \text{and} \quad \lim_{t\to\infty} \hat{\mathcal{E}}_t^\dagger(\sigma^{(\ell)}) = S \tag{7.10}$$

for $\ell = 1, 2, \dots, m$, where $\hat{\mathcal{E}}_t^\dagger = \mathcal{E}_1^\dagger \circ \mathcal{E}_2^\dagger \circ \cdots \circ \mathcal{E}_t^\dagger$.

Proof. The conditions (7.10) clearly imply (7.9). On the other hand, if (7.9) holds for all ρ_0, it is easy to obtain (7.10) by duality, taking the limit inside the trace functional. □

The first of the equalities in (7.10) holds in particular for the natural situation where $\mathcal{E}_t^\dagger(S) = S$ for all t. Similarly to the classical case, average quantum consensus algorithms could be a useful tool towards locally estimating collective quantities of an ensemble of many subsystems. Typically in large-ensemble quantum experiments, only few subsystems might be accessible by a measurement apparatus, and then applying a robust consensus procedure to the final state of the system could allow local measurements to provide a kind of "average state" knowledge of the whole ensemble — including potentially quantum correlations that survive throughout the network, e.g. satisfying pairwise Bell-inequalities, if several subsystems can be conditionally measured. Section 4 discusses this in more detail.

7.2 A Gossip Algorithm for Quantum Consensus

We now propose actual interactions that drive the quantum network to average consensus. As a building block, we focus on the interaction between two

subsystems while the others remain unchanged; all neighborhood-activation options build on this elementary case, as explained above.

7.2.1 Quantum Gossip Interactions

Let us introduce a way to implement gossip-type interactions in a fully quantum way. In a controlled quantum network, one can typically engineer unitary transformations that implement the "identity" evolution and the swapping of two neighboring subsystem states. Let us denote the permutation that swaps subsystems j and k as $\pi_{j,k}$ and its corresponding unitary operator by $U_{(j,k)}$. To develop our analysis, it will be convenient to introduce the *graph* G associated to the multipartite system: its nodes $1,\dots,m$ correspond to the "physical" subsystems, the edge (j,k) is included if the subsystems j and k have a non-zero probability to interact.

Assume edge (j,k) is selected at a certain step t. We then consider an auxiliary two-level system $\mathcal{Q}$ and the joint unitary evolution $I\otimes|\xi_I\rangle\langle\xi_I|+U_{(j,k)}\otimes|\xi_S\rangle\langle\xi_S|$ of the quantum network and the auxiliary system. This conditionally associates the two operations $I, U_{(j,k)}$ on the network to the orthogonal states $|\xi_I\rangle$ and $|\xi_S\rangle$ of $\mathcal{Q}$. Denoting by ρ the initial state of the quantum network and by $\rho_\xi = (1-\alpha)\,|\xi_I\rangle\langle\xi_I| + \alpha|\xi_S\rangle\langle\xi_S| + \beta|\xi_S\rangle\langle\xi_I| + \beta^*|\xi_I\rangle\langle\xi_S|$ the generic initial state of $\mathcal{Q}$, the joint state after the evolution gets:

$$\begin{aligned}\rho\otimes\rho_\xi \quad\to\quad & (1\text{-}\alpha)\,\rho\otimes|\xi_I\rangle\langle\xi_I| + \alpha\,U_{(j,k)}\rho U_{(j,k)}^\dagger\otimes|\xi_S\rangle\langle\xi_S| \\ & +\beta U_{(j,k)}\rho\otimes|\xi_S\rangle\langle\xi_I| + \beta^*\rho U_{(j,k)}^\dagger\otimes|\xi_I\rangle\langle\xi_S|\,.\end{aligned}$$

Taking the partial trace over the auxiliary system, we obtain as evolution for the quantum network a quantum channel that represents our fundamental **quantum gossip interaction**:

$$\rho(t+1) = \mathcal{E}_{j,k}(\rho(t)) = (1-\alpha)\,\rho(t) + \alpha\,U_{(j,k)}\rho(t)U_{(j,k)}^\dagger\,, \tag{7.11}$$

with $\alpha\in(0,1)$. Note that the conditional swapping only involves purely local interactions among subsystems j and k of the quantum network, plus the auxiliary system $\mathcal{Q}$ associated to this pair. The state of $\mathcal{Q}$ after interaction is discarded, hence it does not need measurement equipment. Moreover, any choice of $\alpha\notin\{0,1\}$ is sufficient to introduce some degree of *dissipation* (non-unitary evolution) on $\rho(t)$, which is necessary for convergence [35]. In accurately controlled settings [34], one may assume to have an actual resettable ancillary system associated to each link, or one or more "moving" ancillary

systems that activate the desired links. Then resetting the ancilla to an initial state with $\alpha = 1/2$ would optimize quasi-local mixing.

In the language we have developed in the first part, the quantum gossip interaction can be viewed as a convex combination of actions of the permutation group $\mathfrak{P}_n$ on $\mathfrak{D}(\mathcal{H})$. More precisely, suppose that at time t is selected the edge (j,k), we have that:

$$\rho(t+1) = \mathcal{E}_{j,k}(\rho(t)) = \sum_{\pi \in \mathfrak{P}_m} \mathsf{s}_\pi(t) a(U_\pi, \rho(t)) \tag{7.12}$$

with switching signal $\mathsf{s}(t) \in \mathbb{R}^{m!}$ given by:

$$\mathsf{s}_\pi(t) = \begin{cases} (1-\alpha) & \text{if } \pi = \mathrm{e}_{\mathfrak{P}_m}, \\ \alpha & \text{if } \pi = \pi_{j,k}, \\ 0 & \text{otherwise.} \end{cases} \tag{7.13}$$

The propagator accordingly is given by:

$$\hat{\mathcal{E}}_t(\rho_0) = \sum_{\pi \in \mathfrak{P}_m} \mathsf{p}_\pi(t) a(U_\pi, \rho_0), \tag{7.14}$$

for a suitable vector of convex weights $\mathsf{p}(t) \in \mathbb{R}^{m!}$.

7.2.2 Convergence to Consensus

We study convergence under three types of gossip dynamics: cyclic interactions, expectation of random interactions, and trajectory-wise for the random interactions. In all these cases, quantum gossip can be described by unital CPTP maps. We begin by recalling a characterization of the fixed points of such maps (see e.g. [30]).

Proposition 12. Let $\{V_i\}_{i=1}^K$ the Kraus decomposition of a *unital* CP map $\mathcal{E}(\cdot)$ and define:

$$\mathcal{A}_\mathcal{E} = \{X \in \mathfrak{B}(\mathcal{H}^m) \,|\, XV_i - V_i X = 0 \ \forall\, i = 1, \ldots, K\}\,. \tag{7.15}$$

Then $\bar{X} \in \mathfrak{B}(\mathcal{H}^m)$ is a fixed point of $\mathcal{E}$, i.e. $\mathcal{E}(\bar{X}) = \bar{X}$, if and only if $\bar{X} \in \mathcal{A}_\mathcal{E}$.

This helps determine the set of fixed points for the CP maps of interest in quantum gossip.

Lemma 8. Let $U_{(j,k)}$ denote the pairwise swap operation of subsystems (j,k) on $\mathcal{H}^m$. If the graph G associated to the system is connected, then the set of

fixed points of any CP unital map of the form

$$\begin{aligned} \mathcal{E}(X) &= q_0\, X + \sum_{(j,k)\in E} q_{j,k}\, U_{(j,k)}^\dagger X U_{(j,k)} \quad , \\ \text{with } q_0 + \sum_{(j,k)\in E} q_{j,k} &= 1\,, \quad q_0,\, \{q_{j,k}\} > 0 \end{aligned} \tag{7.16}$$

coincides with the set of permutation-invariant operators.

Proof. According to Proposition 12 above, the fixed points are the X satisfying $XU_{(j,k)} = U_{(j,k)}X$, or equivalently $U_{(j,k)}^\dagger X U_{(j,k)} = X$. The latter expresses that X is invariant with respect to pairwise swaps on all the graph edges. It is well known that sequences of pairwise swaps on the edges of a connected graph generate the full set of permutations on the set of nodes, and so we get the conclusion. □

The last result ensure that the equilibrium of our quasi-local dynamics are indeed the fixed points of the action defined in (6.5).

The following lemma shows how the contribution of the identity, i.e. the trivial permutation, in the CP map plays a crucial role in the proof of convergence.

Lemma 9. Consider a linear completely positive map $\mathcal{E}$ on $\mathfrak{B}(\mathcal{H})$ that admits an operator-sum representation $\{A_k\}$ with one operator proportional to identity, i.e. $A_1 = \sqrt{\alpha} I > 0$. Then, if λ is an eigenvalue of $\mathcal{E}$, $|\lambda| = 1$ implies $\lambda = 1$.

Proof. If $\mathcal{E}$ is a CPTP map it is a contraction in trace norm [26, 47], so its eigenvalues λ_k belong to the closed unit disk. By virtue of the Kraus-Stinespring representation theorem (see e.g. [48]), also $\mathcal{F} = \frac{1}{1-\alpha}(\mathcal{E} - \alpha I)$ is CPTP and thus has eigenvalues μ_k in the closed unit disk. Therefore the eigenvalues $\lambda_k = (1-\alpha)\mu_k + \alpha$ of $\mathcal{E} = (1-\alpha)\mathcal{F} + \alpha I$ in fact belong to the circle of radius $(1-\alpha)$ centered at α, which is strictly inside the unit circle except for a tangency point at $1 \in \mathbb{C}$, see Fig. 7.1. □

In other words, Lemma 9 excludes eigenvalues of unit norm different from $+1$, those which would cause limit cycles.

By combining the above properties, we get the following convergence result for quantum gossip. It shows that S-average SSC can be attained for global operators that are the permutation-invariant average of local ones; this is similar to classical gossip, where distributed computation of the average of individual

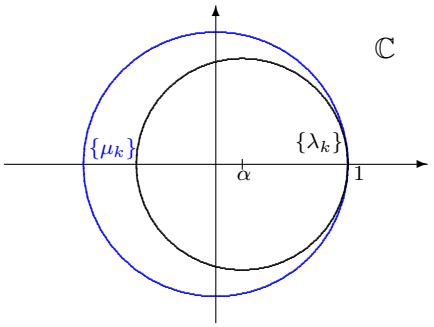

Figure 7.1: Delimitation of the (closed) domains for the eigenvalues $\{\mu_k\}$ of $\mathcal{F}$ (blue) and $\{\lambda_k\}$ of $\mathcal{E}$ (black) [color online].

states actually gives access to the value of any linear permutation-invariant function of these states.

Theorem 6. If the graph associated to the quantum systems is connected, then the quantum gossip algorithm (7.11) ensures global convergence towards SSC:

- deterministically, when the edges on which a gossip interaction occurs at a given time are selected by periodically cycling, in any predefined way, through the set of edges;

- in expectation, when the edges on which a gossip interaction occurs at a given time are selected randomly from a fixed probability distribution $\{q_{j,k} > 0 | \sum_{(j,k)\in E} q_{j,k} = 1\}$;

- with probability one on any trajectory, with the same edge-selection strategy of the previous point. Explicitly, there exists a state $\rho_* \in \mathcal{C}_{SSC}$ for which for any $\delta, \varepsilon > 0$, there exists a time $T > 0$ such that

$$\mathbb{P}[\,\mathrm{Tr}((\rho(T) - \rho_*)^2) \,>\varepsilon\,] \,<\delta\,.$$

In any of the above cases, the system converges to:

$$\rho_* = \frac{1}{m!} \sum_{\pi\in\mathfrak{P}} U_\pi \rho_0 U_\pi^\dagger\,. \tag{7.17}$$

Furthermore, S-average SSC is attained if and only if $S \in \mathfrak{H}(\mathcal{H}^{\otimes m})$ can be written, for some $\sigma \in \mathfrak{H}(\mathcal{H})$, in the form:

$$S = \frac{1}{m} \sum_{i=1}^{m} \sigma^{(i)}. \tag{7.18}$$

Proof. First notice that all the operators in the OSR of the map (7.16) are self-adjoint. This implies that permutation-invariant observables S are fixed points

for the associated dual map, and hence for the gossip interaction associated to any edge (j,k) and $\forall\,\rho$:

$$U_{(j,k)}SU_{(j,k)}^{\dagger}=S \;\Rightarrow\; \mathrm{Tr}[\mathcal{E}_{j,k}(\rho)S]=\mathrm{Tr}[\rho\mathcal{E}_{j,k}^{\dagger}(S)]=\mathrm{Tr}[\rho S]. \tag{7.19}$$

For the *cyclic evolution* map $\mathcal{E}_C$, we notice that all the simple two-subsystem swaps are still present with a weight different from zero in the OSR of the cyclic map (7.7), thanks to the presence of the identity in the OSR of each gossip interaction step. Therefore by Lemma 8 the fixed points are the permutation-invariant operators. Now consider the dynamics associated to $\mathcal{E}_C$ as a linear, time-invariant map acting on the subspace of hermitian matrices. From Lemma 9 and the fact that the time-invariant linear map leaves $\mathfrak{D}(\mathcal{H}^m)$ invariant (excluding unstable Jordan blocks), we have that all the *modes* of the LTI system are asymptotically stable except those corresponding to the fixed-point set, namely the permutation-invariant set: every initial state converges to a fixed point ρ_∞ in this set. Thus the SSC set is globally asymptotically stable, and in fact exponentially stable since the map is linear. Let us now prove that ρ_∞ has the form (7.17). For all *permutation invariant* X, from (7.19) we have that:

$$\mathrm{Tr}[X\mathcal{E}_C(\rho_0)] = \mathrm{Tr}[X\rho_0] \quad \forall\; t. \tag{7.20}$$

Combining the latter with the fact that ρ_∞ is permutation-invariant, that the set of all permutations is self-adjoint, and using (6.4), we get for arbitrary $Q\in\mathfrak{H}(\mathcal{H}^m)$:

$$\begin{aligned}
\mathrm{Tr}[Q\rho_\infty] &= \mathrm{Tr}[Q\frac{1}{m!}\sum_{\pi\in\mathfrak{P}}U_\pi\rho_\infty U_\pi^\dagger] \\
&= \mathrm{Tr}[\frac{1}{m!}\sum_{\pi\in\mathfrak{P}}U_\pi Q U_\pi^\dagger\rho_\infty] \\
&= \mathrm{Tr}[\frac{1}{m!}\sum_{\pi\in\mathfrak{P}}U_\pi Q U_\pi^\dagger\rho_0] \\
&= \mathrm{Tr}[\frac{1}{m!}\sum_{\pi\in\mathfrak{P}}Q U_\pi\rho_0 U_\pi^\dagger].
\end{aligned}$$

This implies that indeed $\rho_\infty=\rho_*$ as stated.

For the *expectation of random evolution*, the CPTP map $\mathcal{E}$ is exactly of the form of Lemma 8 and the same reasoning can be repeated.

For the *random trajectory evolution*, we repeat a proof similar to that of Proposition 1. Since $\mathcal{E}$ for a single evolution step is linear, *self-adjoint* with respect to the Hilbert-Schmidt inner product, and thus with eigenvalues in the

closed unit disk, it is a contraction for the Frobenius norm distance $\mathrm{Tr}((\rho_A - \rho_B)^2)$ between any two states $\rho_A, \rho_B \in \mathfrak{D}(\mathcal{H}^m)$. Indeed, $\mathcal{E}$ has non-increasing orthonormal modes, so by writing any operator $X \in \mathfrak{H}(\mathcal{H}^m)$ in the modal basis we directly get $\mathrm{Tr}(\mathcal{E}(X)^\dagger \mathcal{E}(X)) \leq \mathrm{Tr}(X^\dagger X)$; taking $X = \rho_A - \rho_B$ yields the contraction [1]. Now taking in particular $\rho_A = \rho$ and $\rho_B = \rho_*$, we get that the Frobenius distance from ρ to ρ_* can never increase. Moreover, by transitivity of the permutation operators, $\frac{1}{m!}\sum_{\pi\in\mathfrak{P}} U_\pi \rho U_\pi^\dagger = \frac{1}{m!}\sum_{\pi\in\mathfrak{P}} U_\pi \rho_0 U_\pi^\dagger = \rho_*$ for any ρ along the trajectory of the gossip algorithm. Now given the convergence under cyclic evolution, there must exist some $\lambda < 1$ and integer $M > 0$ such that

$$\mathrm{Tr}((\mathcal{E}_C^M(\rho) - \rho_*)^2) \leq \lambda \mathrm{Tr}((\rho - \rho_*)^2)$$

for any ρ for which $\frac{1}{m!}\sum_{\pi\in\mathfrak{P}} U_\pi \rho U_\pi^\dagger = \rho_*$. The proof then concludes along the same lines as Proposition 1, namely the probability to obtain an edge sequence which includes successions of M cyclic evolutions a sufficiently large number of times to have ε-convergence, gets arbitrarily close to 1 if we wait long enough.

Finally let us prove that we attain S-average consensus if and only if S can be decomposed as in (7.18). We know from the first part of the proof that all permutation-invariant observables S are fixed points for the associated dual map $\mathcal{E}_t^\dagger$. Then according to Proposition 11 we have S-average consensus if and only if there exists a local observable σ such that:

$$\lim_{t\to\infty} \hat{\mathcal{E}}_t^\dagger(\sigma^{(\ell)}) = S \tag{7.21}$$

for $\ell = 1, 2, \ldots, m$. Because of (7.19) and (7.17), by duality we have that for every local operator $\sigma^{(\ell)}$:

$$\lim_{t\to\infty} \hat{\mathcal{E}}_t^\dagger(\sigma^{(\ell)}) = \frac{1}{m!}\sum_{\pi\in\mathfrak{P}} U_\pi^\dagger \sigma^{(\ell)} U_\pi = \frac{1}{m}\sum_{i=1}^{m} \sigma^{(i)}. \tag{7.22}$$

Note that in the last inequality there are only m different contributions, this because since every local observable $\sigma^{(l)}$ is composed by the tensor product of $m-1$ identical operator (i.e. the identity matrices) and this decrease the number of permutations that act non trivially to m. This is the form (7.18), concluding the proof. □

Note that this latter proof could also have been formulated in the framework of the lifted dynamics. In the assumptions we have required that the graph

[1] This is analogous to the non-increasing Euclidean norm $x^T x = \|x\|^2$ under a classical consensus iteration with an *undirected* graph, and the related contraction of $\|x_A - x_B\|^2$.

associated to the quantum system is connected, hence the corresponding set of transpositions is a generating set for the group $\mathfrak{P}_m$. Furthermore, all the selection mechanisms considered satisfy the *primitivity assumption* stated in assumption 1. This is sufficient to conclude that $\mathsf{p}(t)$ converges toward the uniform vector of convex weights in $\mathbb{R}^{m!}$. This in turn implies that:

$$\lim_{t\to\infty} \hat{\mathcal{E}}_t(\rho_0) = \frac{1}{m!} \sum_{\pi\in\mathfrak{P}_m} U_\pi \rho_0 U_\pi^\dagger \qquad \forall\, \rho_0 \in \mathfrak{D}(\mathcal{H}^m), \tag{7.23}$$

as proved in the previous theorem.

Remark: This shows that the mean value of a (global) observable $S = \frac{1}{m}\sum_{\ell=1}^m \sigma^{(\ell)}$, with arbitrary σ, can be asymptotically retrieved from the state of any single subsystem after having applied one of the quantum gossip algorithms.

On the other hand, unlike for classical consensus, there are permutation-invariant operators that do not attain S-average consensus, because they cannot be written in the form (7.18). This is the case among others if S is orthogonal to the linear span of all the local observables. For instance if $\tilde{S} = \sigma_z^{\otimes^m}$, given the orthogonal basis $\{\sigma_k\}_{k=0,x,y,z}$ for $\mathfrak{B}(\mathcal{H})$, we have:

$$\mathrm{Tr}[\tilde{S}\,\sigma_k^{(\ell)}] = 0 \quad \forall\, k \in \{0,x,y,z\} \text{ and } \forall\, l \in \{1,\ldots,m\}.$$

Therefore $\tilde{S}$ cannot be written in the form (7.18), hence although $\tilde{S}$ is conserved by the gossip algorithm, the latter cannot lead to $\tilde{S}$-average consensus in the sense of Definition 16.

7.2.3 Classical equivalent to observable consensus dynamics

We next show how the quantum gossip algorithm (7.11) in fact implements in a quantum fashion the classical gossip as we restrict to σEC. According to Definition 10, a quantum state ρ belongs to $\mathcal{C}_{\sigma\mathrm{EC}}$ if:

$$\mathrm{Tr}[\sigma^{(1)}\rho] = \ldots = \mathrm{Tr}[\sigma^{(m)}\rho]. \tag{7.24}$$

In view of this, it seems reasonable to attempt a convergence study of the algorithm (7.11) directly in terms of the evolution of the expectation values of the $\sigma^{(\ell)}$ operators. This is not possible for arbitrary quantum evolutions, since a quantum state is far from fully specified by a *single* set of commuting observable expectations, and different states with the same expectation may lead to very different evolutions. However, our quantum gossip algorithm remarkably allows us to write a model for the average dynamics of the $\sigma^{(\ell)}$ in closed form.

More precisely, let us define $z_\ell(t) := \mathrm{Tr}[\mathcal{E}^t(\rho_0)\sigma^{(\ell)}] = \mathrm{Tr}[\rho_t\sigma^{(\ell)}]$. Note that for one subsystem swap $U_{(j,k)}$, we have:

$$\mathrm{Tr}[\sigma^{(\ell)}U_{(j,k)}\,\rho\,U_{(j,k)}^\dagger] = \begin{cases} z_\ell & \text{if } \ell \notin \{j,k\} \\ z_k & \text{if } \ell = j \\ z_j & \text{if } \ell = k\,. \end{cases} \tag{7.25}$$

According to (7.25) and (7.11), the random gossip algorithm update yields, with probability $q_{j,k}$, i.e. when the edge (j,k) is selected:

$$\begin{aligned}(z_j(t{+}1), z_k(t{+}1)) &= (1\text{-}\alpha)(z_j(t), z_k(t)) + \alpha(z_k(t), z_j(t)) \\ z_\ell(t{+}1) &= z_\ell(t) \quad \text{for all } \ell \notin \{j,k\}\,.\end{aligned}$$

This last expression is exactly the classical gossip algorithm (2.4). Therefore, Proposition 1 readily implies:

Corollary 3. Under all the various edge selection strategies for quantum consensus algorithm (7.11), the $z_\ell(t)$, $\ell = 1, 2, ..., m$ asymptotically converge towards:

$$\lim_{t\to\infty} z_\ell(t) = \frac{1}{m}\sum_{k=1}^{m} z_k(0) \quad \text{for all } \ell \in \{1, 2, ..., m\}\,.\square$$

We remark that this only proves average σ-*Expectation Consensus* of the quantum gossip algorithm, while our previous Theorem 6 shows that the algorithm in fact ensures the stronger average *Symmetric State Consensus*.

7.2.4 Gossip algorithm example

In this section we briefly discuss the evolution induced by random quantum gossip interactions (7.11) on a four-qubit network whose associated graph is a path[2]. We observe its convergence toward average σEC, average RSC and average SSC. In particular we consider as a "target" global observable:

$$S = \frac{1}{4}\left(\sigma_z^{(1)} + \sigma_z^{(2)} + \sigma_z^{(3)} + \sigma_z^{(4)}\right)\,. \tag{7.26}$$

Let the initial state be:

$$\rho = |1,0,1,0\rangle\langle 1,0,1,0|, \tag{7.27}$$

which is pure, and does not satisfy any of the consensus definitions provided in Chapter 6.

[2]I.e. the available neighborhoods, labeling the subsystems as $\{1,2,3,4\}$, are $\{1,2\}$, $\{2,3\}$ and $\{3,4\}$.

By Theorem 6 the state asymptotically converges to:

$$\begin{aligned}\rho_\infty &= \lim_{t\longrightarrow\infty}\rho(t) = \frac{1}{4!}\sum_{\pi\in\mathfrak{P}_4} U_\pi\rho_0 U_\pi^\dagger \\ &= \frac{1}{6}\ \ (|1,1,0,0\rangle\langle 1,1,0,0| + |1,0,1,0\rangle\langle 1,0,1,0| \\ &\quad + |1,0,0,1\rangle\langle 1,0,0,1| + |0,1,1,0\rangle\langle 0,1,1,0| \\ &\quad + |0,1,0,1\rangle\langle 0,1,0,1| + |0,0,1,1\rangle\langle 0,0,1,1|).\end{aligned} \tag{7.28}$$

This expression is clearly invariant under all the subsystem permutations, i.e. ρ_∞ is in SSC, and therefore also in RSC and σEC for all σ. The expectation value of S is preserved at any step, and by Theorem 6 the algorithm drives the system to S-average consensus, with $\sigma = \sigma_z$.

However, ρ_∞ is not in σSMC for any $\sigma \neq \alpha I$. Indeed, according to Proposition 5, ρ_∞ is in σSMC if and only if $\mathrm{Tr}[\rho_\infty \Pi_{sym}] = 1$. Now let $\{\Pi_i\}_{i=1}^6$ denote the orthonormal rank-one projectors in (7.28) and define the orthonormal projector $\bar{\Pi} = \sum_{i=1}^6 \Pi_i$, such that $\rho_\infty = \frac{1}{6}\sum_{i=1}^6 \Pi_i = \frac{1}{6}\bar{\Pi}$. We then get

$$\mathrm{Tr}[\rho_\infty \Pi_{sym}] = \frac{1}{6}\mathrm{Tr}[\sum_{i=1}^{6}\Pi_i\Pi_{sym}] = \frac{1}{6}\mathrm{Tr}[\Pi_{sym}\bar{\Pi}]. \tag{7.29}$$

This last expression is equal to 1 if and only if $\mathrm{Tr}[\Pi_{sym}\bar{\Pi}] = 6$. However, excluding the trivial case $\sigma = \alpha I$, for qubit networks Π_{sym} is always a two dimensional projector, so $\mathrm{Tr}[\Pi_{sym}\bar{\Pi}] \leq 2$. Hence ρ_∞ cannot be in σSMC for any non-trivial σ.

Figure 7.2 shows the evolution of the expectation values of the local and of the global observables related to σ_z as the iterations proceed for one run. The edges are selected at random with uniform probability, and the mixing parameter α is taken to be 1/2. With this particular choice, the reduced density operators of two subsystems that have just interacted are equal; this explains why a maximum of three points are visible on the graph at any time. The plot shows that asymptotically the expectation of the local observables σ_z tend to the expectation value of the global observable S, while the expectation value of S is preserved at each step.

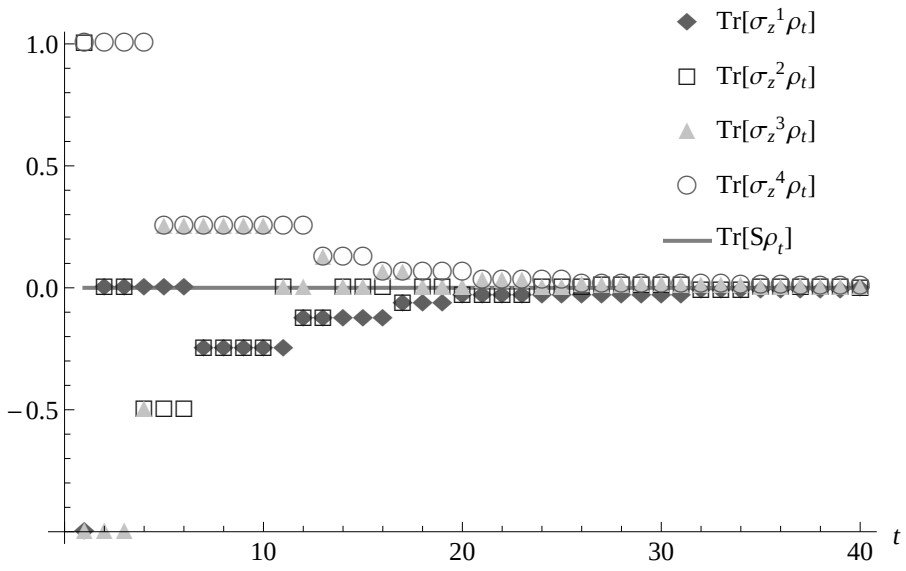

Figure 7.2: Evolution toward σ-Expectation Consensus for a four-qubit network arranged in a path graph.

8

Applications in the Quantum Domain

We now outline some possible applications of the ideas developed in the paper. We want to emphasize the wide flexibility and intrinsic robustness in engineering dynamics that leads to consensus. For example, in our algorithm the strength of mixing, the order of the interactions and the neighborhood topology can be allowed to vary, within the limits imposed by Theorem 6 and by the Primitivity Assumption 1. In this sense, consensus is a robust behavior, that does not have to be tightly controlled. It could e.g. naturally appear in a large lattice of sites where quantum particles can be found and, because of free or (purposefully) perturbed dynamics, particles are allowed to stochastically move around the lattice, hence effectively exchanging states between lattice nodes.

In the examples below, we shall assume that such a consensus-yielding process is present in a large network of interest, while accurate control and/or measurement is only possible on a limited number of subsystems — say, those on the boundary of the lattice, or temporarily removed from it to allow interactions with other pieces of laboratory equipments. With experimental quantum systems, this is typically the case when measurement processes are concerned. Thanks to our consensus results, we show how the mixing induced by the consensus dynamics can be exploited to achieve some network-wide tasks with such restricted local control access.

8.1 Gossip Symmetrizing Probability Distributions II

Let us start by noting that SSC corresponds to the quantum version of the symmetrization of join probability distributions exposed in Example 4.2. A classical random variable, in fact, can be viewed as a special, commutative case in the framework of non-commutative probability theory. Consider again a multipartite quantum system, composed of m isomorphic subsystems with total Hilbert space given by $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes ... \otimes \mathcal{H}_m$, a multipartite state $\rho \in \mathfrak{D}(\mathcal{H})$ play the role of a joint probability distribution. By using the action of the permutation group defined in (6.3), the quantum gossip interaction applied to an observable $X \in \mathfrak{B}(\mathcal{H})$ is given by:

$$X(t+1) = (1-\alpha)X(t) + \alpha U^{\dagger}_{(j,k)} X(t) U_{(j,k)}, \quad \alpha \in (0,1)\,.$$

Then both the cyclic and randomized versions of this quantum gossip algorithm will drive any initial X to:

$$\hat{X} = \frac{1}{m!} \sum_{\pi \in \mathfrak{P}} U^{\dagger}_{\pi} X U_{\pi}\ .$$

Physically, this implies that the measurement of any joint property on a subset of $n < m$ quantum systems will give the same statistics irrespective of the particular n subsystems that are selected, realizing thus the goal of Equation (4.1). Equivalently, we could gain the same conclusion considering the quantum gossip algorithm acting on ρ. Finally, Example 4.2 is retrieved when all considered operators are diagonal in a fixed basis, and the diagonal of the density operator is then equivalent to a classical probability density.

8.2 Estimation of a Global Variable from a Subsample

Consider a quantum system composed of a large number m of identical quantum subsystems, initially prepared by some experiment in an unknown global state ρ. We are interested in estimating the "average value of a physical property Q over all subsystems", that is the expected value $\bar{q}$ of the observable $\bar{Q} = \frac{1}{m}\sum_{j=1}^{m} Q^{(j)}$, for the state ρ. But we are allowed to *perform measurements only on a fixed subset of $p \geq 1$ subsystems* that are accessible to our measurement apparatus.

Let q_j be the random variable (RV) describing the outcome of a local measurement of $Q^{(j)}$ on ρ. Since $Q^{(j)}$ and $Q^{(k)}$ commute for $j \neq k$, we can perform a joint measurement of $Q^{(1)}$ to $Q^{(p)}$ on a single realization of ρ. A natural estimator for the expectation $\bar{q} = \mathrm{Tr}(\rho\bar{Q})$ would then be of course the sampled average:

$$\hat{q}(p) = \frac{1}{p}\sum_{j=1}^{p} q_j \,. \tag{8.1}$$

If we can repeat the same experiment, producing k times the same ρ and denoting $q_j(z)$ the respective measurement outcomes of RV q_j for $z = 1, 2, ..., k$, then a better estimate would be

$$\hat{q}(p,k) = \frac{1}{pk}\sum_{j=1}^{p}\sum_{z=1}^{k} q_j(z) \,. \tag{8.2}$$

By letting k grow large enough, we can make the variance of $\hat{q}$ arbitrarily small. However, in all practical situations where there might be local variations in the network, the p accessible subsystems are not bound to be representative of the whole ensemble, and $\hat{q}$ is unavoidably *biased*, unless $p = m$.

This problem is resolved if our gossip-type algorithm can be first enacted on the whole network of m subsystems. Indeed, the measurement statistics obtained from the p *fixed subsystems* after reaching the consensus state $\rho_* = \frac{1}{m!}\sum_{\pi\in\mathfrak{P}_m} U_\pi \rho U_\pi^\dagger$ are equal to the measurement statistics if we had access to p *randomly selected subsystems* before consensus[1]. In particular, the expected value of an estimate $\bar{q}(p)$ of Q from our p subsystems becomes:

$$\begin{aligned}
\mathbb{E}[\bar{q}(p)] &= \frac{1}{p\,m!}\mathrm{Tr}\left[\textstyle\sum_{j=1}^{p}\sum_{\pi\in\mathfrak{P}_m} Q^{(j)}\, U_\pi \rho U_\pi^\dagger\right] \\
&= \frac{1}{p\,m!}\mathrm{Tr}\left[\textstyle\sum_{j=1}^{p}\sum_{\pi\in\mathfrak{P}_m} Q^{(\pi(j))}\rho\right] \\
&= \frac{1}{m}\mathrm{Tr}\left[\textstyle\sum_{j=1}^{m} Q^{(j)}\rho\right] \;=\; \mathrm{Tr}[\bar{Q}\,\rho] \,.
\end{aligned}$$

Thus $\bar{q}(p)$ provides an unbiased estimator for $\bar{q}$, irrespective of the value of p.

Further computations along the lines of statistical sampling without replacement then allow to analyze the variance of $\bar{q}(p)$.

For k independent measurements, the result can be extended as in (8.2). Note that RSC would be sufficient to guarantee that the estimation is unbiased for *any* local Q. In addition, SSC would allow to generalize the present setting to situations where the local operator Q is replaced by a collective operator on

[1] This follows from the standard statistical mixture interpretation of a convex combination of density operators.

less than p particles. We could then probe different types of *average, symmetric* correlations in the experimental state ρ.

8.3 Purifying and cooling of a sample by local feedback actions

It is known that unitary control and projective measurements are enough for the preparation of any pure state for a *single system* [51]. We here show how local access to an arbitrarily small subset of a quantum network, *in conjunction with our gossip interactions,* is enough to asymptotically prepare a class of factorized, pure states on the whole network.

Consider again a set of m identical subsystems, of which only the first p are accessible via measurements of identical, non-degenerate $Q = \sum_{x=1}^{d} \lambda_x \Pi_x$. Further assume that, after each measurement, we can apply any desired unitary control on these p subsystems, possibly dependent on the measurement result [50, 51], while some (weaker, unsupervised) control procedure allows us to reach global SSC on the m subsystems. For simplicity we will make explicit reference to our gossip interactions.

We are interested in preparing the whole quantum network in a pure state. It can be shown that *any pure, factorized state of the form:*

$$\hat{\rho} = |\psi\rangle\langle\psi| \otimes \cdots \otimes |\psi\rangle\langle\psi|,$$

can be asymptotically obtained with the control resources described above, by cyclically iterating the following two steps:

Step I Apply a finite number $M > 0$ of gossip interactions.

Step II Perform measurements of Q on each of the p probe subsystems. This brings the network into a state

$$\Pi_{x_1} \otimes \Pi_{x_2} ... \otimes \Pi_{x_p} \otimes \rho_{\setminus p} ,$$

where x_1 to x_p are the measurement results and $\rho_{\setminus p}$ is an unknown state on $m - p$ subsystems. Then for each $k = 1, 2, ..., p$, use a unitary control action U_k on subsystem k such that $U_k \Pi_{x_k} U_k^\dagger = |\psi\rangle\langle\psi|$.

At each iteration of Step II the expectation $V(\rho) = 1 - \mathrm{Tr}[\hat{\rho}\,\rho]$ is either left unchanged, if the p subsystems are already all prepared in $|\psi\rangle\langle\psi|$, or else it

must decrease. During Step I, $V(\rho)$ is not changed, since $\hat{\rho}$ is a permutation-invariant operator. If the conditions of Theorem 6 hold, then it is easy to show that the largest invariant set for the whole procedure is contained in the kernel of $I - \hat{\rho}$. Hence, by LaSalle invariance theorem we conclude that $\hat{\rho}$ is prepared asymptotically.

If the global Hamiltonian of the network is of the form $H_{\text{tot}} = \sum_{j=1}^{m} Q^{(j)}$ or, more generally, admits a ground state of the form $\hat{\rho}$, then this procedure can be used to obtain ground-state cooling.

By variations of the above protocol, the same control capabilities can be used to engineer dynamics that asymptotically drive the state of the quantum network to have support on an arbitrary target subspace of the network's joint Hilbert space, provided it is invariant with respect to subsystem permutations.

8.4 Estimating the size of a sample

Consider again a set of m identical subsystems, with the same control capabilities as in the previous application: only the first p are accessible via measurements of identical, non-degenerate observables Q and feedback unitary control, while SSC can be reached on the whole network. We are now interested in estimating the number m of subsystems in the quantum network.

For this, we will first prepare the network in a state ρ' that has support in *a subspace which is orthogonal to some "marker" eigenstate $|\psi\rangle$ of Q,* such that $\text{Tr}[\rho' \, (|\psi\rangle\langle\psi|)^{(j)}] = 0$ for all j. Such a state can be asymptotically reached with an easy adaptation of the protocol described in the last section.

In order to estimate the size of the sample, we next apply the following procedure.

Step 1 Perform measurements of Q on each of the p probe subsystems, and use fast unitary control on each of them *in order to prepare them all in the marker eigenstate $|\psi\rangle$ of Q*;

Step 2 Let the network evolve with gossip to SSC;

Step 3 Perform again measurements of Q on the p probe subsystems, recording how many times $|\psi\rangle$ is obtained.

Step 1 prepares the network into a state

$$|\psi\rangle\langle\psi| \otimes ... \otimes |\psi\rangle\langle\psi| \otimes \rho_{\setminus p} \, ,$$

where $\rho_{\setminus p}$ is an unknown state on $m-p$ subsystems (with m unknown), but still satisfying $\mathrm{Tr}[\rho_{\setminus p}\,(|\psi\rangle\langle\psi|)^{(j)}] = 0$ for all $j > p$. As shown in Section 8.2, the statistics of measuring Q on the p probe subsystems after Step 2, equals the statistics of measuring Q before Step 2 on p uniformly randomly selected subsystems. In the latter case, whenever one of the first p subsystems was selected we would get outcome $|\psi\rangle$, while whenever a subsystem $j > p$ is selected we would certainly not get $|\psi\rangle$. The random variable K counting the number k of times $|\psi\rangle$ is detected in Step 3 therefore follows a hypergeometric distribution,

$$K = k \text{ with probability } \tbinom{p}{k}\tbinom{m-p}{p-k} \,/\, \tbinom{m}{p}$$

where $\binom{b}{a} = b!/(a!(b-a)!)$. We thus have:

$$\mathbb{E}[K] = p^2/m,$$

$$\mathrm{Var}(K) = \mathbb{E}[(K - \mathbb{E}(K))^2] = \frac{p^2(m-p)^2}{m^2(m-1)}.$$

Then, the estimator can be chosen to be $\hat{m} = p^2/\hat{K}$, where $\hat{K}$ is the sampled value of K. It is then easier to study the statistical properties of $\hat{m}^{-1}$, being just a rescaling of the measured $\hat{K}$. The *relative error* $\frac{\hat{m}^{-1}-m^{-1}}{m^{-1}}$ of $\hat{m}^{-1}$ then has mean zero, i.e. it is an unbiased estimator. We can then compute its variance:

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{\hat{m}^{-1}-m^{-1}}{m^{-1}}\right)^2\right] &= \mathbb{E}\left[m^2\left(\frac{\hat{K}}{p^2}-\frac{1}{m}\right)^2\right] \\
&= \frac{m^2}{p^4}\mathrm{Var}(K) \\
&= \frac{m^2}{p^4}\cdot\frac{p^2(m-p)^2}{m^2(m-1)} \\
&= \frac{(m-p)^2}{p^2(m-1)}.
\end{aligned}$$

This shows that if we pick $p = \alpha \cdot m$ to be a *fixed yet unknown* fraction of the total population, when the population increases the relative accuracy of $\hat{m}^{-1}$ improves since the above variance goes to zero as $1/m$. Then for the limit of large m, we can conclude that the variance of $\hat{m}$ also goes to zero as $1/m$.

8.5 Dynamical decoupling

Quantum Dynamical Decoupling (DD) is a set of open-loop control techniques

that are primarily used to reduce the effect of unknown Hamiltonian drifts, or couplings to the environment, on a target quantum system [52]. The main idea is to apply a sequence of unitary rotations to the system, such that effects of the undesired dynamics before and after a unitary rotation compensate each other. This task can be translated into a symmetrization task [53], and we show here how our results suggest a robust DD scheme. For the sake of simplicity, we restrict ourselves to the suppression of the drift Hamiltonian in finite dimensional systems. The extension to decoupling from the environment is straightforward.

When the Hamiltonian of a quantum system is time-dependent, the corresponding unitary transformation must be computed as an ordered product of exponentials over infinitesimal intervals, see Chapter 5. For a fixed time T, the resulting unitary operator can be associated to an effective Hamiltonian H_{eff} such that

$$U_T = e^{-iH_{\text{eff}}T} .$$

A DD strategy consists in a time-dependent control Hamitlonian $H_c(t)$ such that, for any constant H_d in a class of expected perturbations, the effective Hamiltonian associated to $H_d + H_c(t)$ is "close" to a scalar matrix after a predefined time T: $H_{\text{eff}} \approx \lambda I$ with $\lambda \in \mathbb{R}$. Indeed, this would suppress any physical effect of H_d at time T since global phases of the form $U_t = e^{i\lambda t}$ are irrelevant for predictions in quantum mechanics [44]. DD in its simplest form entails a sequence of fast, impulsive control operations that induce a group of "instantaneous" unitary transformations on the system, and achieves first-order suppression of H_d. The relevant time interval $[0, T)$ is subdivided into N subintervals of length $dt = T/N$ and instantaneous controls are applied at the end of each sub-interval so that the effective Hamiltonian, in the interaction picture [39], for subinterval $[(k-1)\,dt,\ k\,dt)$ is $g_k H_d g_k^\dagger$ with $g_k \in \mathcal{G}$. Then, the Magnus expansion [46] allows to approximate the exact evolution from time 0 to T to first order as:

$$\begin{aligned} e^{-i\,dt\,g_1 H_d g_1^\dagger}\, e^{-i\,dt\,g_2 H_d g_2^\dagger} \ldots e^{-i\,dt\,g_N H_d g_N^\dagger} \;&\approx \\ \approx\; e^{-i\,dt\,\sum_{k=1}^{N} g_k H_d g_k^\dagger} &=: e^{-i\,T\,\bar{H}} . \end{aligned} \tag{8.3}$$

Accuracy improves as the product of H_d with dt gets smaller. Hence, given a class $\mathfrak{H}_0$ of drift Hamiltonians on some finite-dimensional Hilbert space $\mathcal{H} \cong \mathbb{C}^n$, first-order DD follows from identifying a finite subgroup $\mathcal{G}$ of unitaries such

that

$$\tfrac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} g\,H_d\,g^\dagger = \lambda I \tag{8.4}$$

for all $H_d \in \mathfrak{H}_0$. In the language of the symmetrization framework, DD achieves symmetrization with respect to a group $\mathcal{G}$, and the latter is selected such that the action $a(g,H) = g\,H\,g^\dagger$ on the space $\mathcal{X}$ of all Hamiltonians H satisfies $\bar{\mathcal{F}}(\mathfrak{H}_0) \subseteq \{\lambda I, \lambda\in\mathbb{R}\}$.

Achieving symmetrization in (8.3) means choosing each $g \in \mathcal{G}$ an equal number of times over the N subintervals. An obvious choice is just to take $N = m|\mathcal{G}|$ and iterate m times a predefined path through the elements of $\mathcal{G}$. However, when H_d is not really constant for a duration $|\mathcal{G}|\,dt$ or when considering higher-order Magnus terms, the potential advantage of randomized [31, 54] or concatenated [55] sequences of g_k has been recognized. Our general dynamics (3.15) allows to retrieve and combine these two variants of DD and, in particular, to highlight their robustness.

Consider an iterative construction of the sequence of unitaries g_k, where at the n-th iteration the time interval $[0,T)$ is subdivided into $N = 2^n$ subintervals. Denote $\mathcal{S} \subseteq \mathcal{G}$ the set of available control actions. We start at $n = 0$ from the situation with no control pulses, so $g_1 = \mathrm{e} \cong I_{\mathcal{H}}$ over $[0,T)$ and $\bar{H} = H_d$. Increasing n, we then choose one element $h(n) \in \mathcal{S}$, we divide each subinterval $\big[(m-1)\frac{T}{2^n},\, m\frac{T}{2^n}\big)$ into two equal time intervals $\big[(2m-2)\frac{T}{2^{n+1}},\, (2m-1)\frac{T}{2^{n+1}}\big)$ and $\big[(2m-1)\frac{T}{2^{n+1}},\, 2m\frac{T}{2^{n+1}}\big)$, and we update the sequence as follows for $m = 1, ..., 2^n$:

$$\begin{aligned} &\text{At} \quad n: \;\; g_m = \bar{g} \\ &\text{At} \quad n+1: \;\; g_{2m-1} = \bar{g}\,, \;\; g_{2m} = h(n)\bar{g}\,. \end{aligned} \tag{8.5}$$

Denoting by $\mathsf{p}_g(n)$ the fraction of time $[0,T)$ during which $g_k = g \in \mathcal{G}$, the procedure (8.5) correponds to (3.15) with t replaced by n, and the switching signal:

$$\begin{aligned} \mathsf{s}_g(n) &= 1/2 \;\text{ for } g \in \{\mathrm{e}_{\mathcal{G}},\, h(n)\}\,, \\ \mathsf{s}_g(n) &= 0 \;\text{ for all other } g \in \mathcal{G}\,. \end{aligned} \tag{8.6}$$

In action form, the average Hamiltonian at the n-th iteration is

$$\bar{H}_n = \textstyle\sum_{g\in\mathcal{G}} \mathsf{p}_g(n)\, a(g, H_d) \;=\; \sum_{g\in\mathcal{G}} \mathsf{s}_g(n-1)\, a(g, \bar{H}_{n-1})\,.$$

Our theorems ensure the convergence of $\bar{H}_n$ towards the $\mathcal{G}$-symmetrized form (8.4) of H_d as n is increased, if Assumption 1 holds. This is valid both for

deterministic or random choices of the $h(n)$. Furthermore, our results indicate a remarkable generality and robustness of the procedure: (i) the control actions $h(n)$ don't have to be chosen uniformly in $\mathcal{G}$, actually any deterministic choice or probabilistic distribution over enough elements will work; (ii) the set $\mathcal{S}$ of control actions does not have to be all $\mathcal{G}$, e.g. a set of generators would be sufficient; and (iii) the subdivision can be more general than a "perfect average": any $\mathsf{s}_{h(n)}(n) = 1 - \mathsf{s}_{\mathrm{e}}(n) = \alpha$ with $\alpha \in (0, 1)$ would asymptotically work, not just (8.6) where $\alpha = 1/2$.

8.6 Speed-up for Random State Generation

In this section we present a proof of principle on how to generate random quantum state over a given set driven by a quantum random walk that achieve an almost quadratic speed up with respect to its classical counterpart presented in Section 4.4. Notably, we will show how this scheme can be applied to a set of non-orthogonal quantum states. This feature is particular desirable for quantum key distribution application. In fact, the security of many protocol, such for example the BB84, relies on encoding a pice of information on a set non orthogonal quantum states [56].

The relations between a finite group and a subset of its elements can be explored by considering its Cayley graph.

Definition 17 (Cayley graph)**.** Let $\mathcal{G}$ be a finite group and let $\mathcal{S}$ be a set of generators. Let us also consider $E_{\mathcal{S}} \subseteq \mathcal{G} \times \mathcal{G}$:

$$E_{\mathcal{S}} = \{(g,h) \in \mathcal{G} \times \mathcal{G} \text{ if } \exists\, s \in \mathcal{S}\; sh = g\}. \tag{8.7}$$

The Cayley graph of $\mathcal{G}$ generated by $\mathcal{S}$ is the graph defined as:

$$\Gamma := \Gamma(\mathcal{G}, E_{\mathcal{S}}) \tag{8.8}$$

The graph is $|\mathcal{S}|$-regular and if $\mathcal{S} = \mathcal{S}^{-1}$ it is also undirected, Furthermore it hold the following result:

Proposition 13. The Cayley graph of the finite group $\mathcal{G}$ generated by $\mathcal{S} \subset \mathcal{G}$ is connected if and only if $\mathcal{S}$ is a generator set of $\mathcal{G}$.

In Section 3.1.1 we have pointed out how the lifted dynamics can be seen as a random walk on a Cayley graph with node given by the group element a edges related to the group elements available at each step. More precisely, let W be a random variable on $\mathcal{G}$, such random walk is defined by the transition probability:

$$\mathsf{p}(W(t+1) = g | W(t) = h) = s_s(t) \;\text{ with } s \in \mathcal{S} \text{ s.t. } sh = g \tag{8.9}$$

We proceed by quantizing the Cayley graph and the dynamics that takes place on it. A straightforward way is to associate to each group element an element of an orthonormal basis of a $|\mathcal{G}|$ Hilbert space $\mathcal{H}_{\mathcal{G}}$, so we will denote $|g\rangle$ the base element associated to g. Namely we have that:

$$\mathcal{H}_{\mathcal{G}} := \text{span}\{|g\rangle,\, g \in \mathcal{G}\} \tag{8.10}$$

Let us now consider the dynamics for the discrete time case.

Discrete time case: Following the intuition from classical random walks let us first add an additional degree of freedom, the so-called "coin" that describes the possible directions of the walker, i.e. the different group elements available for each step, see for example [57] and [58]. Such additional degree of freedom will act on Hilbert space $\mathcal{H}_{\mathcal{S}}$ defined as:

$$\mathcal{H}_{\mathcal{S}} := \text{span}\{|s\rangle,\ s \in \mathcal{S}\}. \tag{8.11}$$

The total unitary evolution will take place in the Hilbert space $\mathcal{H} := \mathcal{H}_{\mathcal{G}} \otimes \mathcal{H}_{\mathcal{S}}$.

Each step of the discrete time unitary evolution will be thus the composition of two unitary operations. First a coin operation acting non-trivially only on $\mathcal{H}_{\mathcal{S}}$ and a shift operation that updates the state of the walker conditionally to the state of the coin register. More precisely, let us define the unitary operator over $\mathcal{H}$:

$$C_{(\mathcal{G},\mathcal{S})} := \mathbb{I}_{\mathcal{G}} \otimes U^{\mathcal{S}}, \tag{8.12}$$

where $U^{\mathcal{S}}$ is a unitary operator over $\mathcal{H}_{\mathcal{S}}$ that take the name of the *coin operator*. In the literature different choices for $U^{\mathcal{S}}$ are explored leading to different behavior and performances of the quantum walker [59] and [58].

Next, the *shift operator* defined as:

$$T_{(\mathcal{G},\mathcal{S})} := \sum_{s\in\mathcal{S},g\in\mathcal{G}} |sg\rangle\langle g| \otimes |s\rangle\langle s|. \tag{8.13}$$

Note that the previous expression corresponds to a conditional unitary operation on the space $\mathcal{H}_{\mathcal{G}}$. This can be seen by considering the unitary operator:

$$U_s^{\mathcal{G}} := \sum_{g\in\mathcal{G}} |sg\rangle\langle g|, \tag{8.14}$$

hence we have that the shift operator can be written as:

$$T_{(\mathcal{G},\mathcal{S})} := \sum_{s\in\mathcal{S},g\in\mathcal{G}} U_s^{\mathcal{G}} \otimes |s\rangle\langle s|. \tag{8.15}$$

One full step of the walker is thus given by:

$$U_{(\mathcal{G},\mathcal{S})} := T_{(\mathcal{G},\mathcal{S})}(\mathbb{I}_{\mathcal{G}} \otimes U_G^{\mathcal{S}}) \tag{8.16}$$

A unitary dynamics does not allows for convergence, in fact, it send orthogonal vectors in orthogonal vectors. Nonetheless, mixing can be induced

by performing suitable measurement on the systems at random times. More precisely, let suppose the initial state of the system is given by $|\psi_0\rangle := |g\rangle \otimes |l\rangle$. Let us denote for sake of simplicity $U := U_{(\mathcal{G},\mathcal{S})}$ and write for the evolution up to time t:

$$|\psi_t\rangle\langle\psi_t| = U_t|\psi_0\rangle\langle\psi_0|U_t^\dagger \tag{8.17}$$

We can compute the probability for the walker to be found in the node corresponding to the group element g after t steps given the initial state $|\psi_0\rangle$

$$\mathsf{p}_t(g|\psi_0) := \mathrm{Tr}[|g\rangle\langle g|\mathrm{Tr}_{\mathcal{S}}[|\psi_t\rangle\langle\psi_t|]] \tag{8.18}$$

where with $\mathrm{Tr}_{\mathcal{S}}[\,\cdot\,]$ we denote the partial trace over the space $\mathcal{S}$. Note that we are implicitly assuming that there exists some observable $\hat{N}$ such that $\hat{N} = \sum_g \lambda_g |g\rangle\langle g|$ with $\lambda_g \neq \lambda_h$ for any $g, h \in \mathcal{G}$.

If we set a fixed time T and perform the latter measurement at a time t extracted uniformly at random from $[0, T-1]$ the probability of being on the node g is going to be distributed according to:

$$\bar{\mathsf{p}}_T(g|\psi_0) := \frac{1}{T}\sum_{t=0}^{T-1} \mathsf{p}_t(g|\psi_0) \tag{8.19}$$

We collect all the probabilities into a single probability vector $\hat{\mathsf{p}}_T(|\psi_0\rangle) = (\bar{\mathsf{p}}_T(g_1|\psi_0), \ldots, \bar{\mathsf{p}}_T(g_{|\mathcal{G}|}|\psi_0))$.

Definition 18 (Quantum Walk)**.** We call *coined quantum walk* the unitary evolution defined in (8.16) interrupted uniformly at random over the interval $[0, T-1]$ by the measurement defined in (8.18).

In [58] it has been proved that the latter probability distribution asymptotixally converges to some *limiting distribution* on the element of $\mathcal{G}$ that in general:

- depends on the initial state $|\psi_0\rangle$,
- it is not the uniform distribution over $\mathcal{G}$.

Whenever the limiting distribution is uniform or not depending on the particle dynamics and of the group at hand is still matter of deep investigation see for example [60]. Anyway, in [58] it has been proved the following result.

Proposition 14. The coin quantum walk on the Cayley graph of an abelian group does not depend on the initial state if the the eigenvalues of U are all distinct.

In [58] it has been also proved that if the graphs is an n-cycle with n odd and the coin is the Hadamard operator the walker asymptotically converges to the uniform distribution over the nodes. More precisely, let us denote with u the uniform distribution over the nodes i.e. $\mathsf{u} = 1/n(1, \ldots, 1)$, we have that:

$$\lim_{T\to\infty} \hat{\mathsf{p}}_T(|\psi_0\rangle) = \mathsf{u}. \tag{8.20}$$

Finally they prove a bound for the speed of convergence of the quantum walk. In order to do so it is considered an *amplificated* version of the quantum walk. More precisely, they consider the following algorithm:

Algorithm 1. 1. Choose uniformly at random T' times from $[0, T]$:

$$\{t_1, \ldots, t_{T'}\}. \tag{8.21}$$

2. For $j = 1, \ldots, T'$

   (a) Apply $U^{t_j}_{(\mathcal{G},\mathcal{S})} := (T_{(\mathcal{G},\mathcal{S})}(\mathbb{I}_{\mathcal{G}} \otimes U^{\mathcal{S}}_{G})^{t_j}$

   (b) Take the partial trace over $\mathcal{H}_{\mathcal{S}}$.

   (c) Measure $\hat{N} = \sum_g \lambda_g |g\rangle\langle g|$ with $\lambda_g \neq \lambda_h \ \forall\, g,\, h \in \mathcal{G}$, the probability of finding the walker in $g \in \mathcal{G}$ is thus given by $\hat{\mathsf{p}}^{t_j}_T(g|\psi_0)$.

   (d) Re-initialize the coin register with a random state.

Note that each step of the outer loop requires, on average $T/2$ actual steps of the total unitary dynamics $U_{(\mathcal{G},\mathcal{S})}$. In [58] it is proved that for a suitable $T = \tilde{T}_\epsilon$ the distribution outputted by the two loop algorithm is ϵ-close[2] to the uniform distribution for a number of iteration of the unitary operator $U_{(\mathcal{G},\mathcal{S})}$ given by $T'\tilde{T}_\epsilon \leqslant \mathcal{O}(n \ln n \ln \epsilon^{-1})$.

The classical simple random walk on the n-cycle with n odd requires $\Theta(n^2 \ln \epsilon^{-1})$ steps to get ϵ-close to the uniform distribution (see for example [61]), hence the coined quantum walk presents an almost quadratic speed-up.

We now show how by using this speed-up result we can propose a scheme to generate a random element over a set of non-necessarily orthogonal quantum state based on a quantum walk that outperform the classical version presented in Section 4.4.

Let us begin by considering a situation where that set we want to sample uniformly is given by a subset of the qubit states $\mathfrak{D}(\mathcal{H}_2)$. More precisely let us

[2] in total variation distance.

consider that portion of the Bloch sphere spanned by $\{\sigma_z, \sigma_x\}$ i.e. the circle $\mathcal{H}_{z,x} \in \mathfrak{D}(\mathcal{H}_2)$, hence $\mathcal{X} \simeq \mathcal{H}_{z,x}$. A rotation of angle θ around the the axis spanned by σ_y is defined as:

$$R_y(\theta) := e^{-i\frac{\theta}{2}\sigma_y}. \tag{8.22}$$

Let us assume n odd, the set of rotations:

$$\mathcal{R} := \{\mathbb{I}_2, R_y(2\pi/n), R_y(4\pi/n), \ldots, R_y(2\pi(n-1)/n)\}, \tag{8.23}$$

form a unitary group, the inverse of each element being given by:

$$R_y(2\pi l/n)^{-1} := R_y(2\pi(n-l)/n)^{-1}, \tag{8.24}$$

furthermore the group is an Abelian group (rotation among the same axis commute) and it is isomorphic to the cyclic group generate by $h := R_y(2\pi/n)$. Consider any $|\phi_e\rangle \in \mathcal{H}_{x,z}$, its orbit under the action of the group $\mathcal{R}$ we have that $|\mathrm{Orb}_{\mathcal{R}}(|\phi_e\rangle)| = |\mathcal{R}|$ as required in Section 4.4. Let us set the notation:

$$|\phi_l\rangle = a(h^l, |\phi_e\rangle) = R_y(2\pi l/n)|\phi_e\rangle. \tag{8.25}$$

Following Section 4.4 assume that at each sept t we are able to draw elements of $\mathcal{R}$ with a fixed probability distribution $\mathsf{s}(t) = \mathsf{s}$ such that:

$$\mathsf{s}_r = \begin{cases} 1/2 & \text{if } r = \{h, h^{-1}\} \\ 0 & \text{otherwise.} \end{cases} \tag{8.26}$$

then the if we set $\mathcal{S} = \{h, h^{-1}\}$ lifted dynamics take place on the Cayley graph $\Gamma = \Gamma(\mathcal{G}, \mathcal{S})$. It is apparent that the Cayley graph is given by a ring with $|\mathcal{R}|$ nodes, hence the dynamics $\mathsf{p}(t)$ is equivalent to a simple random walk on a circle with $|\mathcal{R}|$ nodes. More precisely:

$$\mathsf{p}_r(t+1) = \sum_{z\in\mathcal{R}} \mathsf{s}_z \mathsf{p}_{z^{-1}r}(t) = \frac{1}{2}\left(\mathsf{p}_{hr}(t) + \mathsf{p}_{h^{-1}r}(t)\right) \tag{8.27}$$

Hence using the result on classical random walk on a cycle we have that after $\theta(|\mathcal{R}|^2 \ln \epsilon^{-1})$ steps we are ϵ-close to the uniform distribution over the set $\{|\phi_0\rangle, \cdots, |\phi_{|\mathcal{R}|-1}\rangle\}$.

Let us now consider a version of the protocol in which the probability of selection of a group element yields from a quantum walk. We quantize the Cayley graph by associated a set of orthonormal vectors to its nodes:

$$\mathcal{H}_{\mathcal{R}} := \mathrm{span}\{|r\rangle,\, r \in \mathcal{R}\} \tag{8.28}$$

The coin register will be accordingly a two dimensional Hilbert space, at each step in fact two group elements namely $h = R_y(2\pi/n)$ and $h^{-1} = R_y(2\pi/n)^{-1}$ can be implement, now we aim to implement them in a coherent superposition rather than on a probabilistic mixture as it is done in the classical case. The coin register if thus given by $H_{\mathcal{S}} := \text{span}\{|h\rangle, |h^{-1}\rangle\}$.

In order to implement the action of the unitary operation on the qubit states we need and additional third register, we call it *physical register*, such register will have in general dimension equal to the space $\mathcal{X}$ on which the group of unitary transformations act. In this particular application we have $\mathcal{H}_P \simeq \mathcal{H}_{x,z}$. The total space is thus given by $\mathcal{H} \simeq \mathcal{H}_{\mathcal{P}} \otimes \mathcal{H}_{\mathcal{R}} \otimes \mathcal{H}_{\mathcal{S}}$. The protocol is initialized with the state $|\psi_0\rangle = |\phi_e\rangle \otimes |r\rangle \otimes |s\rangle$ with $|s\rangle$ given by some coherent superposition of $|h\rangle$ and $|h^{-1}\rangle$.
The coin operation is given by:

$$C_{(\mathcal{P},\mathcal{R},\mathcal{S})} := \mathbb{I}_{\mathcal{P}} \otimes \mathbb{I}_{\mathcal{R}} \otimes U^{\mathcal{S}}. \tag{8.29}$$

The shift operator is given by:

$$\begin{aligned} T_{(\mathcal{P},\mathcal{R},\mathcal{S})} &= \mathbb{I}_{\mathcal{P}} \otimes \sum_{s\in\mathcal{S}, r\in\mathcal{R}} |sr\rangle\langle r| \otimes |s\rangle\langle s| \\ &= \mathbb{I}_{\mathcal{P}} \otimes \sum_{r\in\mathcal{R}} \left(|hr\rangle\langle r| \otimes |h\rangle\langle h| + |h^{-1}r\rangle\langle r| \otimes |h^{-1}\rangle\langle h^{-1}| \right). \end{aligned} \tag{8.30}$$

Finally the third operator that actual implement the unitary transformations is given by:

$$A_{(\mathcal{P},\mathcal{R},\mathcal{S})} = \sum_r a(r,\cdot) \otimes |r\rangle\langle r| \otimes \mathbb{I}_{\mathcal{P}} \tag{8.31}$$

One full evolution step reads:

$$U = A_{(\mathcal{P},\mathcal{R},\mathcal{S})} \otimes T_{(\mathcal{P},\mathcal{R},\mathcal{S})} \otimes C_{(\mathcal{P},\mathcal{R},\mathcal{S})} \tag{8.32}$$

If we select uniformly at random $t \in [0, T-1]$ and take the partial trace over $\mathcal{S}$ and over $\mathcal{R}$ we have that the state on $\mathcal{H}_{x,z}$ is thus described by:

$$\sum_{r\in\mathcal{R}} \mathsf{p}^T(r|\phi_0) a(r,\phi_0) = \sum_{l=0}^{|\mathcal{R}|-1} \mathsf{p}^T(r_l|\phi_0) R_y(2\pi l/n)|\phi_0\rangle\langle\phi_0| R_y(2\pi l/n)^\dagger \tag{8.33}$$

Where $\mathsf{p}^T(r|\phi_0)$ is the probability of having implemented the group element r, i.e. of finding the walker in r. By using the algorithm described in (1) we can thus ensure that we have an almost quadratic speed up with respect to the classical case.

Conclusions and Research Directions

In the first part of this dissertation we have shown how the simple dynamics of linear gossip consensus can inspire robust iterative procedures for tasks that can be formulated as *symmetrization with respect to a finite group.* Our results offer not only a formal generalization of the well-known classical consensus problem, but also directly extend the desirable features of the consensus-type algorithms to various applications in diverse fields. We have proved convergence for a general symmetrization algorithm with either deterministic or stochastic choices of the individual iterations. We have included a selection of illustrative applications of our framework to a variety of existing problems, and we expect that in many other applications the *robustness* of the consensus formulation can be advantageously carried over to symmetrization tasks, e.g. including actions on infinite-dimensional spaces.

In the second part we have presented a detailed application of our symmetrizing framework to the consensus problem in the quantum domain focusing in particular on the gossip algorithm and we have illustrated how it could be used for distributed control and estimation problems. For most of the results regarding the convergence of our algorithms we have provided independent proofs using the tools of quantum mechanics. In particular, we have built on the statistical property of the states with respect to local observables and their symmetry with respect to permutation operations to derive four different generalizations of a consensus state to quantum systems and we have established their hierarchy. We have highlighted at each step the symmetry considerations underlying the results, making explicit connections with the usual multi-agent consensus problem. With respect to the existing work on non-commutative consensus [43], our approach follows the analogy with the classical setting as

closely as possible, maintaining an operational viewpoint and working with a multipartite system (a quantum network). We propose and analyze a quantum gossip-type algorithm that asymptotically prepares symmetric-state consensus states while preserving the expectation of *any permutation invariant observable*.

Natural directions for expanding our results include an in-depth study of convergence *speed* for specific protocols. On this regard, we report that the set of results on the rate of convergence of a specific class of random walks on the permutation group published in [62] could be employed to compute the speed of convergence of the gossip algorithm in the case of the complete graph. The development of (approximate) symmetrization procedures for infinite and uncountable groups, by replacing the linear action on a vector field by abstract algebraic structures could also offer a rewarding way to widen the scope of the results, is also a possible research direction. Furthermore, we believe that it would be particularly interesting to further explore the link between single σ-measurement consensus states and entangled states [26], and to determine if, and under which conditions, it is possible to achieve this type of consensus with a distributed algorithm. This could potentially lead to a class of algorithms that prepare entangled states in a robust and distributed way. Another interesting point is to assess the potential of continuous-time dynamics for consensus: a first dissipative proposal has been presented in [63], but we believe it would be worth exploring also time-averages of Hamiltonian dynamics, which could lead to connections with physically relevant many-body dynamics. The resemblance of the gossip interaction to the Glauber dynamics [64] might suggests that our picture could be employed as a tool to describe the thermalization of classical and even quantum systems. Lastly, a fully quantum implementation, by using a quantum walk, is definitely worth further investigation in view of potential speed-up as suggested by our proof of principle presented in Section 8.6 and by a consistent body of literature on quantum walks [58, 65, 66, 67].

A

Element of Graphs Theory and Finite Groups

A.1 Graph Theory

Here we recover some fundamental results about graph theory, for a comprehensive text on the subject see for example [19]. A graph is an ordered pair $G(V, E)$, composed by a set of nodes, $V = \{1, \dots, n\}$, and a set of edges $E \subseteq \{(u, v) : u, v \in V\}$, with $|V|$ we indicate the number of elements of V, i.e. the *order* of V. We will always assume the order of V to be finite. A graph is said undirected if it holds that:

$$(u, v) \in E \iff (v, u) \in E, \tag{A.1}$$

in an undirected graph we can hence identify (u, v) with (v, u). Here we will restrict ourself only to undirected graphs.

Given a node of the graph u we define its *neighbors* as the set of nodes that share an edge of the graph with u:

$$\mathcal{N}(u) := \{v \in V \backslash \{u\} : (u, v) \in E\} \tag{A.2}$$

we also define the degree of a node as the number of its neighbors $d(v) = |\mathcal{N}(v)|$.

The *adjacency matrix* of a graph is defined as:

$$\{A_G\}_{i,j=1,\dots,n} := \begin{cases} 1 & \text{if } (i, j) \in E \\ 0 & \text{otherwise.} \end{cases} \tag{A.3}$$

It follows that the adjacency matrix of an undirected graph is symmetric.

The *degree matrix* of a graph is a diagonal matrix defined as:

$$D_G := \text{diag}\{d(v_1), \dots, d(v_n)\} \tag{A.4}$$

A graph is said regular if is every nodes has the same number of neighbors, i.e. $d(v) = d \; \forall \, v \in V$. If a graph is regular the degree matrix is proportional to the identity matrix.

Definition 19 (Bipartite Graph)**.** A graph is bipartite if its set of vertices V, can be partitioned into two non overlapping subsets A, B:

$$V = A \cup B \quad \text{and} \quad A \cap B = \emptyset, \tag{A.5}$$

such that every edge of the graph connects a vertex of A to a vertex of B.

It can also be formulated a more operative condition to check whenever a graph is bipartite:

Proposition 15. A Bipartite graph does not contain any odd-length cycles.

In many application is important to be able to broadcast information between arbitrary nodes among the graph. If this is the case we speak of a *connect graph*. More precisely:

Definition 20 (path)**.** A path in a graph G from $u \in V$ to $v \in V$ is a sequence of distinct vertices starting in u and ending in v such that consecutive vertices are adjacent, i.e. are connected by an edge of the graph.

Definition 21 (Connected Graph)**.** An undirected graph is connected if there is at least a path between each pair of vertices of the graph.

Example 10. The odd-ring graph is connected and not bipartite while the even-ring graph is strongly connected but not bipartite.

Among the connects graphs the are those graph in which each pair of vertices is connected by exactly one path, a tree by construction aperiodic.

It is worth recalling an important relation regarding between the number of nodes and edges of a graph and its connectivity properties:

Proposition 16 ([19])**.** For a connected graph is holds that:

$$|E| \geqslant |V| - 1 \tag{A.6}$$

and the equality holds if the graph is a tree. The trees are thus graphs with minimal connectivity.

A.2 Finite Groups

In this section we briefly recall some properties of finite group and of their representation, see e.g. [18, 19, 68]. We will use them in some proofs.

Definition 22. A finite group $\mathcal{G}$ is a set with a binary operation or composition law $\cdot$ that satisfied the following axioms:

- **Closure:** $g \cdot h \in \mathcal{G}$ for every $g, h \in \mathcal{G}$.
- **Associativity:** for every $g, h, w \in \mathcal{G}$ $(g \cdot h) \cdot w = g \cdot (h \cdot w)$.
- **Identity:** there exist a unique element $e_{\mathcal{G}}$ such that $g \cdot e_{\mathcal{G}} = e_{\mathcal{G}} \cdot g$ for every $g \in \mathcal{G}$.
- **Inverse element:** for every $g \in \mathcal{G}$ there exist $g^{-1} \in \mathcal{G}$ such that $g^{-1} \cdot g = g \cdot g^{-1} = e_{\mathcal{G}}$.

Let $\mathcal{G}$ be a finite group. We denote the order of the group, i.e. the number of its element, as $|\mathcal{G}|$.

Definition 23 (Abelian group)**.** a group is said Abelian if:

$$g \cdot h = h \cdot g \qquad \forall\, g, h \in \mathcal{G}. \tag{A.7}$$

For sake of simplicity we will drop the $\cdot$ to indicate the group operation, i.e. $g \cdot h := gh$.

Definition 24 (Generating set)**.** A subset S of $\mathcal{G}$ is a generating set for the group if every element of $\mathcal{G}$ can be written as a composition, under the group operation, of elements of S.

Definition 25 (Minimal generating set)**.** A generating set $\mathcal{S}$ of the group $\mathcal{G}$is said minimal if every proper subset of $\mathcal{S}$ generates a proper subgroup of $\mathcal{G}$.

Let $\mathcal{X}$ be a vector space over a field $\mathbb{F} = \{\mathbb{R}, \mathbb{C}\}$, we denote the group of linear isomorphism on $\mathcal{X}$ with $\mathrm{GL}(\mathcal{X})$. If V is endowed with an inner-product operation:

$$\langle\ ,\ \rangle : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{F}, \tag{A.8}$$

it is called *inner-product space*.

Definition 26 (Group homomorphism)**.** Given two groups $\mathcal{G}$ and $\mathcal{F}$ a group homomorphism from $\mathcal{G}$ to $\mathcal{F}$ is a function $f : \mathcal{G} \to \mathcal{F}$ such that for every $g, h \in \mathcal{G}$:

- $f(gh) = f(g)f(h)$,
- $f(g^{-1}) = f(g)^{-1}$.

An homomorphism that maps a group on itself is called *endomorphism* and if it is bijective *automorphism*.

Definition 27 (Representation of a group)**.** A representation of a group $\mathcal{G}$ is a map:

$$D: \mathcal{G} \longrightarrow \mathrm{GL}(\mathcal{X}) \tag{A.9}$$

for some finite-dimensional complex vector space $\mathcal{X}$, such that for every $g, h \in \mathcal{G}$

- $D(gh) = D(g)D(h)$,
- $D(g^{-1}) = D(g)^{-1}$,
- $D(e_{\mathcal{G}}) = \mathbb{I}_{\mathcal{X}}$.

Where e is the identity element of $\mathcal{G}$ and $\mathbb{I}_{\mathcal{X}}$ is the identity automorphism over $\mathcal{X}$. The map D is an homomorphism of $\mathcal{G}$ into $\mathcal{X}$. The advantage of working with a representation is that each element of $\mathrm{GL}(\mathcal{X})$ can be represented with a finite dimensional matrix.

The dimension of $\mathcal{X}$ is called *degree* or *dimension* of the representation D. If the homomorphism is injective the representation is *faithful*.

Proposition 17 ([68])**.** Every finite groups admit a finite dimensional representation.

Definition 28 (Equivalent representations)**.** Let $D : \mathcal{G} \to GL(\mathcal{X})$ and $R : \mathcal{G} \to GL(\mathcal{Y})$ be two representations of $\mathcal{G}$. D and R are said to be equivalent if there exists an isomorphism $T : \mathcal{X} \to \mathcal{Y}$ such that $R(g) = TD(g)T^{-1}$ for every $g \in \mathcal{G}$. In this case we write $D \sim R$.

Definition 29 (Unitary representation)**.** Let $\mathcal{X}$ be an inner product space. The representation $D : \mathcal{G} \to GL(\mathcal{X})$ is said to be unitary if for every $g \in \mathcal{G}$:

$$\langle D(g)u, D(g)v\rangle = \langle u, v\rangle \quad \forall\, u, v \in \mathcal{X}, \tag{A.10}$$

i.e. if $D(g)$ is a unitary operator for every $g \in \mathcal{G}$.

For our purposes will be important the following result.

Proposition 18 ([68])**.** Every representation of a finite group is equivalent to a unitary operation.

One representation with linearly independent elements is the *regular representation*.

Definition 30 (Regular representation)**.** Consider $\mathcal{X} = \mathbb{R}^{|\mathcal{G}|}$, index the vectors of the canonical basis of $\mathcal{X}$ by $\{\, v(g) \in \mathcal{X} : v(g)_h = \delta_{h,g} \;\; \forall\, g, h \in \mathcal{G} \,\}$ where $\delta_{g,h}$ is the Kronecker delta and define the linear action of $\mathcal{G}$ on $\mathcal{X}$ by $a(h, v(g)) = v(hg)$ for all $g, h \in \mathcal{G}$.

To see that the actions associated to different $h \in \mathcal{G}$ are all linearly independent, it suffices to notice that $a(h, v(\mathrm{e}_{\mathcal{G}})) = v(h)$, i.e. they are in one to one correspondence with a set of orthogonal vectors.

Definition 31 (linear action)**.** The *linear action* of $\mathcal{G}$ on $\mathcal{X}$, that is a linear map:

$$a : \mathcal{G} \times \mathcal{X} \to \mathcal{X}, \tag{A.11}$$

such that, for all $x, y \in \mathcal{X}$ and all $g, h \in \mathcal{G}$:

$$\begin{aligned} a(g, \alpha x + \beta y) &= \alpha a(g, x) + \beta a(g, y) \quad \alpha, \beta \in \mathbb{F}, && \text{(A.12)} \\ a(hg, x) &= a(h, a(g, x)), && \text{(A.13)} \\ a(\mathrm{e}_{\mathcal{G}}, x) &= x. && \text{(A.14)} \end{aligned}$$

where $\mathrm{e}_{\mathcal{G}}$ is the identity of $\mathcal{G}$.

Definition 32. $\bar{x} \in \mathcal{X}$ is a fixed point of the action of the group $\mathcal{G}$ if

$$a(g, \bar{x}) = \bar{x} \quad \forall\, g \in \mathcal{G}. \tag{A.15}$$

We will denote the set of fixed points as $\mathcal{F}^{\mathcal{G}} \subseteq X$.

Let us denote with $\mathrm{Aut}(\mathcal{X})$ the group of the automorphism on the inner-product space $\mathcal{X}$.

Proposition 19. The set $\mathcal{A} = \{a(g, \cdot)\}_{g \in \mathcal{G}}$ is a subgroup of $\mathrm{Aut}(\mathcal{X})$.

Proof. For any $h, g \in \mathcal{G}$ we have that:

$$a(h, \cdot) \circ a(g, \cdot) = a(h, a(g, \cdot)) = a(hg, \cdot) \in \mathcal{A}, \tag{A.16}$$

and for any $a(g, \cdot) \in \mathcal{A}$ there exist $a^{-1}(g, \cdot) \in \mathcal{A}$ s.t.:

$$a(g, \cdot) \circ a^{-1}(g, \cdot) = a^{-1}(g, \cdot) \circ a(g, \cdot) = a(e, \cdot), \tag{A.17}$$

such element is given by $a^{-1}(g, \cdot) = a(g^{-1}, \cdot)$. □

Proposition 20. The map $D : \mathcal{G} \to \mathcal{A} \subset \mathrm{Aut}(\mathcal{X})$:

$$D(g) = a(g, \cdot) \tag{A.18}$$

is a representation of $\mathcal{G}$.

Proof. For every $g, h \in \mathcal{G}$ it holds that:

- $D(gh) = a(gh, \cdot) = a(g, \cdot) \circ a(h, \cdot) = D(g)D(h)$,
- $D(g^{-1}) = a(g^{-1}, \cdot) = a^{-1}(g, \cdot) = D(g)^{-1}$,
- $D(e_{\mathcal{G}}) = a(e_{\mathcal{G}}, \cdot) = I_X$,

□

Given $a(g, \cdot)$ (or a general element of $\mathrm{Aut}(\mathcal{X})$) we can define its adjoint with respect to the inner product as the unique operator that satisfies:

$$\langle a(g, x), y\rangle = \langle x, a^{\dagger}(g, y)\rangle \quad \forall\, x, y \in X. \tag{A.19}$$

Previously we have pointed out that every representation of a finite group is equivalent to a unitary representation:

$$U : \mathcal{A} \longrightarrow \mathfrak{U}(\mathcal{X}), \tag{A.20}$$

where $\mathfrak{U}(\mathcal{X})$ is the set of unitary operator over $\mathcal{X}$. Since $\mathcal{A}$ hosts a representation of $\mathcal{G}$ we will denote the representative of an action as $U(g)$ where g is the group element that generates the action. Since $U(g)$ is unitary for every $g \in \mathcal{G}$ we have that:

$$a^{\dagger}(g, \cdot) = U(g)^{\dagger} = U(g)^{-1} = a^{-1}(g, \cdot) = a(g^{-1}, \cdot). \tag{A.21}$$

Hence $\mathcal{A}$ is closed under conjugation.

Let us now recall some properties of the *permutation group* [18]. Through this dissertation we make use of relevant permutation groups and of their representation in order to prove many of our results.

Consider a non empty set of objects $\Xi = \{1, \ldots, n\}$, a bijection from Ξ to itself is called *permutation of* Ξ. The set of all permutations of Ξ forms a group with composition law given by the composition of maps, we donate it wither with $\mathfrak{P}_{\Xi}$ or with $\mathfrak{P}_n$. The order of the permutation group of n objects is $n!$. The permutation group admits various sets of generators. We are most interested in the generating sets given by the *transpositions of* Ξ, i.e. those permutations that exchanges two elements and left all the others unchanged.

A useful representation of $\mathfrak{P}_n$ is given in terms of the so called the *permutation matrices*. The representation is hosted in the space of the $n \times n$ real matrices. Consider the canonical basis $\{e_i\}_{i=1}^n$, each permutation $\pi \in \mathfrak{P}_n$ is associated to a $n \times n$ matrix P_π given by:

$$P_\pi := \begin{pmatrix} e_{\pi(1)}^T \\ e_{\pi(2)}^T \\ \dots \\ e_{\pi(n)}^T \end{pmatrix}. \tag{A.22}$$

We have that every P_π is a matrix in which the only entries different from zero are those such that $(i, \pi(i))$ and their values is equal to one. It is easy to see that the latter is a valid representation. For every $\pi, \sigma \in \mathfrak{P}_n$ we have, in fact:

$$P_{\pi\circ\sigma} = P_\pi P_\sigma, \tag{A.23}$$

$$P_\pi^T = P_\pi^{-1} = P_{\pi^{-1}}. \tag{A.24}$$

The latter relation can be established by noting that:

$$P_\pi^T = \begin{pmatrix} e_{\pi(1)}, & e_{\pi(2)}, & \dots, & e_{\pi(n)} \end{pmatrix}, \tag{A.25}$$

and by keeping into account that, if $i \neq j$ than $\pi(i) \neq \pi(j)$ being π a bijection. Note that permutation matrices are isometries.

The relations between a finite group and one of its subsets of element can be addressed by considering its Cayley graph [19].

Definition 33 (Cayley graph)**.** Let $\mathcal{G}$ be a finite group and let $\mathcal{S} \subseteq \mathcal{G}$. Let us also consider $E_\mathcal{S} \subseteq \mathcal{G} \times \mathcal{G}$:

$$E_\mathcal{S} = \{(g, h) \in \mathcal{G} \times \mathcal{G} \text{ if } \exists\, s \in \mathcal{S}\; sh = g\}. \tag{A.26}$$

The Cayley graph of $\mathcal{G}$ relative to $\mathcal{S}$ is the graph defined as:

$$\Gamma := \Gamma(\mathcal{G}, E_\mathcal{S}) \tag{A.27}$$

Note that if $\mathcal{S} = \mathcal{S}^{-1}$ the graph Γ is undirected, the graph is $|\mathcal{S}|$-regular. Furthermore it hold the following result:

Proposition 21 ([19])**.** The Cayley graph of the finite group $\mathcal{G}$ generated by $\mathcal{S} \subset \mathcal{G}$ is connected if and only if $\mathcal{S}$ is a generator set of $\mathcal{G}$.

B

Proofs of Results in Section 6

B.1 Proof of Proposition 10

By considering an hermitian basis for $\mathfrak{B}(\mathcal{H})$, it is clear that a state is RSC if it is σEC for all $\sigma \in \mathfrak{B}(\mathcal{H})$.

If $\bar{\rho}_k$ is a pure state for each k, then necessarily $\rho = |\psi\rangle\langle\psi|$ with $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes ... \otimes |\psi_m\rangle$ for some $|\psi_k\rangle$, $k = 1, 2, ..., m$. If in addition we require RSC, then we need $|\psi_k\rangle\langle\psi_k| = \bar{\rho}_k = \bar{\rho}_j = |\psi_j\rangle\langle\psi_j|$ for all j, k, thus $|\psi_k\rangle = |\psi_j\rangle$ up to an irrelevant phase factor for all j, k and particle permutation indeed leaves $|\psi\rangle$ invariant. □

B.2 Proof of Theorem 5

Note that the properties $\mathrm{Tr}(\Pi_{\mathrm{sym}}\rho) = 1$ and $\Pi_{\mathrm{sym}}\rho\Pi_{\mathrm{sym}} = \Pi_{\mathrm{sym}}\rho = \rho$ are equivalent because Π_{sym} is an orthonormal projector and ρ is self-adjoint positive semi-definite with unit trace. Assume (6.7) to hold. Along with the identities $\Pi_j^{(k)}\Pi_{\mathrm{sym}} = \Pi_j^{(k)}\Pi_j^{\otimes m} = \Pi_j^{\otimes m}$, this gives:

$$\begin{aligned}\mathrm{Tr}(\Pi_j^{(\ell)}\Pi_j^{(k)}\rho) &= \mathrm{Tr}(\Pi_j^{(\ell)}\Pi_j^{(k)}\Pi_{\mathrm{sym}}\rho) = \mathrm{Tr}(\Pi_j^{(\ell)}\Pi_j^{\otimes m}\rho) \\ &= \mathrm{Tr}(\Pi_j^{(\ell)}\Pi_{\mathrm{sym}}\rho) = \mathrm{Tr}(\Pi_j^{(\ell)}\rho)\,,\end{aligned}$$

for all j, k, ℓ. Hence, the σSMC Definition (6.6) indeed holds.
On the other hand, suppose that (6.7) does not hold. This means that $\mathrm{Tr}((I - \Pi_{\mathrm{sym}})\rho) > 0$. We want to show that this implies

$$\mathrm{Tr}(\Pi_j^{(k)}\Pi_j^{(\ell)}\rho) \neq \mathrm{Tr}(\Pi_j^{(\ell)}\rho)$$

for some j, k, ℓ. Let us write

$$I - \Pi_{\rm sym} = \sum_{\substack{j_1, \ldots, j_m \in \{1,2,...,d\} \\ \text{except } \{j_1 = \ldots = j_m\}}} \Pi_{j_1} \otimes \ldots \otimes \Pi_{j_m}.$$

Since $\mathrm{Tr}((I - \Pi_{\rm sym})\rho) > 0$ implies that $\mathrm{Tr}(\Pi_{j_1} \otimes \ldots \otimes \Pi_{j_m}\rho) > 0$ for at least one of the terms in the above sum, let us take one such term, denote the corresponding indices as $\{\bar{j}_s\}$ and denote by k, ℓ two subsystems such that $\bar{j}_k \neq \bar{j}_\ell$ in that term. Now writing $\Pi_{\bar{j}_1} \otimes \ldots \otimes \Pi_{\bar{j}_m} = \Pi^{(1)}_{\bar{j}_1} \Pi^{(2)}_{\bar{j}_2} ... \Pi^{(m)}_{\bar{j}_m}$, where all factors commute, we have:

$$\mathrm{Tr}(\Pi^{(k)}_{\bar{j}_k} \Pi^{(\ell)}_{\bar{j}_\ell} \rho) \geq \mathrm{Tr}(\Pi^{(1)}_{\bar{j}_1} \Pi^{(2)}_{\bar{j}_2} ... \Pi^{(m)}_{\bar{j}_m} \rho) > 0.$$

By mutual orthogonality of $\{\Pi^{(k)}_{\bar{j}_a} : a = 1, 2, ..., m\}$ for fixed k, and knowing that the trace of an operator cannot increase under multiplication by a projection operator, we thus get:

$$\begin{aligned} \mathrm{Tr}(\Pi^{(k)}_{\bar{j}_\ell} \Pi^{(\ell)}_{\bar{j}_\ell} \rho) &\leq \mathrm{Tr}((1 - \Pi^{(k)}_{\bar{j}_k}) \Pi^{(\ell)}_{\bar{j}_\ell} \rho) \\ &= \mathrm{Tr}(\Pi^{(\ell)}_{\bar{j}_\ell} \rho) - \mathrm{Tr}(\Pi^{(k)}_{\bar{j}_k} \Pi^{(\ell)}_{\bar{j}_\ell} \rho) \\ &\leq \mathrm{Tr}(\Pi^{(\ell)}_{\bar{j}_\ell} \rho) - \mathrm{Tr}(\Pi^{(1)}_{\bar{j}_1} \Pi^{(2)}_{\bar{j}_2} ... \Pi^{(m)}_{\bar{j}_m} \rho) \\ &< \mathrm{Tr}(\Pi^{(\ell)}_{\bar{j}_\ell} \rho)\,. \end{aligned}$$

For (a), we have since (6.6) holds for all k, ℓ:

$$\mathrm{Tr}(\Pi^{(k)}_j \rho) = \mathrm{Tr}(\Pi^{(k)}_j \Pi^{(\ell)}_j \rho) = \mathrm{Tr}(\Pi^{(\ell)}_j \rho).$$

By linearity, we thus have:

$$\begin{aligned} \mathrm{Tr}(\sigma^{(k)} \rho) &= \sum_{j=1}^d s_j \mathrm{Tr}(\Pi^{(k)}_j \rho) = \sum_{j=1}^d s_j \mathrm{Tr}(\Pi^{(\ell)}_j \rho) \\ &= \mathrm{Tr}(\sigma^{(\ell)} \rho). \end{aligned}$$

A counterexample for the converse is state ρ^A in Example 1.

Counterexamples for the converse of (b) and (c) are respectively states ρ^B and ρ^C in Example 1. For the direct statements, given Proposition 4, we know that if (c) is true, then (b) must be true as well. Let us then focus on (c). Take the representation of ρ in the basis associated to $\sigma = \sum_{j=1}^n s_j |j\rangle\langle j|$, where thus $|j\rangle\langle j| = \Pi_j$, that reads

$$\rho = \sum_{\substack{j_1, j_2, ... j_m \in D, \\ k_1, k_2, ... k_m \in D}} r^{\,j_1, j_2, ... j_m}_{\,k_1, k_2, ... k_m} \, |j_1, j_2, ..., j_m\rangle\langle k_1, k_2, ..., k_m|\,,$$

with $D = \{1, 2, ..., n\}$. From Proposition 5, the condition for σSMC writes

$$\sum_{k,j\in D} (|k\rangle\langle k|)^{\otimes m}\, \rho\, (|j\rangle\langle j|)^{\otimes m} = \rho\,,$$

so (B.1) must reduce to

$$\rho = \sum_{k,j\in D} p_{kj} |kk\ldots k\rangle\langle jj\ldots j| \tag{B.1}$$

for some $p_{kj} \in \mathbb{C}$. It is straightforward to see that a ρ of this form satisfies SSC, since any element in the sum is invariant w.r.t. subsystem permutations. Regarding point (e), the definition of σSMC involves $\mathrm{Tr}(\Pi_j^{(k)}\Pi_j^{(\ell)}\rho)$, which takes the partial trace over the state of all subsystems except the pair $\{k, \ell\}$. So we can effectively discard all but two subsystems, and show without loss of generality that it is impossible to make σSMC hold for all σ on two subsystems $k = 1$, $\ell = 2$. In Proposition 5, we say that σSMC for a particular σ requires $\Pi_{\mathrm{sym}}\rho = \rho$ with $\Pi_{\mathrm{sym}} = \sum_{j=1}^{d} \Pi_j^{\otimes m}$, and $\{\Pi_j\}$ the spectral projectors associated to σ. So if σSMC has to hold for both σ and σ', we must have in particular

$$\Pi_{\mathrm{sym}}\Pi'_{\mathrm{sym}}\Pi_{\mathrm{sym}}\rho = \rho\,,$$

where Π'_{sym} is associated to σ'. Since $H := \Pi_{\mathrm{sym}}\Pi'_{\mathrm{sym}}\Pi_{\mathrm{sym}}$ and ρ both are self-adjoint positive semidefinite, the only way to have $H\rho = \rho \neq 0$ is if H has at least one eigenvalue ≥ 1. Now take in particular $\sigma = \sum_{k=1}^{n} k\, |x_k\rangle\langle x_k|$ and $\sigma' = \sum_{k=1}^{n} k\, |p_k\rangle\langle p_k|$, with $p_k = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} e^{jk2\pi i/n} |x_{j+1}\rangle$ (thus the $|p_k\rangle$-basis is related to the $|x_k\rangle$-basis by Fourier transform). A few computations show that H then has all eigenvalues < 1, except for $n = 2$ that is the case of two qbits. For the latter particular case, one can prove the property by showing e.g. that there is no state which would satisfy σSMC for all $\sigma \in \{\sigma_x, \sigma_y, \sigma_z\}$. □

B.3 On Detecting Quantum Consensus

In the quantum setting, there exists no state ρ for which *all* measurement outcomes are deterministically defined. Even a maximal information state, i.e. $\rho = |\psi\rangle\langle\psi|$ a rank one projector, leads to probabilistic outcomes for all observables of which $|\psi\rangle$ is not an eigenstate. As we are thus compelled to use probabilistic notions, consensus can only be inferred from stochastic measurement records, and checking different types of consensus requires different types of measurement statistics.

The σEC, requiring only equal expectations for a particular observable σ on the different subsystems, simply requires measurements of local σ-measurements results, but no correlations between measurement results on different subsystems are needed. Checking RSC requires statistics for a basis of observables for each subsystem; as for σEC, correlations between measurement results on different subsystems play no role. On the other hand, distinguishing SSC from RSC does require to inspect correlations between measurement outcomes at different subsystems.

Proposition 22. Except for the case of reduced pure states considered in Proposition 10, SSC can only be distinguished from RSC by inspecting correlations between measurement outcomes at different subsystems.

Proof. The statement builds on the standard fact that the statistics of a local observable $\sigma_1 \otimes \sigma_2 \otimes ... \otimes \sigma_m$ only depend on reduced states $\bar{\rho}_1, \bar{\rho}_2, ..., \bar{\rho}_m$. So repeated local measurements can, at their best, fully characterize the $\bar{\rho}_k$. Checking RSC, i.e. that these $\bar{\rho}_k$ are all equal, is thus straightforward. On the other hand, reduced states $\bar{\rho}_k$ are the best that can be extracted by local measurements in trying to distinguish RSC from SSC states. If $\bar{\rho}_1 = \bar{\rho}_2 = ... =: \bar{\rho}$ have rank one, we have the special case that is always SSC. If instead $\bar{\rho}$ has rank at least 2, we can write it as $\bar{\rho} = p_1 R_1 + p_2 R_2$ where R_1, $R_2 \in \mathfrak{B}(\mathcal{H})$, p_1, p_2 are positive scalars, R_2 is positive semidefinite, and R_1 is a projector on a 2-dimensional subspace $\mathcal{V}_2$. Consider $R_1 = |e_1\rangle\langle e_1| + |e_2\rangle\langle e_2| = |f_1\rangle\langle f_1| + |f_2\rangle\langle f_2|$, where $|e_1\rangle, |e_2\rangle$ and $|f_1\rangle, |f_2\rangle$ are two orthonormal bases for $\mathcal{V}_2$ with $\langle e_1|f_1\rangle \notin \{0, 1\}$. Then $\bar{\rho}$ could equally well reflect the state

$$\rho = \bar{\rho}^{\otimes m} ,$$

which is SSC, or e.g. a state of the form:

$$\rho = p_2\, R_2^{\otimes m} + \; p_1(|e_1\rangle|f_1\rangle + |e_2\rangle|f_2\rangle)(|e_1\rangle|f_1\rangle + |e_2\rangle|f_2\rangle)^\dagger \otimes R_1^{\otimes(m-2)},$$

where the first two subsystems are entangled. This state is not SSC, even for $m = 2$. Thus the local knowledge of $\bar{\rho}$ does not allow to distinguish if the state is SSC or not. □

For instance, considering the state ρ^B of Example 1, measurements of σ_z on the three subsystems would quickly show that the results on subsystems 2 and 3 are always perfectly correlated, and show no correlation at all with the results on the first subsystem. This difference in correlations rules out ρ^B as a candidate for SSC. The definition of σSMC is all about correlations between measurement outcomes at different subsystems: the latter must be fully correlated for a particular observable σ. Positively detecting states in SSC but not in SMC, however, appears to be less obvious (except through full state tomography).

Bibliography

[1] D. Kempe, A Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*, pages 482–491, Oct 2003. 17

[2] X. Lin, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. In *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*, pages 63–70, April 2005. 17

[3] J. Zhao, R. Govindan, and D. Estrin. Computing aggregates for monitoring wireless sensor networks. In *Sensor Network Protocols and Applications, 2003. Proceedings of the First IEEE. 2003 IEEE International Workshop on*, pages 139–148, May 2003. 17

[4] L. S. Smith, M. E. Broucke, and B. A. Francis. A hierarchical cyclic pursuit scheme for vehicle networks. *Automatica*, 41(6):1045 – 1053, 2005. 17

[5] R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time delays. *IEEE Trans. Automatic Control*, 49(9):1520–1533, 2004. 17, 44, 54

[6] L. Moreau. Stability of multiagent systems with time-dependent communication links. *Automatic Control, IEEE Transactions on*, 50(2):169–182, Feb 2005. 17

[7] A Jadbabaie, Jie Lin, and AS. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *Automatic Control, IEEE Transactions on*, 48(6):988–1001, June 2003. 17

[8] S. Muthukrishnan, B. Ghosh, and M. H. Schultz. First- and second-order diffusive methods for rapid, coarse, distributed load balancing. *Theory of Computing Systems*, 31(4):331–354, 1998. 17

[9] R. Diekmann, A. Frommer, and B. Monien. Efficient schemes for nearest neighbor load balancing. *Parallel Computing*, 25(7):789 – 812, 1999. 17

[10] G. Cybenko. Dynamic load balancing for distributed memory multiprocessors. *Journal of Parallel and Distributed Computing*, 7(2):279 – 301, 1989. 17

[11] R. Olfati-Saber, J.A Fax, and R.M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, Jan 2007. 17, 26

[12] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE/ACM Trans. Netw.*, 14(SI):2508–2530, June 2006. 17, 18, 22, 24, 25, 51

[13] F. Fagnani and S. Zampieri. Randomized consensus algorithms over large scale networks. *Selected Areas in Communications, IEEE Journal on*, 26(4):634–649, May 2008. 17, 18, 22

[14] F. Garin and L. Schenato. A survey on distributed estimation and control applications using linear consensus algorithms. In Alberto Bemporad, Maurice Heemels, and Mikael Johansson, editors, *Networked Control Systems*, volume 406 of *Lecture Notes in Control and Information Sciences*, pages 75–107. Springer London, 2010. 17

[15] L. Elsner, I. Koltracht, and M. Neumann. On the convergence of asynchronous paracontractions with application to tomographic reconstruction from incomplete data. *Linear Algebra and its Applications*, 130(0):65 – 82, 1990. 22

[16] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, USA, third edition, 2002. 25

[17] L. Moreau. Stability of multi-agent systems with time-dependent communication links. *IEEE Trans. Automatic Control*, 50(2):169–182, 2005. 25, 44, 51, 81

[18] J.D. Dixon and B. Mortimer. Permutation groups. volume 163 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1996. 28, 115, 118

[19] C. D. Godsil and Gordon. Royle. *Algebraic graph theory*. Springer New York, 2001. 28, 29, 33, 113, 114, 115, 119

[20] J.N. Tsitsiklis and M. Athans (advisor). Problems in decentralized decision making and computation. *PhD Thesis, MIT*, 1984. 44, 51

[21] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Automatic Control*, 48(6):988–1001, 2003. 44

[22] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 1991. 45

[23] R. Olfati-Saber, J.A. Fax, and R.M. Murray. Consensus and cooperation in networked multi-agent systems. *Proc. IEEE*, 95(1):215–233, 2007. 51

[24] G. Birkhoff. Three observations on linear algebra. *Univ. Nac. Tucuan. Revista A*, 5:147–151, 1946. 54

[25] B.A. Berg, D.P. Landau, W.S. Kendall, R. Chen, and E.A. Thompson. *Markov Chain Monte Carlo: innovations and applications*. World Scientific Publishing, 2005. 56

[26] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Information*. Cambridge University Press, Cambridge, 2002. 56, 57, 61, 68, 70, 79, 86, 112

[27] J. Emerson, Y. Weinstein, M. Saraceno, S. Lloyd, and D. Cory. Pseudo-random unitary operators for quantum information processing. *Science*, 302:2098–2100, 2003. 56, 57

[28] J. Emerson, E. Livine, and S. Lloyd. Convergence conditions for random quantum circuits. *Physical Review A*, 72(6):060302, 2005. 56, 57

[29] E. Knill and R. Laflamme. Theory of quantum error-correcting codes. *Phys. Rev. A*, 55(2):900–911, Feb 1997. 59

[30] R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola. Information preserving structures: A general framework for quantum zero-error information. *preprint*, arxiv:1006.1358v1, 2010. 59, 85

[31] L. Viola and E. Knill. Random decoupling schemes for quantum dynamical control and error suppression. *Phys. Rev. Lett.*, 94:060502, 2005. 59, 102

[32] F. Ticozzi and L. Viola. Quantum information encoding, protection and correction via trace-norm isometries. *Phys. Rev. A*, 81(3):032313, 2010. 59

[33] S. Lloyd and L. Viola. Engineering quantum dynamics. *Phys. Rev. A*, 65:010101:1–4, 2001. 59, 81

[34] J. T. Barreiro, M. Muller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt. An open-system quantum simulator with trapped ions. *Nature*, 470:486–491, 2011. 59, 84

[35] F. Verstraete, M. M. Wolf, and J. I. Cirac. Quantum computation and quantum-state engineering driven by dissipation. *Nature Physics*, 5:633 – 636, 2009. 59, 84

[36] B. Kraus, S. Diehl, A. Micheli, A. Kantian, H. P. Büchler, and P. Zoller. Preparation of entangled states by dissipative quantum markov processes. *Phys. Rev. A*, 78(4):042307, 2008. 59

[37] F. Ticozzi and L. Viola. Stabilizing entangled states with quasi-local quantum dynamical semigroups. *Phil. Trans. R. Soc. A*, 370(1979):5259–5269, 2012. 59, 80

[38] C. Altafini and F. Ticozzi. Modeling and control of quantum systems: An introduction. *IEEE Trans. Aut. Cont,*, 57(8):1898 –1917, 2012. 59, 61

[39] H. M. Wiseman and G. J. Milburn. *Quantum Measurement and Control.* Cambridge University Press, 2009. 59, 61, 70, 101

[40] J. Gough and M. R. James. The series product and its application to quantum feedforward and feedback networks. *IEEE Trans. Aut. Contr*, 54(11):2530 –2544, 2009. 59

[41] R. F. Werner. Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, 1989. 59

[42] T. Eggeling and R. F. Werner. Separability properties of tripartite states with $u \bigotimes u \bigotimes u$ symmetry. *Phys. Rev. A*, 63:042111, 2001. 59

[43] R. Sepulchre, A. Sarlette, and P. Rouchon. Consensus in non-commutative spaces. *Proc. 49th IEEE Conf. Decision & Control*, pages 6596–6601, 2010. 60, 111

[44] J. J. Sakurai. *Modern Quantum Mechanics*. Addison-Wesley, New York, 1994. 61, 79, 101

[45] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory*. Cambridge University Press, 2012. 61, 63, 64, 65, 67

[46] W. Magnus. On the exponential solution of differential equations for a linear operator. *Commun. Pure and Appl. Math.*, 7:649–673, 1954. 69, 101

[47] R. Alicki and K. Lendi. *Quantum Dynamical Semigroups and Applications*. Springer-Verlag, Berlin, 1987. 70, 86

[48] K. Kraus. *States, Effects, and Operations: Fundamental Notions of Quantum Theory*. Lecture notes in Physics. Springer-Verlag, Berlin, 1983. 70, 79, 86

[49] C. B. Mendl and M. M. Wolf. Unital quantum channels - convex structure and revivals of birkhoff's theorem. *Commun. Math. Phys.*, 289:1057–1096, 2009. 79

[50] F. Albertini and F. Ticozzi. Discrete-time controllability for feedback quantum dynamics. *Automatica*, 47(11):2451 – 2456, 2011. 81, 98

[51] S. Bolognani and F. Ticozzi. Engineering stable discrete-time quantum dynamics via a canonical QR decomposition. *IEEE Trans. Aut. Contr.*, 55(12):2721 –2734, 2010. 98

[52] L. Viola, E. Knill, and S. Lloyd. Dynamical decoupling of open quantum system. *Phys. Rev. Lett.*, 82(12):2417–2421, 1999. 101

[53] P. Zanardi. Symmetrizing evolutions. *Phys. Lett. A*, 258:77, 1999. 101

[54] L. F. Santos and L. Viola. Enhanced convergence and robust performance of randomized dynamical decoupling. *Phys. Rev. Lett.*, 97(15):150501, 2006. 102

[55] K.Khodjasteh and D.A.Lidar. Fault-tolerant quantum dynamical decoupling. *Phys. Rev. Lett.*, 95(18):180501, 2005. 102

[56] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev. The security of practical quantum key distribution. *Rev. Mod. Phys.*, 81:1301–1350, Sep 2009. 104

[57] D. Aharonov, A. Ambainis, J. Kempe, and U. Vazirani. Quantum walks on graphs. In *Proceedings of the Thirty-third Annual ACM Symposium on Theory of Computing*, STOC '01, pages 50–59, New York, NY, USA, 2001. ACM. 105

[58] J. Kempe. Quantum random walks: An introductory overview. *Contemporary Physics*, 44(4):307–327, 2003. 105, 106, 107, 112

[59] B. Tregenna, W. Flanagan, R. Maile, and V. Kendon. Controlling discrete quantum walks: coins and initial states. *New Journal of Physics*, 5(1):83, 2003. 105

[60] H. Gerhardt and J. Watrous. Continuous-time quantum walks on the symmetric group. In *Proc. RANDOM-APPROX (Sanjeev*, pages 290–301. SpringerVerlag, 2003. 106

[61] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2006. 107

[62] P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57(2):159–179, 1981. 112

[63] F. Ticozzi L. Mazzarella and A. Sarlette. Symmetrization for quantum networks: a continuous-time approach. *to appear in MTNS 2014 proceedings*, 2014. 112

[64] F. Martinelli. Lectures on glauber dynamics for discrete spin models. In Pierre Bernard, editor, *Lectures on Probability Theory and Statistics*, volume 1717 of *Lecture Notes in Mathematics*, pages 93–191. Springer Berlin Heidelberg, 1999. 112

[65] M. Szegedy. Quantum speed-up of Markov chain based algorithms. *Proc. 45th Annual IEEE Symp. Foundations of Computer Science*, pages 32–41, 2004. 112

[66] C. P. Richter. Quantum speedup of classical mixing processes. *Physical Review A*, 76(4):042306, 2007. 112

[67] C. P. Richter. Almost uniform sampling via quantum walks. *New Journal of Physics*, 9(3):72, 2007. 112

[68] W. Fulton and J. Harris. *Representation Theory: A First Course*. Graduate Texts in Mathematics / Readings in Mathematics. Springer New York, 1991. 115, 116, 117