Habilitationà Diriger des Recherches (HDR)

Alain Sarlette

Habilitation à Diriger des Recherches (HDR)

Alain Sarlette*
submitted to Ecole Doctorale de Mathématiques 386
Université Pierre et Marie Curie, Paris, France

June 5, 2019

Abstract

This manuscript is meant to provide an overview of most of my research activities, taking a step back to highlight the links between my contributions and suggest perspectives for future research directions. A central element on which I have focussed turns out to be *controlled interactions*, with an approach at the boundary between applied mathematics, theoretical physics, and control theory. Other recurrent elements are *nonlinear manifolds*, and most importantly tools towards *developing quantum technology*. The overarching theme is the interplay between *information* and physical systems modeled by differential equations.

A first type of work establishes some *fundamental limitations*, as a help to guide further research focus. These results include
- fundamental bounds on rejecting disturbances in linear systems due to distributed sensing (with European Extremely Large Telescope primary mirror application);
- impossibility results on achieving so-called "string stability" in long chains of systems, with the first results in presence of strong nonlinearities (including any form of digital communication);
- a structural bound, showing that a first-order memory already achieves optimal convergence speed in consensus algorithms when the network spectrum can span a given interval (with an opening to accelerate the convergence for clustered spectra)
- comprehensive and tight results on the ultimately achievable mixing speed of Quantum Random Walks (QRWs), compared to Markov Chains and so-called Lifted Markov Chains: once memory is allowed, in principle QRWs provide no further speedup.

A second type of work has been developing *tools for the analysis of few-body*, controlled quantum systems, like those built in the experiments of my physics colleagues. Results along these lines include:
- results on Lyapunov functions for proving convergence of dissipative quantum systems towards a unique steady state, with an aim for general results and for covering infinite-dimensional systems;
- procedures for high-accuracy model reduction in open quantum systems, on the basis of series expansions, extending the usual adiabatic elimination method;
- characterizing the presence of low-dimensional deterministic manifolds on which some typical quantum stochastic differential equations remain confined with probability 1.

A third type of work has been the proposal of *particular designs or algorithmic procedures*.

*QUANTIC lab, INRIA Paris. 2 rue Simone Iff, 75012 Paris, France; alain.sarlette@inria.fr

- In the sense of generalizing simple linear procedures, we have proposed principles for applying integral control on nonlinear manifolds; and to formulate "consensus" as an abstract symmetrization procedure that can be extended among others to quantum systems;
- Towards stabilizing quantum systems, we have proposed a way to tune the repeated interaction with a stream of auxiliary systems in order to stabilize a target quantum system into a desired state, like so-called Schrödinger cat states of harmonic oscillators, Fock states or squeezed sates
- For the same goal, we have developed some improved quantum feedback schemes based on Quantum Non-Demolition measurements: stabilizing highly entangled states between remote systems despite decoherence along the line; and proving how injecting noise of controlled amplitude can achieve the so-far elusive *exponential* stabilization of eigenstates under weak continuous-time measurement.
- While working towards a better understanding of quantum algorithms, in particular QRWs, we have proposed a Quantum Fast-Forwarding algorithm (QFF) which provides an accelerated evolution of a classical Markov chain, without requiring to compute the associated lifted Markov chain as would be needed for designing a classical forwarding algorithm. This algorithm is shown to accelerate primitives on graphs provided via quantum query.

The main drive behind all this research has been to understand the essence and limitations of various procedures towards taking decisions in complex dynamical systems, on the basis of continuous physical variables. In this sense, the extraction of some most fundamental elements of quantum technology, combining both hardware level and logical level into a consistent and robust system-theoretic picture, remains a guiding objective for the years to come. This shall involve further intense collaboration with quantum physics experimentalists, and also with algorithm experts as I have started doing more recently. On the way, we plan to develop fundamental tools for approximate analysis and control design guidelines, tailored to the needs of interacting quantum systems.

Much of this work has been carried out in collaboration with outstanding and friendly scientific colleagues, whom I most warmly thank for this experience. Those whose work is missing here should please just consider that I had to leave out some things in the process of keeping the present document precise and concise.

Contents

1 Introduction

The interplay between information and physical systems has triggered interesting discussions for a long time, going back to concepts like entropy, Maxwell's demon and spontaneously emerging order in driven processes. On the technological side, it has become the enabler of communication technologies, information processing machines, certain approaches to feedback control, and algorithm developments.

While the Turing machine has formalized in a universal way how discrete information would be treated by machines, it leaves many doors open for exploration. The way in which continuous physical variables, governed by differential equations, could be exploited directly and robustly for information processing is rather implicit with this formalism. Design under constraints – e.g. locality of information in subsystems, unknown parts in an algorithm – also still has to be investigated. Most recently, the development of quantum computing and quantum error correction leads to a new take on how the interplay between abstract information processing and physical systems should be approached.

The language of mathematics is essential to link the physical world to what we call information. The concepts involved include differential equations modeling the behavior of continuous physical variables, and the related analysis tools; geometric concepts allowing to consider more abstract instances of "systems"; and the theory of probabilities and of stochastic processes, modeling contingencies, what is (un)known a priori and a posteriori, and how to relate them. My motivation for working in applied mathematics has been to develop insights for information-related aspects like:

- scalability, i.e. systems which work with reasonable complexity and accuracy when the number of components becomes large;

- making use of equivalences, symmetries and invariants to reduce system complexity or increase generality of some concepts;

- helping to tame complicated dynamical systems via focusing on partial insight only (e.g. Lyapunov function), significant system approximations, robustness bounds;

- examining and quantifying hard limits on how much we may control a physical system;

- applying this insight to analyze systems considered by physicists in experiments, and suggest new designs or procedures to obtain better system performance.

In terms of academic "checkboxes", my work would typically be categorized as Systems and Control Theory. In my humble opinion, this mainly reflects a set of tools and typical assumptions in my historical background. I like to challenge myself to expand this towards a more comprehensive treatment of information in, for and by physical systems. A better, comprehensive theoretical understanding should allow to design better algorithms and designs for challenging applications.

The following account summarizes most of my work by theme, not in chronological order. The length of each piece does not reflect its importance in terms of work effort, publications or expected impact; it is rather a mere consequence of what can be well described within the bounds of the present document. The reader is strongly encouraged to check the associated publications for any topic of interest. Their list can be found at https://scholar.google.be/citations?user=Skw_3L0AAAAJ&hl=en&oi=ao and freely accessible versions should be on arXiv. Readers less familiar with quantum dynamics will find a primer in appendix.

2 Fundamental limitations

The question of what is possible or what is fundamentally impossible has been recurrent in my research and has motivated the following developments. Their common point is the setting of distributed dynamical systems with local interactions, thus studying the emergent behavior of a network of subsystems interconnected in a tailored way. The common type of contribution is to clarify *which elements most essentially* lie behind the limitation in system performance.

2.1 A bound on disturbance rejection through distributed sensing

This works has been carried out in the framework of a collaboration between Liège University (my postdoc) and the European Southern Observatory for control of their EELT segmented primary mirror. The theoretical part that I have developed and review here has been published in [11].

The setting is a distributed plant where the controller can be centralized (i.e. it has access to measurements from the whole plant), but sensors only measure differences between neighboring subsystem outputs. Such "distributed sensing" can be a technological necessity in applications where system size exceeds accuracy requirements by many orders of magnitude. In Laplace-domain we consider the model:

$$y(s) = G(s) \, [\, u(s) + d(s) \,] \tag{1}$$
$$z(s) = [\, B + \Delta \,] \, y(s) + n(s) \tag{2}$$
$$u(s) = -C(s) \, z(s) \tag{3}$$

to represent $M \gg 1$ coupled N-dimensional subsystems. Components $kN+1$ to $(k+1)N$ of $y(s), u(s), d(s) \in \mathbb{C}^{N_y}$ denote outputs, inputs and disturbances of subsystem k in the Laplace domain, with $N_y = NM$. For easier reference we define $q_k = \{kN+1, kN+2, ..., (k+1)N\}$. We assume that the plant governed by $G(s)$ is stable. Output $z(s) \in \mathbb{C}^{N_z}$ is obtained through the static map $[B + \Delta] \in \mathbb{R}^{N_z \times N_y}$, where B is the nominal sensor behavior and Δ a sensor model error. Each sensor measurement is corrupted by zero-mean independent identical Gaussian white noise, represented by $n(s)$ with covariance matrix $\sigma^2 I_{N_z}$. The purpose of controller $C(s) \in \mathbb{C}^{N_y \times N_z}$ is to reject disturbances $d(s)$ from y. Importantly, we do not restrict the controller (3) to be distributed, i.e. we allow $C(s)$ to be a full matrix. We also allow the disturbances on different subsystems to be correlated, by investigating how a general vector $d(s)$ affects the controlled plant. This differs from studies on locally coupled systems or involving independent local disturbances, as discussed in the next section.The central element of our investigation is *local relative measurement*.

Assumption [distributed sensing]: Each component $l \in \{1, ..., N_z\}$ of $B \, y$ measures the difference between a convex combination of outputs of subsystem j and a convex combination of outputs of subsystem k, i.e. for each l there exist q_j, q_k such that:

1. $(B)_{l,m} = 0 \quad$ for $m \notin q_j \cup q_k$
2. $\sum_{m \in q_j} |(B)_{l,m}| = \sum_{m \in q_j} (B)_{l,m}$
    $= \sum_{m \in q_k} |(B)_{l,m}| = -\sum_{m \in q_k} (B)_{l,m} = 1 \,.$

Moreover, there exists a spatial range $\rho \geq 1$ and a spatial structure associating a position $p(k) \in \mathbb{R}^{\gamma}$ to each subsystem index k with $\|p(k) - p(j)\| \geq 1$ for $j \neq k$, such that the measurements performed by B all involve pairs (j, k) satisfying $\|p(k) - p(j)\| \leq \rho$.

Note that local sensing has no meaning if it is not relative: sensors giving "absolute" e.g. positions actually physically measure positions with respect to a common ("central") reference physically shared among all sensors.

Our main contribution is to clarify how **distributed sensing alone poses severe constraints on being robust to sensing inaccuracies** Δ and n, without assuming any controller restrictions.

Proposition 1: *Consider* (1)-(3) *with distributed sensing and* $\Delta = 0$. *The closed-loop relation between disturbances and outputs of interest can be summarized by the generalized sensitivity / complementary sensitivity relation:*

$$y(s) = [\, I_{N_y} + K(s)\,]^{-1}\, \delta(s) + [\, I_{N_y} + K(s)\,]^{-1} K(s)\, \nu(s) \tag{4}$$

where $K(s)$ summarizes all transfer functions; $\delta(s)$ are the open-loop deviations of $y(s)$ implied by $d(s)$; and $\nu(s)$ is measurement noise with amplitudes divided by the singular values $\sigma_k > 0$ of B.
Let D_m the maximal number of sensors connected to a given subsystem. Then for any (small) c, $N_c > 0$, there exists a (large enough) number of subsystems M such that at least N_c singular values σ_k are smaller than c.

<u>Proof idea:</u> The "bad modes" with low singular values of B correspond to long-range modes; their existence can be inferred as spanning situations where only a small fraction of the sensors indicate significant nonzero values. $\qquad\square$

This result thus implies that there can be many modes, for large M, on which effective measurement noise $1/\sigma_k$ is huge. This can lead to a difficult tradeoff between sensitivity and complementary sensitivity in particular for systems where input disturbances are strong on long-range modes. The EELT application, with d representing wind gusts, is such a case.

Proposition 2: *Consider* (1)-(3) *with $n = 0$ and relative uncertainty ε on each component of B, i.e.*

$$(\Delta)_{k,l} \in [\, -\varepsilon\, (B)_{k,l}\, ,\ +\varepsilon\, (B)_{k,l}\,] \quad \text{for all } k,l\ . \tag{5}$$

Then with singular value decomposition $B = U\Lambda^{1/2}Q^T$ and $\Phi = \Lambda^{-1/2}U^T\Delta Q$ we have

$$y(s) = [\, I_{N_y} + K(s)\, [I_{N_y} + \Phi]\,]^{-1}\, \delta(s)\ . \tag{6}$$

If $K(s)^{-1}$ can become arbitrarily small for some real $s > 0$, as e.g. using integral control, then for any $\epsilon > 0$ there exists Δ making the closed-loop system unstable.

<u>Proof idea:</u> This statement is a summarized version of the one in [11]. The main technical point is to construct an example of such Δ in the general case. $\qquad\square$

It is important to note that the model disturbances Δ making the system poorly robust are those *breaking the symmetry of perfect relative measurement*. This can happen as the stable subsystems are still connected to some reference (the mirror support in the EELT application) with respect to which tiny effects of absolute position could appear in the sensors (due to mirror curvature for the EELT application). Proposition 2 thus identifies that these effects can be very detrimental, and thus require to limit the control gain $K(s)$ (for the EELT application this has led to the use of an integrator with mode-dependent leakage). If

Δ satisfies the same invariance as B with respect to a global translation symmetry, then the robustness problem can be avoided. In other words, a tiny break of the symmetry in B can severely limit system performance.

The interested reader can find in [11] further details about the case of spatially invariant systems (spatial Fourier modes) and about the EELT segmented mirror application.

2.2 About the impossibility to achieve string stability

This work has been carried out during the PhD thesis of my student Arash Farnam at Ghent University, successfully defended in Fall 2018. It has been published in his thesis and in [1, 2] for the main part, with variations in [26]. It touches upon another fundamental limitation in locally coupled systems, somehow dual to the one above.

The concept of string stability is related to the behavior of a chain of N subsystems as N grows to infinity. It has been introduced for vehicle chains, with vehicles modeled as second-order integrators i.e. position controlled by force:

$$\frac{d^2}{dt^2} x_k(t) = u_k(t) + d_k(t) \quad , k = 1, 2, ..., N \; , \tag{7}$$

where x_k denotes the vehicle position, u_k its control action, and d_k are disturbances. By convention vehicle $k - 1$ would be in front of vehicle k. The aim is to stabilize a reference inter-vehicle distance $x_{k+1}(t) - x_k(t) = r$. Under disturbances d_k satisfying a certain norm bound, the controller should ensure that deviations of inter-vehicle distance from r remain bounded in a certain norm. The challenge is to achieve this: (i) for the worst disturbance satisfying the bound; (ii) relying on *local relative state measurements only*, like $x_{k+1} - x_k$ and its derivative but without any reference to absolute position or absolute velocity; and (iii) *for any chain length N*.

Note that compared to the previous section, there is no observability problem here since the variables of interest, i.e. the errors $x_{k+1}(t) - x_k(t) - r$, are directly observed. However, distributed sensing must now help to locally coordinate a swarm where *the individual subsystems are no longer stable*.

It has been noticed a few decades ago [59] that string stability can be problematic to achieve. The simplest setting considers the linear system in frequency domain:

$$u_k(s) = K(s) \left(x_{k-1}(s) - x_k(s) + r \right) \tag{8}$$

where $K(s)$ is a transfer function specifying how each vehicle k reacts to deviations of its position from its predecessor's position. Defining $e_k = x_{k-1} - x_k - r$, the closed-loop system writes

$$e_k(s) = \frac{K(s)}{s^2 + K(s)} e_{k-1}(s) + \frac{1}{s^2 + K(s)} (d_k - d_{k-1}) \; .$$

Recognizing that $\frac{K(s)}{s^2+K(s)}$ takes the form of a sensitivity function with second-order open loop, we know by Bode's integral that it will necessarily have a norm $g_\omega > 1$ at some frequency $s = j\omega$. If thus for instance $d_1 = \cos(\omega t)$ and $d_{k>1} = 0$, then the perturbation d_1 will be amplified $(g_\omega)^k$ as it arrives at agent k, and for chain length N tending to infinity the effect on relative position e_N grows unbounded — definitely not an acceptable situation regarding

swarm cohesion and collision avoidance. The important insight is that this is unavoidable for any linear controller of the form (8).

Note that if absolute-velocity drag is added, i.e. $\frac{d^2}{dt^2}x_k(t) = u_k(t) - \gamma\frac{d}{dt}x_k(t) + d_k(t)$, then the situation can be resolved. However, situations where absolute velocity has no or too little influence can appear: think of space vehicles, of the operational benefits of avoiding to rely on a global absolute velocity reference for all systems in such situations, of preferably minimizing energy-loss and thus natural drag. More fundamentally, the behavior should ideally be stable in natural processes with Galilean invariance.

In the meantime, several variants of string stability have been examined, with various choices of norms on d_k and e_k and various interconnection strategies: looking several vehicles in front and behind, possibly communicating. Some authors consider perturbations in initial conditions rather than as an input signal; see e.g. [57] for a review. A particular modification has been to make the desired distance r dependent on absolute velocity, in the so-called *time-headway policy* $r(t) = r_0 + h\frac{d}{dt}x(t)$, $h > 0$ a real parameter, see e.g. [32]. The rationale is that at higher speeds, on would typically wish a larger intervehicle distance. The drawback is that large h would imply poorly compact vehicle platoons at higher speeds; and, e.g. in space vehicles or referring to Galilean invariance, that absolute velocity would have to be given an operational meaning, i.e. defined by measurement with respect to a global reference. The advantage is that this type of velocity-dependence does not involve drag forces, and still allows to solve the issue: for well-tuned $a, b, h \in \mathbb{R}$, the PD controller

$$u_k(s) = (a + bs)\left(x_{k-1}(s) - x_k(s) - hsx_k(s)\right)$$

does allow to achieve string stability according to various definitions (see more precision below).

Our contributions on this problem revolve around two axes. First, among the various settings for string stability, several natural options had not been characterized and we have clarified this. Second and maybe most importantly, we have taken string stability out of the realm of linear systems, by allowing nonlinear controllers and any type of nonlinear communication in a discrete-time setting. The result proving impossibility to satisfy string stability even in this setting, with a completely new proof unrelated to linear systems studies, suggests that something deeper about considering worst-case bounded disturbances and marginally unstable open-loop subsystems must be at play, than suggested by the linear study. This could provide important fundamental insight for future studies of quite general distributed systems.

2.2.1 About linear unidirectional controllers, see [1]

Considering the possibility to use time-headway, we define the configuration error

$$e_k(s) = x_{k-1}(s) - x_k(s) - hsx_k(s)$$

for some $h \geq 0$, keeping $h = 0$ for the case without relying on absolute velocity. We focus on the effect of disturbance inputs, with the usual 2-norm over time. While the practical criterion would most likely correspond to BIBO and thus ∞-norm over time, the 2-norm is a too useful proxy with its reformulation in frequency domain and it has been considered by most authors.

For a vector v of signals $v_k(t)$, $i = 1, 2, ..., N$, the L_2 norm of v_k is defined as $\|v_k(.)\|_2 = \sqrt{\int_{-\infty}^{+\infty} (v_k(t))^2 \, dt}$ and the (L_2, l_2) norm of v is defined as $\|v(.)\|_2 = \sqrt{\sum_{k=1}^{N} \int_{-\infty}^{+\infty} (v_k(t))^2 \, dt}$, which by Parseval both have equivalent expressions in frequency domain. When evaluating the vector norm at a specific value of t in time domain or of s in frequency domain we write e.g. $\|v(s)\|_2 = \sqrt{\sum_{k=1}^{N} |v_k(s)|^2}$.

Definition 1 [(L_2, l_2) **string stability**]: *The chain* (7), *controlled with feedback signals* u_k *to be designed, is called* (L_2, l_2) *string stable if there exists a constant value* c_1 *such that*

$$\|e(.)\|_2 \leq c_1 \, \|d(.)\|_2$$

for all bounded signals d, *and all chain lengths* N.

Definition 2 [L_2 **string stability**]: *The chain* (7), *controlled with feedback signals* u_k *to be designed, is called* L_2 *string stable if there exists a constant value* c_1 *such that*

$$\|e_j(.)\|_2 \leq c \, \|d_k(.)\|_2 \quad \text{for all } k ,$$

for any situation with bounded disturbance signal d_k *at some* k *and zero disturbances on all* $j \neq k$, *and for all* N.

The difficulty, for both definitions, is to make the bound uniform in N. Definition 2, considered in e.g. [28, 32], essentially requires that the transfer function from any d_k to any e_j is H-infinity bounded independently of N. This appears to be a desirable necessary property for practical applications, but not strictly sufficient. Indeed, L_2 string stability does not bound independently of N what happens when disturbances are present at the inputs of *all* the vehicles: if each individual disturbance d_k can induce a nonzero bounded error on each e_j, without further constraints, then when disturbances act on all the vehicles the sum of their effects could make e_j grow unbounded with N. Avoiding this is the motivation behind the stronger Definition 1, which is in line with the definitions considered by [16] among others.

We have clarified two main points in this context.

- Results with time headway $h > 0$ had focused on satisfying the weaker Definition 2. We have clarified that Definition 1 can be satisfied too, provided the controller has unbounded DC gain (thus unlike a PD controller).

- Results with inter-vehicle communication always included time-headway as well. We wanted to know if communication alone, with $h = 0$, can enable to achieve string stability. We have proved a negative result in this sense.

We have also proved a few side results that I will not explain in detail here.

The most general controller that we consider writes:

$$
\begin{aligned}
u_k(s) &= K(s)e_k(s) + H(s)r_k(s) \\
r_k(s) &= W(s)\,q_{k-1}(s) \\
q_k(s) &= F(s)e_k(s) + G(s)r_k(s) \\
e_k &= x_{k-1}(s)x_k(s) - hsx_k(s) ,
\end{aligned}
\tag{9}
$$

for $k = 2, 3, ..., N$. Here q_k is a signal sent by vehicle k to vehicle $k + 1$, and the latter receives it in the form of r_{k+1} after transition through the communication channel. Thus $W(s)$ is a linear model of the communication channel, while $H(s)$, $F(s)$ and $G(s)$ are a combination of encoding/decoding/gain filters applied to the communication signal. We further assume the following elements.

Assumption [controller & communication channel]:
- $K(0) \neq 0$, i.e. impossibility to have a perfect derivative of the measurement output in order to cancel the pole at $s = 0$ in the open-loop system (**??**).

- Unlike in [33], we do not allow $K(s)$ to grow unboundedly with N and we assume all transfer functions independent of N.

- The poles of F, G, H, W must all have negative real parts.

- $G(j\omega)$, $F(j\omega)$ and $\frac{H(j\omega)}{K(j\omega)}$ are bounded for all ω.

- System behavior must be robust to replacing W by any $\tilde{W}$ close to W, without adapting the controller F, G, H, K.

Several of these assumptions, beyond typical practice, can be motivated by the fact that the string stability formulation only explicitly models input perturbations: modeling inaccuracies, measurement noise, communication noise or bias are not included. The usual care about robust control design applies to avoid that these unmodeled inaccuracies would be amplified by the controller. Another typical limitation would be a finite bandwidth in W associated with additive communication noise, but we do not even need this assumption for our results.

Remark: Examples of string stability involving poorly robust designs abound. Canceling the open-loop pole at $s = 0$ would yield a relative-degree-one transfer function in closed loop, whose Bode diagram can have a magnitude ≤ 1 at all frequencies. Another example would assume that communication can be made perfect by choosing $G(s) = 1/W(s)$. Then communicating $r_1 = x_0 - x_1$, $r_k = r_{k-1} + e_k$ for $k > 1$, vehicle k would in fact receive $r_k = x_0 - x_k$. In other words, vehicle k would have access to the perfect value of $x_0 - x_k$, as if the coupling with the leader was global. On this basis each vehicle could just stabilize its position with respect to the leader with some stable controller $u_k(s) = H(s)r_i(s) = H(s)(x_0 - x_k)$, independently of all the other vehicles in the chain. However, assuming perfect transmission of arbitrary signals throughout the whole vehicle chain is not what a realistic communication channel can do, and implementing the associated strategy has failed in any realistic simulations.

The results of [1] can be summarized by the following table. For better context we mention in italics three existing results. CACC-type communication refers to Cooperative Adaptive Cruise Control, a standard communication setting considered in papers like [39, 43] where q_k takes the particular form of a filtered version of the input command u_k:

$$q_k(s) = \tfrac{1}{B(s)}\big(K(s)e_k(s) + H(s)r_k(s)\big) \ .$$

The theorem numbers refer to [1], where the proof can be found. They involve no particularities beyond a fitting formulation of the linear dynamics, allowing to derive bounds from low-dimensional systems analysis.

Standard impossibility; [59]	$h = W = 0$	Def.1 and Def.2 fail
Time-headway solution; [32]	$W = 0$, $h > 0$ with PD controller (a fortiori also for any $W \neq 0$)	Def.2 holds
Theorem 1	$W = 0$, $h > 0$ with $K(0)$ bounded	Def.1 fails
Theorem 2	$W = 0$, $h > 0$ with PID controller (a fortiori also for any $W \neq 0$)	Def.1 and Def.2 hold
CACC with time-headway; [39, 43]	$W \neq 0$, $h > 0$	Def.2 holds with possibly better scaling, lower h
Theorem 3a	CACC-type $W \neq 0$, $h > 0$, $K(0)$ bounded	Def.1 fails
Theorem 3b	CACC-type $W \neq 0$, $h = 0$	Def.1 and Def.2 fail
Theorem 4a	any $W \neq 0$, $h = 0$, $K(0)$ bounded	Def.1 fails
Theorem 4b	any $W \neq 0$, scalar q_k, $h = 0$	Def.1 and Def.2 fail

From these results, a remaining possibility under the linear control model (9) could involve some complicated MIMO communication from k to $k + 1$; however, there is no concrete hint how this would indeed solve the string instability issue. These doubts get further support from the following result, establishing a very general impossibility in an equivalent discrete-time model when $h = 0$.

2.2.2 A general string instability result, see [2]

A major restriction in the above, and in most if not all previous approaches, is the linear model. Especially when communication is involved, one can doubt whether this is the proper setting. More fundamentally, one may wonder how allowing nonlinearity would affect the string stability issue.

Maybe surprisingly, we were able to give a quite general result in this sense, provided we assume a discrete-time controller, which should be compatible with nowadays' digital control systems.

We thus assume that each subsystem is controlled with a discrete-time control logic, computing at each time $t = n \, \Delta t$ with $n \in \mathbb{Z}$, a control signal that will be applied during the whole interval $(t, t + \Delta t]$ as input to each double integrator; thus Δt is the time increment of the discrete-time controller. We can then integrate exactly the second-order dynamics over the time interval $(t, t + \Delta t]$ to obtain:

$$
\begin{aligned}
v_k(t + \Delta t) &= v_k(t) + u_{k,1}(t) + d_{k,1}(t) \\
x_k(t + \Delta t) &= x_k(t) + v_k(t) \, \Delta t + u_{k,2}(t) + d_{k,2}(t) \, ,
\end{aligned}
\tag{10}
$$

where $x_k, v_k \in \mathbb{R}$ for $k = 0, 1, 2, ..., N$ denote position and velocity respectively; the control inputs $u_{k,1}, u_{k,2}$ result from integrating the control signal applied during the interval $(t, t + \Delta t]$ respectively once and twice; and the disturbances $d_{k,1}, d_{k,2}$ result from similarly integrating the continuous-time input disturbances.

Regarding string stability definitions, integrating the signals would yield the following equivalent definitions; thanks to our general nonlinear proof, we can cover more variants.

Definition 3: *For positive integers p, q, the (ℓ_p, ℓ_q) string stability requires that there exist $C_1, C_2 > 0$ independent of N such that: for any disturbances satisfying*

$$\sum_{s=1,2} \sum_{k=0}^{N} \left(\sum_{n \in \mathbb{Z}} |d_{k,s}(n\,\Delta t) \tfrac{1}{\Delta t^s}|^p \Delta t \right)^{q/p} < C_1 \,,$$

it is ensured that $\sum_{k=0}^{N} \left(\sum_{n \in \mathbb{Z}} |e_k(n\,\Delta t)|^p \Delta t \right)^{q/p} < C_2$.

Definition 4: *For positive integer p, the (ℓ_p, ℓ_∞) string stability requires that there exist $C_1, C_2 > 0$ independent of N such that: for any disturbances satisfying*

$$\sum_{n \in \mathbb{Z}} |d_{k,s}(n\,\Delta t) \tfrac{1}{\Delta t^s}|^p \Delta t < C_1 \,, \quad s = 1, 2 \ \text{ for all } k \,,$$

it is ensured that $\sum_{n \in \mathbb{Z}} |e_j(n\,\Delta t)|^p \Delta t < C_2$ for all j.

Definition 5: *The $(\ell_\infty, \ell_\infty)$ string stability requires that there exist $C_1, C_2 > 0$ independent of N such that: for any disturbances satisfying $|d_{k,s}(t)|/\Delta t^s < C_1$ for $s = 1, 2$ and all k, t, it is ensured that $|e_j(t)| < C_2$ for all j, t.*

In terms of the ordering of the constants (like "for each C_1 there exists a C_2"), we have chosen the weakest constraint, such that the impossibility result will be the strongest. With this setting we allow the following very general control law, where $\dot{e}_k = v_{k-1} - v_k$ and $e_{k:\ell}$ denotes for $\ell > k$ the set of values $e_k, e_{k+1}, ..., e_\ell$:

$$
\begin{aligned}
u_{k,1} &= f_1\big(e_{(k-m_1):(k+m_2)}, \dot{e}_{(k-m_1):(k+m_2)}, \\
&\qquad\qquad c_{k,+}, c_{k,-}, \xi_k, N, t\big) \,, \\
u_{k,2} &= f_2\big(e_{(k-m_1):(k+m_2)}, \dot{e}_{(k-m_1):(k+m_2)}, \\
&\qquad\qquad c_{k,+}, c_{k,-}, \xi_k, N, t\big) \,, \\
c_{k+1,+} &= g_1\big(e_{(k-m_1):(k+m_2)}, \dot{e}_{(k-m_1):(k+m_2)}, \\
&\qquad\qquad c_{k,+}, c_{k,-}, \xi_k, N, t\big) \,, \\
c_{k-1,-} &= g_2\big(e_{(k-m_1):(k+m_2)}, \dot{e}_{(k-m_1):(k+m_2)}, \\
&\qquad\qquad c_{k,+}, c_{k,-}, \xi_k, N, t\big) \,, \\
\xi_k(t+\Delta t) &= h\big(\xi_k(t), e_{(k-m_1):(k+m_2)}, \\
&\qquad\qquad \dot{e}_{(k-m_1):(k+m_2)}, c_{k,+}, c_{k,-}, N, t\big) \,.
\end{aligned}
\tag{11}
$$

Here m_1 (resp. m_2) is a finite number of agents ahead (resp. behind); $c_{k,+}, c_{k,-} \in \mathbb{R}^{n_c}$ with n_c some bounded integer are communication signals from $\{k - m_1, ..., k - 1\}$ to k and from $\{k+1, ..., k+m_2\}$ to k respectively; the $\xi_k \in \mathbb{R}^{n_\xi}$ for some finite integer n_ξ allow for dynamical controllers with finite memory; and f_1, f_2, g_1, g_2, h are arbitrary functions, with minimal regularity just to ensure that the solution to the dynamical system is well-defined at all times. In particular, by specifying a particular profile of $u(\tau)$ over the discretization time span $\tau \in [t, t + \Delta t)$ between two measurement updates, one can command $f_1 = \int_t^{t+\Delta t} u(\tau) d\tau$ and $f_2 = \int \int_t^{t+\Delta t} u(\tau) d\tau^2$ independently. The controllers (11) are applied by all vehicles $k \in (m_1, N - m_2)$, whereas adapted versions are applied by the m_1 leading and m_2 last vehicles. The adapted versions play no role in the proof. The functions f_1, f_2, g_1, g_2, h can be nonlinear and time-dependent (e.g. modulated at specific frequencies), communication can involve quantization and coding, thereby vastly extending the traditional LTI setting. Our only true restrictions on control design are:

Assumptions [nonlinear string]:

- *The controllers must be homogeneous along the chain,* i.e. the functions f_1, f_2, g_1, g_2, h do not depend on vehicle index k and the internal variables are initialized with the same default values for each k.

- *The digital controller has a finite update time Δt,* fixed independently of chain length N.

- *The control commands are based on* relative *state information only*, i.e. there is no dependence on absolute states that would be obtained with respect to some common reference, like absolute velocity.

Theorem: *For the system* (10),(11), *there exists no controller design satisfying string stability. More precisely, there exist disturbances $d_{k,1}$ and $d_{k,2}$ satisfying the required respective bounds according to the definitions and such that, irrespectively of any (well-defined) controller choice, for large N:*

[Definition 3]: $\sum_{k=0}^{N} \left(\sum_n |e_k(n\,\Delta t)|^p \right)^{q/p}$ *grows as* $N^q \Delta t^{2q}$;

[Definition 4]: $\sum_n |e_j(n\,dt)|^p$ *grows as* $N^p \Delta t^{2p}$;

[Definition 5]: $|e_j(t)|$ *grows as* $N\Delta t^2$.

<u>Proof idea:</u> The point is to identify a particularly bad disturbance. Consider

$$
\begin{array}{ll}
d_{k,1}(t) = d_{k,2}(t) = 0 & \text{for all } t < 0, \quad k = 0, 1, ..., N; \\[2mm]
\left. \begin{array}{l} d_{k,1}(t) = \frac{\alpha k \Delta t}{N}, \\[1mm] d_{k,2}(t) = \frac{\alpha k \Delta t^2}{N} \end{array} \right\} & \begin{array}{l} \text{for} \qquad\quad \text{all} \\ t \;\; = \;\; 0, \Delta t, ..., T, \\ k = 0, 1, ..., N \; ; \end{array} \\[4mm]
d_{k,1}(t) = d_{k,2}(t) = 0 & \text{for all } t > T, \quad k = 0, 1, ..., N \; ,
\end{array}
\tag{12}
$$

with constants $\alpha > 0$ and $T > 0$ specified according to the needs of the proof. The evolution of the system under this disturbance and with zero initial conditions can be easily written in closed form, *for the vehicles in the center of the swarm*: the key is that the corrections $u_{k,1}, u_{k,2}$ applied by vehicle k are exactly the same as those applied by vehicle $k-1$, so they bring nothing on e_{k-1} and the perturbation keeps building up an increasing value e_{k-1}, until information from the boundary of the chain reaches vehicle $k-1$ or k. From there, the conclusion follows by a counting argument until this happens and fixing the worst α, T. $\square$

The assumption of homogeneity appears non-essential, in the sense that inhomogeneities may probably be cancelled with adversarial noise reshaping, but this remains to be checked. The assumption of fixed Δt appears compatible with a finite reaction speed, be it in terms of digital computation power or in physical bandwidth of the components.

2.2.3 Discussion and outlook

Our results can motivate several conclusions. According to Section 2.2.1, time headway can be put to good use, but with care regarding other perturbation sources (see $K(0)$ unbounded). Also, it appears that communication does not really help on its own, it is a matter of quantitative performance only. In related work, we also highlight that care must be taken about which norms one uses as a proxy for the practical constraint.

From Section 2.2.2, one should wonder what keeps systems with Galilean invariance and nonlinear interactions, e.g. physical materials, string stable in practice. The result indeed extends to swarms in higher dimension, and to interaction strengths which vanish asymptotically with increasing index-distance among the subsystems.

In the end, one may wonder whether requiring string stability may be too much. At any given finite N, there exist controllers limiting the growth of e_k – is this happening also in huge swarms of e.g. physical materials? after all our proof examines the system on the time of propagation throughout the chain, i.e. size of the sample divided by speed of sound. Another point is that the worst-case disturbances may become less and less probable with increasing N, such that huge swarms would in fact be string stable "in probability", see e.g. [15]. Anyways, it seems worth clarifying the best engineering setting for controller design in large swarms.

2.3 First-order memory optimally accelerates consensus

This work has been published in [47], following up on a conjecture by Prof. Sandro Zampieri. We here consider a very simple task and focus, instead of on perturbation rejection, rather on the optimal speed of convergence under local coupling.

The basic consensus problem goes as follows. The variables of interest are $x_k \in \mathbb{R}$, with arbitrary initial values $x_k(0)$, $k = 1, 2, ..., N$. The goal is to make all the x_k converge towards the same end value $x_1 = x_2 = ...x_N$, sometimes with the restriction that this end value must be equal to the average of initial values. In practice x_k can represent e.g. a value estimate, a working load, or a slack variable of local agent k. The dynamics is very simple, e.g. first-order integrators

$$\frac{d}{dt}x_k(t) = u_k(t) \quad \text{or} \quad x_k(t+1) = x_k(t) + u_k(t) \tag{13}$$

with u_k a control input that we have have to design for each agent k. There are two main restrictions on designing u_k. The first one is that agents interact according to some imposed interaction graph $G(t)$, whose nodes are the agent indices and where $u_k(t)$ cannot depend on $x_j(t)$ if $G(t)$ contains no edge from j to k. The second restriction is that driving each agent individually to e.g. $x_k = 0$ should not be an option: the u_k should depend on *relative states* $x_j - x_k$ only. Mathematically this corresponds to invariance of the dynamics under the common translation $x_k(t) \mapsto x_k(t) + a$ for all k and for any fixed $a \in \mathbb{R}$.

Various questions can be asked about consensus, e.g.: How does the formulation generalize to other spaces than $\mathbb{R}^n$ (subject of my PhD thesis)? What are the minimal guarantees we must make on the time-varying interconnection graph $G(t)$ to ensure convergence to consensus? How does it generalize to more complicated dynamics (be careful about invariance to keep it meaningful w.r.t. relative states)? What is the fastest possible convergence towards consensus?

The result of [47] concerns the convergence speed, for a fixed graph $G(t) = G$, in the discrete-time setting with linear controllers $u_k(t)$. We denote $x, u \in \mathbb{R}^N$ the vectors of components x_k, u_k respectively. The most simple controller would be *static* and write

$$u(t) = K x(t)$$

for some acceptable gain matrix K, satisfying graph locality and ensuring stability. Convergence would then be exponential, governed by the second-largest eigenvalue in norm of $(I + K)$ where I is the identity matrix. (The largest eigenvalue equals one and expresses the

11

symmetry under common translation $x_k(t) \mapsto x_k(t) + a$.) At the opposite end, if knowing and analyzing the whole graph, one can design various time-dependent exchanges

$$u(t) = K(t)\, x(t)$$

which ensure *exact* convergence to consensus in no more time than the diameter of G [30]. Another easy way to see this finite-time convergence, at least for symmetric graphs, is to write L any symmetric positive semidefinite interaction matrix; compute $\mu_1 > \mu_2 \geq ... \geq \mu_N = 0$ its eigenvalues, where $\mu_N = 0$ is again the unavoidable consequence of common translation symmetry; and apply $K(t) = -\frac{1}{\mu_t}L$ for $t = 1, 2, ..., N-1$, such that at each time we completely cancel the eigenspace related to a given eigenvalue of L. (Usually for L we take the Laplacian of the graph.)

In view of such a primitive application as achieving consensus, one may doubt however whether this last approach, involving a precise spectral analysis of the whole network matrix L, is acceptable for design. Some adaptive approaches have been proposed, estimating the interconnection matrix L online [58]. An intermediate complexity between these two extremes involves a *dynamic controller with simple memory*:

$$x(t+1) = x(t) - \alpha L x(t) + \sum_{m=1}^{M} \beta_m (x(t-m) - x(t)) \,. \tag{14}$$

Here L again represents a given interconnection matrix, while α and $\beta_1, ... \beta_M$ are gains $\in \mathbb{R}$ to be freely tuned. Thus each agent k now has M local memory slots, storing previous values of x_k. Such algorithms have been proposed e.g. in [53]. A fundamental fact, well-known and exploited in more general settings too, is that the role of the dynamical memory is typically to *extrapolate* the movement: in order to achieve a speedup, the additional memory e.g. for $M = 1$ should push $x_k(t)$ out of the interval spanned by $x_k(t-1)$ and $x_k(t-2)$. A basic question is to know how the convergence speed can be improved by adding more memory slots, i.e. increasing M. It was in fact not even clear a priori that increasing M keeps improving the speed. Numerical exploration led to the following conjecture, which I prove in [47].

Theorem [47]: *Consider dynamics* (14) *with L symmetric and having its nonzero eigenvalues in the interval $[\underline{\lambda}, \bar{\lambda}] \subset (0, +\infty)$. Denote by ν the slowest eigenvalue of the converging part of the dynamics, i.e. excluding the norm-1 eigenvalues related to translation symmetry or to L modeling a disconnected graph. Then $\nu \in (0, 1)$ is minimized over all $\alpha, \beta_1, \beta_2, ... \beta_M \in \mathbb{R}$, when using just a single memory slot optimally tuned as in [53a], i.e. taking:*

$$
\begin{aligned}
\beta_m &= 0 \ \text{for all } m > 1 \,, \\
\beta_1 &= \beta_{1*} := -\left(1/\mu - \sqrt{1/\mu^2 - 1} \right)^2 , \\
\alpha &= \alpha_* := 2(1 - \beta_{1*}) / (\bar{\lambda} + \underline{\lambda}) \,,
\end{aligned}
$$

where $\mu = \frac{\bar{\lambda} - \underline{\lambda}}{\bar{\lambda} + \underline{\lambda}}$. The corresponding convergence speed guarantee is

$$\nu = 1/\mu - \sqrt{1/\mu^2 - 1} = \sqrt{-\beta_{1*}} \,.$$

<u>Proof idea:</u> We write the system in the eigenbasis of L to decouple it into N independent systems, parametrized by the eigenvalues. From there we must check the worst of a continuum of systems, indexed by eigenvalue $\lambda \in [\underline{\lambda}, \bar{\lambda}]$. To this aim we write the z-transform of each subsystem, we normalize it directly by the optimal solution, and we study it in the complex plane in the search for the renormalized polynomial taking a root of modulus ≥ 1. The study of the location of roots in the complex plane, as a function of λ, goes through a sequence of propositions where we ultimately use the Cauchy Argument Principle. We show that unavoidably, any deviations from the optimum solution leads to the presence of a renormalized root of norm ≥ 1, implying an eigenvalue for the convergence of (14) of norm $\geq \nu$. $\qquad\square$

The bound ν increases with μ, which is the convergence speed for $M = 0$ i.e. no added memory. For $\mu = 1\text{-}\varepsilon$, with ε the *spectral gap*, we have $\nu < 1\text{-}\sqrt{\varepsilon}$ and the bound gets tight as $\varepsilon \to 0$, showing a quadratic speedup reminiscent of many other quadratic speedups, including in quantum systems. All this ultimately seems connected with a diffusive vs. transport behavior (see further in this document), introduced thanks to the "extrapolating" step in taking β_1 negative. This transport behavior also makes the whole state matrix asymmetric, logically preventing to reapply the technique on itself. Adding a single memory slot thus appears to introduce just enough transport vs. diffusion, for what we know about L.

In [6], we have generalized this setting by studying how the convergence can be further accelerated if we know that the spectrum of L is confined to a set of narrow, disjoint intervals, as can appear for particular types of graphs. We propose a polynomial filter, which is a variant of the eigenvalue-cancellation strategy of [30], and we suggest how for a given graph the eigenvalues of L may be explicitly "clustered" into a few narrow intervals.

The general point was thus to clarify: If we only want to rely on *limited knowledge about* L — here an upper bound and a lower bound on its spectrum — then also the algorithm which achieves fast convergence in all cases can be simpler — here involving a single memory slot per agent. This appears like a concrete instance of a possibly more general approach to controller design: justifying simplicity of optimal solutions when robustness is part of the constraints.

2.4 Mixing with Quantum Random Walks compared to classical processes

This work constitutes the first part of the PhD thesis of my student Simon Apers at Ghent University, successfully defended in January 2019. It has been carried out in collaboration with Francesco Ticozzi from the University of Padova. It has been published in Simon Apers' thesis and in [9, 8, 10].

The square root improvement in convergence speed observed in the previous section is reminiscent of some quantum speedups. In particular, the so-called Quantum Random Walk (QRW) has been advertised as allowing up to a square root improvement in mixing time, compared to classical random walks. However, this QRW uses more memory than a classical walker, and we were hence prompted to compare it more directly to so-called Lifted Markov Chains (LMCs) which essentially come down to allowing a finite-state memory in classical stochastic processes as well. We have clarified that:

- The finite-time evolution of any QRW over a finite number of steps can be obtained

as well with a properly tuned LMC; thus there is no inherent quantum advantage in leveraging the constraint of graph locality.

- The mixing time of invariant QRWs can be simulated by an LMC of finite memory, implying a conductance bound on QRW mixing. The same holds true replacing QRWs by any stochastic process that satisfies a formal locality constraint.

- For the most prominent graph instances where mast mixing QRWs have been designed with simple rules, namely lattices, there exists an equally simple LMC design with fast mixing.

- The mixing time of LMCs depends on the framework in which they are designed: does the target distribution have to be invariant, do we have control on the LMC initialization,... ?

We must note that these conclusions apply to *tight lower bounds* on the mixing time, i.e. the fastest mixing time achievable by any design, without discussing how this fastest design may actually be found. In case of constraints on the computations made to optimize their design, both LMCs and QRWs are likely to be slower and differences between the two settings may appear. We come back to this in Section **??**.

Accordingly, the main message of this work would be: *with respect to QRWs and graph locality, there is no advantage when focusing on convergence speed as such; instead, claiming any quantum advantage must include a clear discussion of design constraints.* In other words, from a less algorithmic perspective: speedups in mixing and transport phenomena are not diagnostic of quantum effects, although superdiffusive spreading may be more prominent with quantum walks. This calls for putting into perspective previous work on potential quantum advantages of QRWs for fast mixing; it also provides a tight answer to the formerly open question, to which point QRWs must obey the conductance bound on mixing time.

2.4.1 Quantum Random Walks and their classical simulation (see [9])

Standard discrete-time QRWs [38, 3, 60] describe the evolution of a quantum particle ("walker") over a discrete set of graph nodes $\mathcal{V}$, conditioned on additional degrees of freedom $\mathcal{C}$ (the walker's coin or spin). The walker state is thus defined on the Hilbert space $\mathcal{H} = \mathcal{H}_C \otimes \mathcal{H}_V = \mathrm{span}\{|c\rangle \otimes |v\rangle | (c, v) \in \mathcal{C} \times \mathcal{V}\}$, while the object of interest is the induced distribution p_t over $\mathcal{V}$. The cycle graph is a simple example where QWs provide a mixing speedup, see Fig. 1. To the nodes $\mathcal{V} = \{1, 2, \ldots, N\}$ of the cycle, the QW adds a binary coin $\mathcal{C} = \{+, -\}$ [4, 3]. Denoting $\mathbf{P}^{\pm}$ the cyclic permutation of position, that is, $\mathbf{P}^{\pm}|v\rangle = |(v \pm 1) \mathrm{mod} N\rangle$ for $v \in \mathcal{V}$, the unitary QRW primitive reads

$$\mathbf{U} = \mathbf{S}\left(\mathbf{C} \otimes \mathbf{I}_N\right), \quad \mathbf{C} = \begin{bmatrix} e^{-i\phi}\sqrt{1-\alpha} & e^{i\theta}\sqrt{\alpha} \\ -e^{-i\theta}\sqrt{\alpha} & e^{i\phi}\sqrt{1-\alpha} \end{bmatrix},$$

where $\mathbf{S} = |+\rangle \langle +| \otimes \mathbf{P}^+ + |-\rangle \langle -| \otimes \mathbf{P}^-$ expresses conditional shifts, while $\mathbf{C}$ is a general unitary "coin toss" on $\mathcal{H}_C$. The conditional motion can also be viewed as spin-orbit coupling. The dynamics $U^t |\psi_0\rangle$ induces a distribution p_t that spreads over $\mathcal{V}$ in $O(N)$ steps and oscillates quasi-periodically. To perfectly *mix*, i.e. make p_t converge and remain close to the uniform distribution over $\mathcal{V}$, it suffices to intersperse some measurement or slight decoherence, see [31]
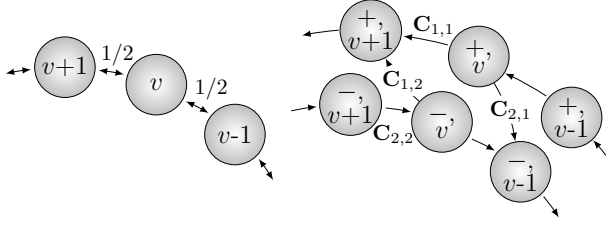
14

Figure 1: (left) random walk $\mathbf{P}_0$ on the cycle; (right) quantum walk unitary with coin toss $\mathbf{C}$, or lifted Markov chain with a stochastic coin toss $\overline{\mathbf{C}}$, suggesting their comparison.

for a survey. For instance, one can perform with probability $q(t)$ a projective measurement in the canonical basis between consecutive applications of $\mathbf{U}$:

$$|\psi_{t+1}\rangle = \begin{cases} \mathbf{U}\,|\psi_t\rangle & \text{with probability } 1 - q(t), \\ |c, v\rangle & \text{with probability } q(t)\,|\,\langle c, v|\mathbf{U}|\psi_t||\rangle^2\,. \end{cases} \tag{15}$$

Taking $q(t) = 1\ \forall t$ projects the state at each step and comes down to classical dynamics. The QW of Eq. (15), even with constant $q = O(1/N)$ and e.g. $\alpha = 1/2, \phi = \theta = 0$, converges towards a uniform p_t in $t = O(N)$ steps, from any initial distribution [4, 3]. In contrast, a classical random walk with transition matrix $\mathbf{P}_0 = (\mathbf{P}^+ + \mathbf{P}^-)/2$ reaches the same distribution only after $O(N^2)$ steps.

Compared to classical random walks, the QRW above adds memory via the coin. Yet, QWs can exhibit memory effects even without coin. Consider the two-node graph $\mathcal{H} = \text{span}\{|1\rangle, |2R\rangle\}$, equivalent to a qubit, and take the Hadamard gate $U_H = (\sigma_x + \sigma_z)/\sqrt{2}$ as QRW primitive. Starting on a given node, after one step, the distribution is uniform over $|1\rangle, |2\rangle$, yet at the second step the initial state is perfectly recovered since $(U_H)^2$ equals identity. This behavior, impossible for any classical memoryless process, comes from the quantum state storing information in its relative phases, or coherences. Hence, to establish an intrinsic quantum advantage, QRWs should be compared to classical local processes with at least some amount of additional memory.

Remarkably, a classical walker with memory that mixes fast on the cycle has already been proposed independently of the QRW literature [23, 24], and it shares striking similarities. This walker moves among *classical* states in $\mathcal{C} \times \mathcal{V}$. Its probability distribution $\hat{p}$ over $\mathcal{C} \times \mathcal{V}$ evolves as $\hat{p}_{t+1} = \mathbf{P}\,\hat{p}_t$, with stochastic transition matrix $\mathbf{P}$ having the same structure as $\mathbf{U}$, yet with $\mathbf{C}$ replaced by a *stochastic* coin toss:

$$\mathbf{P} = \mathbf{S}\left(\bar{\mathbf{C}} \otimes \mathbf{I}_N\right), \quad \bar{\mathbf{C}} = \begin{bmatrix} 1 - \alpha & \alpha \\ \alpha & 1 - \alpha \end{bmatrix}. \tag{16}$$

This can be seen as the mixture of two reversible evolutions: with probability $1 - \alpha$, the state follows the conditional shift $\mathbf{S}$; or, with probability α, the coin is switched before applying $\mathbf{S}$. The coin allows the classical walker to use information about its previous motion direction, in physical terms its momentum. In fact $\mathbf{P}$ in Eq. (16) exactly describes the probabilistic evolution induced by Eq. (15) with $q(t) = 1\ \forall t$, and starting with $|\psi\rangle = |c, v\rangle$ for some $(c, v) \in \mathcal{C} \times \mathcal{V}$. Such $\mathbf{P}$ mixes over the cycle in $O(N)$ steps [23, 24], provided $\alpha = O(1/N)$. This speedup, only due to classical memory, matches the one provided by the QRW. In both cases, an $O(1/N)$ nonunitarity provides a good tradeoff between fast (unitary) motion and

15

losing correlation with the initial condition. From these observations, it appears most natural to compare QRWs like Eq.(15) to classical evolutions with memory like Eq.(16), which are formalized as LMCs [21].

More formally, consider a graph $\mathcal{G}$ with node set $\mathcal{V}$ and edges $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$. Nodes could represent energy levels and edges allowed transitions. The QRW and LMC constructions both start by building a lifted graph, where each node of $\mathcal{G}$ is split into "lifted nodes" or "sublevels". This is done without loss of generality by introducing a coin set $\mathcal{C}$, defining the lifted nodes $\mathcal{C} \times \mathcal{V} = \{(c, v)\}$ and selecting lifted edges in $\{(\,(c, v), (c', v')\,) \mid (v, v') \in \mathcal{E}\}$, thus without introducing transitions that were not allowed before lifting.

General QRWs, including for instance open quantum walks [12], are then described by a quantum channel over the space generated by viewing coin and node as quantum numbers, i.e.,

$$\rho_{t+1} = \sum_k \mathbf{M_k(t)} \rho_t \mathbf{M_k(t)}^\dagger, \tag{17}$$

where ρ_t is a density operator on $\mathcal{H} = \mathrm{span}\{|c, v\rangle \mid (c, v) \in \mathcal{C} \times \mathcal{V}\}$ and the $\mathbf{M_k}$ satisfy $\sum_k \mathbf{M_k}(t)^\dagger \mathbf{M_k}(t) = \mathbf{I}_{\mathcal{C} \times \mathcal{V}}$ for all t, with $\mathbf{I}$ the identity [1]. Graph locality imposes $\langle c', v' | \mathbf{M_k} | c, v =\rangle 0$ if $(v, v') \notin \mathcal{E}$. An initial distribution p_0 over $\mathcal{V}$ is mapped onto the lifted nodes (or sublevels) by $F : p_0 \mapsto \rho = \sum_{v \in \mathcal{V}} p_0(v) |c_v, v\rangle \langle c_v, v|$, thus associating some fixed initial coin state c_v to each v. The object of interest is the distribution p_t over $\mathcal{V}$, the main nodes or levels, obtained with the partial trace as $p_t = \mathrm{diag}(\mathrm{trace}_{\mathcal{C}}(\rho_t))$.

Similarly, a LMC follows $\hat{p}_{t+1} = \mathbf{P} \hat{p}_t$, where $\hat{p}_t$ is a vector representing the probability distribution over $\mathcal{C} \times \mathcal{V}$, and $\mathbf{P}$ is a stochastic matrix expressing the jump probabilities among sublevels. Namely, denoting by $p = e_v$ and $\hat{p} = e_{(c,v)}$ the distributions with probability 1 of being on v and on (c, v), respectively, $e_{(c',v')}^\dagger \mathbf{P} e_{(c,v)}$ is the transition probability from (c, v) to (c', v'). Graph locality imposes $e_{(c',v')}^\dagger \mathbf{P} e_{(c,v)} = 0$ if $(v, v') \notin \mathcal{E}$. Initial lifted nodes are assigned by $F : p_0 \mapsto \hat{p}_0 = \sum_v p_0(v) e_{(c_v, v)}$. The distribution of interest is obtained by marginalizing over $\mathcal{C}$, thus $p_t(v) = \sum_{c \in \mathcal{C}} \hat{p}_t(c, v)$ for all $v \in \mathcal{V}$. A LMC is in fact a particular QRW where populations evolve without coherences. The key to our results is to observe how, conversely, any QRW can be simulated by some LMC (with possibly higher-dimensional coin). In other words, the non-Markovian evolution of p_t under a QRW can be described as a classical Markovian evolution of sublevel populations.

Theorem 1: *Given a QRW and a time horizon τ_0, there exists an LMC with local initialization and satisfying the same graph constraints such that, for all $t \in [0, \tau_0]$ and all initial distribution p_0 over the nodes, the distributions p_t induced by the QRW and by the LMC coincide.*

<u>Proof idea:</u> First, we show that for any initialization of the walk on a fixed starting node v, we can build a sequence of transition matrices $\mathbf{P}_t$ on $\mathcal{V}$ alone, whose distribution exactly follows the one of the QRW; the technical argument is the same as for Theorem 4 below. An time-independent LMC is obtained by adjoining a time index to the state of the walker, initialized at 0, and conditional on which the transitions among nodes are applied. Finally, a single LMC is obtained by similarly adjoining an "initial node" index to the state of the walker, combining the $\mathcal{M}_v$ such that they are applied conditional on this "initial node" index.

---

[1]Some authors add so-called Cesaro averaging on top of this model [3, 35]. We can explicitly capture this and similar extensions via local stochastic maps, see Supplemental Material.

By linearity, any initial distribution over graph nodes is then correctly following the QRW. □

This existence result has clear consequences on the best achievable mixing time of QRWs compared to LMCs. A QRW or LMC is said to mix to some distribution $\bar{p}$ over $\mathcal{V}$ if p_t asymptotically converges to $\bar{p}$, that is, if for any $\epsilon > 0$ there exists a finite $\tau(\epsilon)$ such that $\frac{1}{2}\sum_{v\in\mathcal{V}}|p_t(v) - \bar{p}(v)| \leq \epsilon$ for all $t \geq \tau(\epsilon)$ and all p_0; $\tau(\epsilon)$ is called the *mixing time* for getting ϵ-close to $\bar{p}$. A standard "stabilizing" requirement for a process that converges to $\bar{p}$ is that $p_0 = \bar{p}$ should imply $p_t = \bar{p}$ at all times. This holds automatically for the time-invariant $\mathbf{P}$ in the LMC framework. In the QRW framework, the $\mathbf{M_k(t)}$ depend on time typically only through the measurement mechanism. Such QRWs too preserve $p_t = \bar{p}$ at all times when $p_0 = \bar{p}$. We call this $\bar{p}$-invariance.

Theorem 2: *Given a $\bar{p}$-invariant QRW with mixing time $\bar{\tau}(\epsilon_0)$ for some $\epsilon_0 \leq 1/4$, we can construct an LMC that has mixing time, for all $\epsilon > 0$:*

$$\frac{\tau(\epsilon)}{\bar{\tau}(\epsilon_0)} \leq \left\lceil \frac{\log 1/\epsilon}{\log 1/(2\epsilon_0)} \right\rceil .$$

<u>Proof idea:</u> We use the construction of the LMC, with conditional application of particular transitions as summarized in the proof of Theorem 1, with $\tau_0 = \bar{\tau}(\epsilon_0)$. When reaching the upper limit of the time index, the walker is rerouted towards the lowest time index in such a way that it undergoes the same procedure repeatedly. Invariance of $\bar{p}$ ensures that booked progress is kept and each round exponentially decreases deviations from $\bar{p}$. □

The importance of providing a fully time-independent and invariant LMC construction to simulate the QRW stems from the fact that for such LMCs, the best achievable mixing time can be bounded by a purely graph-dependent quantity called the graph conductance $\Phi_{\bar{p}}$. Partitioning $\mathcal{V}$ into two subsets $\mathcal{X}$ and $\mathcal{X}^c$, let all the stationary population on $\mathcal{X}^c$ be lost; the conductance counts which fraction of the remaining population $\bar{p}(\mathcal{X}) = \sum_{v\in\mathcal{X}} \bar{p}(v)$ jumps back to $\mathcal{X}^c$ in one step, i.e.:

$$\Phi_{\bar{p}}(\mathbf{P}) = \min_{\mathcal{X}:0<\bar{p}(\mathcal{X})\leq\frac{1}{2}} \left(\sum_{v\in\mathcal{X},v'\in\mathcal{X}^c} (e_{v'}^\dagger \mathbf{P}e_v)\bar{p}(v) \right) / \bar{p}(\mathcal{X}) .$$

The maximal $\Phi_{\bar{p}}(\mathbf{P})$ over all Markov chains that keep $\bar{p}$ invariant on a given graph, is the graph conductance $\Phi_{\bar{p}}$.

Theorem 3: *Any $\bar{p}$-invariant QW has a mixing time $\tau(1/4) \geq 1/(4\Phi_{\bar{p}})$, and there exists such a QW with mixing time $\tau(\epsilon) \leq O(\log(1/\min_k \bar{p}_k)\log(1/\epsilon) / \Phi_{\bar{p}})$ for all $\epsilon > 0$.*

<u>Proof idea:</u> Given Theorem 2, the proposition follows from the fact that these properties are known for LMCs, see e.g. [21]. □

This last result answers an open question about proving a conductance bound for QRWs tighter than [3, 63], and thus as tight as for classical LMCs. While the general constructions in the proofs of Theorem 1 and Theorem 2 lead to LMCs containing many lifted nodes, and based on much analysis of the system evolution, it is worth noting that the situations where "simple" QRW speedups have been obtained are all but general. These situations mostly reduce to lattice graphs, and in this case a much simpler LMC design, similar to (16), can achieve the same mixing speed, as we explain in [9].

2.4.2 Extension to general stochastic processes (see [8])

In our LMC constructions for the results of the previous section, from the viewpoint of the distribution p_t, the initial QRW is only used as a way to determine how much population is transferred between nodes. The QRW serves as a way to make these decisions in a more complicated way than a simple random walk, and LMCs can do so as well. In fact, we can show that LMCs can simulate any abstract way of taking these decisions.

We thus consider an abstract stochastic process Ψ_t, mapping p_t to p_{t+1} for $t = 0, 1, 2, \dots$. To express locality on this abstract process, we just impose: for any p_t and any node $v \in \mathcal{V}$, the population according to $\Psi_t(p_t)$ of v is upper bounded by the sum of the populations, according to p_t, of v and of its graph-neighbors.

Theorem 4: *The results of Theorems 1 to 3 also hold when replacing the QRW by an abstract stochastic process satisfying the graph locality requirement.*

<u>Proof idea:</u> The proof follows exactly the same lines as for the QRW. The idea of the first step, i.e. proving that for a given initial node v we can build a sequence of transition matrices $\mathbf{P}_t$ inducing the same distribution over $\mathcal{V}$, relies on translating the locality constraint into an instance of the max-flow min-cut theorem. $\qquad\square$

2.4.3 LMC mixing time depends on the application scenario (see [10])

In Theorems 2 and 3, we have restricted our attention to $\bar{p}$-invariant QRWs without much discussion. In LMC studies like [21] this is not the standard, but the focus lies on convergence of the distribution over the *lifted* nodes. Moreover, gather-and-distribute techniques like those mentioned in the previous section about consensus, under full graph knowledge, allow to achieve perfect convergence in diameter-time. This points to the fact that the precise context of the mixing task should be clarified.

In [10] we give several examples of existing algorithmic procedures that can be classified under the following characteristics.

- *Initialization of the lift:* When a stochastic dynamics is seen as an algorithm, one must specify how to initialize it. We denote (S) the situation where the algorithm designer is allowed to tune the initial values of the latent variables, thus the distribution over lifted nodes associated to each original node; (s) is the situation where the relevant initial conditions for the LMC comprise *all* distributions over the lifted nodes. We denote $\mathcal{S}$ the relevant set of initial conditions

- *Invariance of the target marginal:* The evolution of an LMC depends not only on the distribution over $\mathcal{V}$ but also on the latent variables modeled by lifted nodes. It may happen that for an initial condition in $\mathcal{S}$ for which the distribution over $\mathcal{V}$ matches the target $\bar{p}$, the auxiliary variables drive the system away from $\bar{p}$ before coming back; we denote this situation (I). If having an initial distribution matching $\bar{p}$ implies that $\bar{p}$ is preserved for all times, then the LMC is called invariant and we denote this (i).

- *Marginal vs lift mixing time:* A situation (M) focuses on the convergence of the marginal distribution over $\mathcal{V}$, while a situation (m) requires convergence of the full distribution over lifted nodes; the mixing time is thus defined accordingly.

- *Reducibility of the lift:* One may allow reducible lifts (R) or not (r).
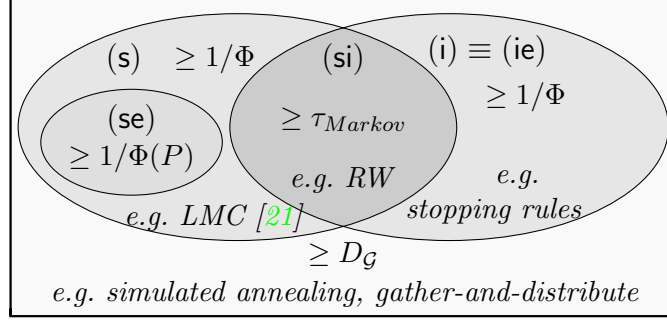
Figure 2: Summary of the main bounds on mixing time $\tau(1/4)$, mentioning only the most relevant constraints and up to small constant factors. $D_{\mathcal{G}}$ denotes the diameter of the graph and Φ the conductance, either $\Phi(P)$ associated to fixed P or Φ allowing to optimize over all P with limit $\bar{p}$. By τ_{Markov} we mean the mixing time of the non-lifted Markov chain $p_{t+1} = P\,p_t$, where in case (E) the transition matrix P can be optimized and in case (e) it is given. The names mention algorithmic procedures that fit each setting. **All bounds are**
- **essentially tight**, in the sense that for all scenarios and all graphs there exists an LMC attaining the lower bound up to a log factor;
- **essentially significant**, in the sense that there exist graphs where the various bounds differ by polynomial factors.

- *Matching ergodic flows:* The literature on optimizing Markov chains often requests to match ergodic flows, which are the expected flows along each graph edge in steady state. Imposing this constraint is denoted (e), discarding it (E).

An LMC and its mixing time can be considered in any combination of these options. Of course there may be others of which we were not thinking, besides our deliberate choice to not involve with which computational complexity the LMC design is achieved. Each time the lowercase letter indicates the most constraining option.

Figure 2 summarizes how those characteristics influence the best achievable mixing time. The proofs are just technical and we refer the reader to [10] for more details. We also provide there examples of actual algorithms and graphs distinguishing these cases. In essence, the focus on marginal or lifted distribution implies no significant difference, nor using reducible LMCs or not. The most determining constraints are the request of invariance, and the possibility to initialize the auxiliary variables with particular values or not. These constraints are verifiable on an algorithm directly. Requesting the strongest on both these constraints, even the non-lifted Markov chain cannot be improved; requesting the weakest on both, a diameter-time convergence is a priori possible; in between, the conductance bound is significant.

We emphasize again that these bounds concern the ultimately achievable mixing time in a given framework, without discussing the complexity of designing the algorithm that would attain this optimal convergence. We nevertheless believe that this could give useful guidelines for algorithm design in terms of dead-ends and opportunities for modifying existing settings.

3 Tools for analyzing open quantum systems

The challenge for quantum technology is to build devices which work with sufficient accuracy despite their unavoidable interaction with an environment. As a means of controlling quantum systems whose open-loop behavior is not perfectly fixed, measurements are carried out. When an environment or a measurement device is integrated into the dynamical model, the quantum system no longer follows a Schrödinger equation and we say that it is an open quantum system. The open quantum system is in general not in a pure maximal information state, its state is hence described by a density operator ρ on its Hilbert space. This ρ can be seen as the probabilistic mixture of various "pure" states (rank-one density operators $\rho = |\psi\rangle\langle\psi|$) in which the system could be, conditioned on the state of its environment.

The average action of an environment on ρ is described by a deterministic Lindblad equation, modeling "decoherence", which generalizes to quantum models the dissipative (as opposed to Hamiltonian) classical processes. On the control side, so-called reservoir engineering techniques can be envisioned to turn such decoherence into information-protecting mechanisms. Closer to standard control means, measurement-based feedback can apply Hamiltonian actions onto the system, depending on measurement results. In this case the dynamical model becomes stochastic.

Aside from proposing control designs and architectures, the precise analysis of the resulting dynamical systems can be challenging because of time-dependence, stochastic effects, and large dimensions (ρ is a $2^n \times 2^n$ complex matrix when describing n qubits). This section gathers results which I have developed towards analyzing quantum systems that are relevant in practical experiments. The last two subsections are part of an ongoing effort in the framework of the ANR-JCJC grant "High-Accuracy Model Reduction for Open Quantum Systems".

3.1 Lyapunov methods

A Lyapunov function is a scalar positive function of the state that decreases along trajectories, and it thereby allows to conclude that the system must asymptotically converge to certain sets of interest, e.g. equilibria. Analyzing the derivative of a scalar function makes system analysis easy; the challenge is to *find* the good Lyapunov function for a given system. For Linear-Time-Invariant systems, quadratic Lyapunov functions are the standard. Quantum systems are affine in ρ, but high-dimensional, and with the additional property of being positivity- and trace-preserving. Therefore other systematic approaches are of interest, see also e.g. [17].

We here list two of our contributions to Lyapunov functions for open quantum systems. The first one derives from contraction properties of Petz metrics. The second one is concerned with the infinite-dimensional character of typical quantum systems like the harmonic oscillator.

3.1.1 Contraction-based Lyapunov functions [46]

This contribution is based on metrics quadratic in ρ, but exploiting the particular structure of the Lindblad equation

$$\tfrac{d}{dt}\rho = -i[H,\rho] + \sum_k L_k\rho L_k^\dagger - \tfrac{1}{2}L_k^\dagger L_k\rho - \tfrac{1}{2}L_k^\dagger L_k\rho \tag{18}$$

where H is the system Hamiltonian, the L_k are operators on the Hilbert space describing the "decoherence channels" acting on the open quantum system, and ρ is the density operator

describing the system state: it is Hermitian, positive semidefinite, and of trace 1 since the diagonal of ρ in a given basis describes the probability distribution for projective measurements in that basis.

The remarkable work [44] characterizes the Riemannian metrics that are contractive under *all* (finite-dimensional) Kraus maps, which are the discrete-time equivalent of (18). We show that the same metrics are contractive under *all* evolutions of the form (18). They can be parameterized by the formula

$$\|\delta\rho\|_\rho^2 = \int_0^1 \text{trace}(\,\delta\rho(\delta\omega_s + \delta\omega_s^\dagger)/2\,)\, m(s)\, ds \tag{19}$$

where $\delta\omega_s$ is solution of the Sylvester equation

$$s\,\delta\omega_s\rho + \rho\delta\omega_s = \delta\rho \tag{20}$$

and $m(s)\,ds$ is a positive finite measure. These include standard distance measures, like the Bures distance for $s = 1$:

$$\frac{d}{dt}\text{trace}(\delta\rho\delta\omega) \;=\; -2\sum_k \text{trace}([\delta\omega, L_k]\,\rho\,[\delta\omega, L_k]^\dagger) \le 0\,, \quad \text{with } \rho\delta\omega + \delta\omega\rho = \delta\rho\,.$$

This result can be used to systematically build a Lyapunov function for *any* (finite-dimensional) Lindblad differential equation, for which an equilibrium ρ_∞ of full rank is known (and nothing more). For this we take ρ_∞ as basis for the tangent vector $(\rho - \rho_\infty)$ in the contraction results.

Corollary: *If the equation* (18) *has a full-rank equilibrium* ρ_∞, *then whatever the form of H and of the L_k, the function*

$$V_{\text{Bures}}(\rho) = \text{trace}(\rho_\infty\, G_\rho^2) \quad \text{with } \rho_\infty G_\rho + G_\rho\rho_\infty = \rho - \rho_\infty$$

is a (non-strict) Lyapunov function for (??) *and* $\frac{d}{dt}V_{\text{Bures}}(\rho) = -\sum_k \text{trace}(\,[G_\rho, L_k]\,\rho_\infty\,[G_\rho, L_k]^\dagger\,)$.

In order to conclude about global convergence of the state to ρ_∞, it remains to examine how the commutator of the set of all L_k relates to ρ_∞, if necessary with a LaSalle-type argument. This can sometimes be very efficient and systematic, as we have shown on a few example systems.

3.1.2 Well-posedness and convergence of an infinite-dimensional system [14]

This contribution is based on metrics linear in ρ; its purpose is to show how to rigorously address the infinite dimensional Hilbert space of typical quantum systems.

The harmonic oscillator is a standard quantum system of infinite dimension [29], whose Hilbert space $\mathcal{H}$ can be seen as the span of orthogonal states $\{|n\rangle : n \in 0, 1, ...\}$ modeling eigenstates of exactly n energy quanta. It models in particular an electromagnetic oscillation mode in circuitQED, which has been the subject of intense study at INRIA / QUANTIC.

In particular, towards stabilizing "Schrödinger cat" states in such circuits (we will come back to this in Section 4), we consider harmonic oscillators with resonant drive and loss processes involving k photons at once, $k \ge 1$:

$$\frac{d}{dt}\rho = L\rho L^\dagger - \tfrac{1}{2}L^\dagger L\rho - \tfrac{1}{2}L^\dagger L\rho =: \mathcal{L}(\rho) \quad \text{with} \quad L = \mathbf{a}^k - \alpha^k I\,,$$

where $\mathbf{a}$ is the photon loss operator $\mathbf{a}|n\rangle = \sqrt{n-1}|n-1\rangle$, I is the identity operator and α is a scalar constant (real positive without loss of generality). Note that the operator L is unbounded. In [14] we establish the well-posedness and convergence properties of this system.

The well-posedness is based on defining the Banach space $\mathfrak{B} = \{\rho : \mathrm{Trace}(S_\nu \rho S_\nu) < \infty\}$, where $S_\nu = (I + L^\dagger L)^\nu$ is well-defined spectrally. We show that the problem is well-posed in this Banach space for any $\nu \geq 1/2$, and the norm is a Lyapunov function. The proof is partially inspired by [22]. It is based on the Hille-Yosida theorem: showing that the Lindblad superoperator is accretive i.e. its domain $D(\mathcal{L})$ is dense in $\mathfrak{B}$ and for every $\lambda > 0$, $I + \lambda\mathcal{L}$ is a bijection from $D(\mathcal{L})$ into $\mathfrak{B}$ with $(I + \lambda\mathcal{L})^{-1}$ a bounded linear operator on $\mathfrak{B}$ with norm less or equal to 1. One of the key and original arguments, unique to infinite-dimensional systems, is that the commutator $[L, L^\dagger]$ defines a non-negative Hermitian (unbounded) operator.

Using $\nu \geq 3/2$, we further obtain that $V(\rho) = \mathrm{Trace}(L\rho L^\dagger)$ decays exponentially, and for any $\nu' < \nu$ each initial condition converges to a unique equilibrium. For $k > 1$, the nullspace of V is a manifold in which quantum information can be encoded. We conclude by providing invariants of the dynamics to tell to which state the system converges, depending on ρ_0.

3.2 High-accuracy adiabatic elimination

This work has been initiated during the PhD thesis of Rémi Azouit, in co-supervision with Pierre Rouchon, and pursued with my postdoc Paolo Forni.

Composite quantum systems are often high-dimensional, as the dimensions of subsytems' Hilbert spaces multiply. Yet often only few dimensions are relevant at a given timescale: some interactions are very weak, some components stabilize quickly to their steady state, residual errors induce slow deviations of quantum information. Adiabatic elimination uses this timescale separation to eliminate the fast contracting dimensions, and summarize their effect on the remaining degrees of freedom. In quantum information applications, ideally no contraction should take place on the manifold of nominal information states, and its weak coupling or leakage to other states can then be studied as an effective perturbation. Complementarily, strong contraction on some physical degrees of freedom can protect encoded information from errors that this contraction counters — this can be viewed similarly as the bistable hardware to stabilize the two states encoding a classical bit. It is thus tempting to study these effects using approximate analysis via timescale separation.

We have developed explicit formulas for carrying out this approximate analysis on the Lindblad equation (18), which we summarize as

$$\tfrac{d}{dt}\rho = \mathcal{L}\rho = (\mathcal{L}_0 + \epsilon\mathcal{L}_1)\rho \ .$$

The timescale separation is modeled with $\epsilon \ll 1$, where $\mathcal{L}_0$ has a particularly ideal structure — e.g. many eigenvalues with zero real part — and $\mathcal{L}_1$ is a perturbation to this ideal situation. For Hamiltonian operators, computing the resulting perturbations on eigenvectors and eigenvalues is standard. For Lindbladian superoperators, the linear systems perspective would be to compute an approximate block-diagonalization, separating the subspaces corresponding to large or to small real parts of the eigenvalues and describing the remaining dynamics inside each block. However, (18) features an additional structure of positivity-preservation and a physical interpretation of this dynamics is possible only through the operators H and L_k. In order to find a reduced system with the same structure, thus guaranteeing a direct physical interpretation, we rather take an approach of center manifold theory.

We therefore write $\rho = \mathcal{K}(\rho_s)$ and $\frac{d}{dt}\rho_s = \mathcal{L}_s\rho_s$, where ρ_s is the density operator associated to the reduced system; the reduced dynamics $\mathcal{L}_s$ should take the same form as (18); and the mapping $\mathcal{K}$ from reduced model to full model should take the form of a Completely Positive Trace-Preserving map (CPTP, or Kraus map). Expanding both $\mathcal{K}$ and $\mathcal{L}_s$ as a series in ϵ and solving the invariance equation

$$\mathcal{K}(\mathcal{L}_s\rho_s) = \mathcal{L}(\mathcal{K}(\rho_s))$$

for each power of ϵ separately, we can obtain formulas approximating the reduced dynamics at various orders. The computational bottleneck becomes the inversion of $\mathcal{L}_0$ on particular operators; we have developed a few tricks for the latter, and positivity guarantees. A detailed account of our results would bring us too far here; in summary:

- Physicists routinely use adiabatic elimination for a few paradigmatic systems, typically composed of a dispersive or resonant Hamiltonian interaction between a high-quality (non-dissipating) system and a low-quality (fast dissipating) auxiliary. At first order, they compute associated Zeno dynamics; at second-order, the fast dissipation in $\mathcal{L}_0$ induces spurious dissipation on the ideal system at the typically known rate g^2/κ when $\mathcal{L}_0$ is of order κ and $\epsilon\mathcal{L}_1$ is of order g. Our goal was to go beyond this standard knowledge.

- For the first-order and second-order approximation in ϵ, we obtain explicit and general formulas. Firstly, for composite systems, we can address arbitrary fast dissipation and arbitrary Hamiltonian couplings, possibly composed of many terms. This allows to explore other models of e.g. quantum electrodynamical circuits than the few canonical ones typically considered in less systematic approaches, comparing induced decoherence channels in a direct and quantitative way. Secondly, we have developed similar formulas for other structures than multi-partite systems under Hamiltonian coupling: cascaded-unidirectional couplings, as appear in quantum systems linked with transmission lines and circulators, and decoherence-free subspaces as are used for quantum information encoding.

- We have generalized the adiabatic elimination to cases where several small systems perturb a target one, as their effects just add up (to second-order, not necessarily beyond), and when a fast Hamiltonian acts on the ideal system. This allows to add drives performing useful operations on the nominal system, and to study the effects induced by fast dissipating material imperfections when such drives are applied.

- We have obtained the formulas for third-order and fourth-order effects in various cases. This allows for instance to study imperfections in reservoir engineering designs, where the second-order dissipation is in fact the target effect ("protecting information"). The third-order has been developed for the general case of weak Hamiltonian coupling between a fast dissipating subsystem and a slowly dissipating one. Our method is the very first to provide any results at this order.

- We conjecture that it should be possible to carry out the developments to arbitrary order. From a linear systems perspective, block-diagonalizing is possible in an analytic way. The issue is to keep the positivity-preserving quantum structure. A gauge degree of freedom appears in the invariance equation, which we should exploit in this direction at the cost of modeling a somewhat reduced subset of the state space.

- On the basis of the general analytic formulas, we can set up optimization of parameters in experimental systems. For instance, it may be possible to add drive to an environment, or to add further dissipation channels than the unavoidable ones, if this would lead to a better effect on the target information. We have shown that often, somewhat counterintuitively, it is best to make all mixing channels as strong as possible on the environment: despite adding more entropy to the overall system by a factor κ, this decreases the effective coupling between ideal system and environment by a factor κ^2, such that overall the target information is less affected. This looks like a fully Markovian version of Quantum Dynamical Decoupling [64].

The related formulas, theorems and examples can be found in the set of papers [13]. A review paper is in preparation.

3.3   Deterministic submanifolds of the Quantum Master SDE

This work has been carried out in collaboration with Pierre Rouchon at INRIA/QUANTIC. It has originated from experimental observations in the group of Benjamin Huard, see [18]. A first theoretical basis has been published in [50] and an extension is being submitted soon.

The aim of this work is also to reduce the dimension of the dynamics to be studied, but in a different context. We consider quantum systems subject to weak continuous measurement, and whose evolution thus depends on the (stochastic) measurement results. We then study how, in typical system configurations, the stochastic process in fact only impacts a small part of the dynamics. In other words, most of the evolution can be precomputed before knowing the measurement results, as at each fixed time the measurement results only govern a distribution on a low-dimensional manifold. Unlike in the previous section, this manifold is typically nonlinear. A direct application is to efficiently implement quantum filters / observers of low dimension.

We consider a quantum system under weak continuous-time measurement:

$$
\begin{aligned}
d\rho_t &= -i[H,\rho]\,dt + \sum_k (L_k\rho L_k^\dagger - \tfrac{1}{2}L_k^\dagger L_k\rho - \tfrac{1}{2}\rho L_k^\dagger L_k)\,dt \\
&\quad + \sum_k \sqrt{\eta_k}(L_k\rho + \rho L_k^\dagger - \operatorname{trace}(L_k\rho + \rho L_k^\dagger)\rho)\,dW_k \ , \\
dY_{t,k} &= \sqrt{\eta_k}\operatorname{trace}(L_k\rho + \rho L_k^\dagger)\,dt + dW_k \ .
\end{aligned}
\tag{21}
$$

Here ρ is the density operator on Hilbert space $\mathcal{H}$ and $d\rho_t$ its Bayesian update according to evolution under Hamiltonian H and measurement results $dY_{t,k}$. The latter involve a measurement efficiency $\eta_k \in [0,1]$ and an operator L_k on $\mathcal{H}$ characterizing the measurement channel. The stochastic nature of quantum measurements is captured by the Wiener processes W_k, appearing both in the measurement and in the related update equation. The latter also contains a deterministic decay, i.e. the usual Lindblad decoherence, expressing that measuring a quantum system in one basis necessarily induces uncertainty in other bases, whether the measurement is read or not. Interactions with an unread environment then correspond to $\eta_k = 0$. The stochastic part must be understood in the Ito sense. Note that the decomposition into a sum of terms is not unique, i.e. several apparently different combinations of $\{L_k, H\}$ can in fact model the same dynamics.

A typical example of this setting has been experimentally studied in [18] in collaboration with Benjamin Huard: the target system is a qubit $\mathcal{H} \simeq \mathbb{C}^2$, drive $H = 0$ and the measurement channel at a rate γ is spontaneous photon emission with phase, corresponding to $L_1 = \sqrt{\gamma}|g\rangle e$ and $L_2 = i\sqrt{\gamma}|g\rangle e$; measurement efficiency η was of the order 0.5. It was observed that when repeating the experiment many times and integrating (26), the quantum state ρ appeared to be distributed on a spherical cap inside the Bloch sphere, converging towards the south pole $\rho = |g\rangle\langle g|$, and this independently of initial condition. We have provided a theoretical confirmation and general theory behind this observation.

The distribution that can be induced via (26) as a function of the processes $\{W_k\}$, suggests a reachability study. And Indeed, the Stroock-Varadhan theorem [56] establishes that the support reached by a stochastic equation in the Stratonovich sense, can be studied by replacing the noises dW_k by control inputs u_t; thus having arbitrary noise is replaced by having the freedom to choose arbitrary control input signals. From there, our analysis of deterministic submanifolds follows a clear strategy:

1. Reformulate (26) in the Stratonovich sense by the standard correction (algebraic manipulations). Viewing the resulting dynamics as a sum of vector fields, one "drift" vector field multiplied by dt and the others "control vector fields" multiplied by $dW_k = u_k dt$ for each k, the dimensions among which the system can spread come down to studying the *strong accessibility* of the dynamics under free control signals u_k.

2. Algebraic conditions exist to give the dimension of the strongly accessible space of a system: compute $\mathfrak{G}_F$ the smallest algebra generated by commutations among control vector fields, and invariant under commutation with the drift vector field. The dimension of $\mathfrak{G}_F$ is the dimension of the manifold, possibly evolving deterministically in time, over which the dW_k can make the system diffuse (algebraic manipulations).
   To check this dimension, for quantum systems, particular care must be taken to evaluate linear independence of a set of vector fields generated by superoperators.

3. If the dimension of $\mathfrak{G}_F$ is lower than $\dim(\mathcal{H})^2 - 1$, which is the number of degrees of freedom in ρ, then the system remains confined with probability 1 to a lower-dimensional, possibly time-dependent manifold. The associated deterministic variables should be sought by obtaining a set of variables following autonomous ODEs, i.e. where the dW_k and influences of other variables disappear. To inform this search, one can take into account the tangent to the manifold, spanned by $\mathfrak{G}_F$.

4. Ideally the separation of variables into deterministic and stochastic ones can be given a closed-form, towards implementing reduced-order filters which must only keep track of the stochastic part.

Setting out this strategy was a first contribution of [50]. We have then further applied it to the system of [18] and computed the manifolds, see Figure 3. Furthermore, as the distribution reduced to 1-dimension, we were able to provide exact analytic expressions for the whole probability distribution over the Bloch sphere at any time; those specific expressions involve several coordinate changes to solve a nonlinear PDE, for details we refer the reader to [50].

More general theoretical results are the following ones.

Theorem 1: *Consider the quantum stochastic differential equation* (26) *for a qubit system* $\mathcal{H} \simeq \mathbb{C}^2$, *with a generic value of the η_k and with $H = 0$. Then there only exist two types*
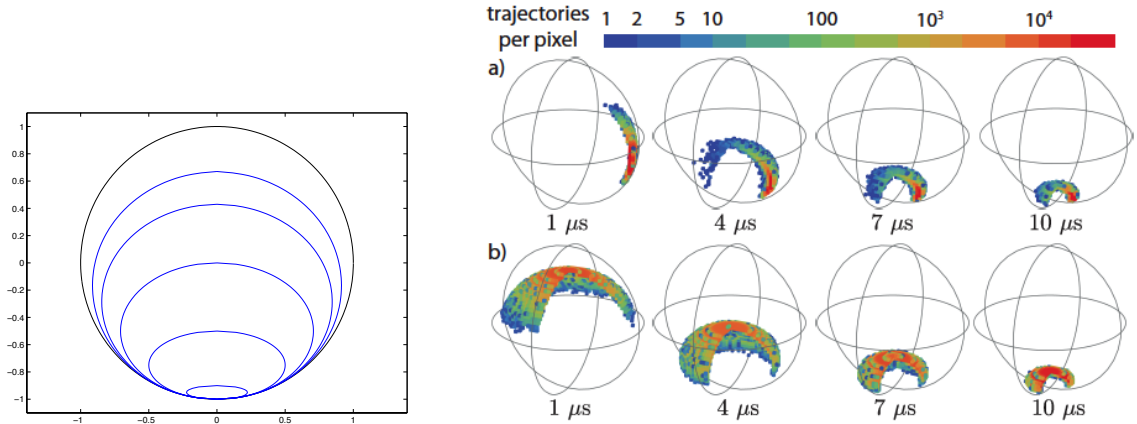
Figure 3: For $L_1 = |g\rangle\langle e|$ and $L_2 = i\,L_1$, with $H = 0$, the qubit system governed by (26) remains confined *independently of the measurement results* to the time-dependent manifold: $\frac{r^2}{2} + c_t(1+z)^2 - (1+z) = 0$, where $c_t = (c_0 - \frac{\eta}{2})\,e^{2t} + \frac{\eta}{2} \in [\frac{1}{2}, +\infty)$. The variables here refer to cylindrical coordinates (r, θ, z) in the Bloch sphere. Cuts of the surfaces corresponding to $c_t \in \{0.5, 0.6, 0.7, 1, 2, 10\}$ are represented on the left figure, with $z \in [-1, 1]$ on the vertical axis and $r \in [-1, 1]$ on the horizontal axis; quantum states correspond to $r^2 + z^2 \leq 1$. The remaining figures show experimental data from [18], on which these nonlinear manifolds were effectively observed (time evolution for two different initial states).

of settings in which the system state remains confined to a deterministic submanifold (of dimension < 3), namely:

either, $L_1 = \sqrt{\gamma}|g\rangle e$ (with possibly $L_2 = iL_1$) i.e. a spontaneous emission channel;

or, $L_1 = \sqrt{\gamma}(|g\rangle\langle g| - |e\rangle\langle e|)$ (with possibly $L_2 = iL_1$) i.e. the continuous-time version of a projective measurement in the basis $(|g\rangle, |e\rangle)$.

All the other such settings are obtained by applying a unitary basis change to one of the above, or with operators L_k that can be equivalently reformulated as one of the above.

Theorem 2: *Consider a quantum master SDE (26) on $\mathcal{H} = \mathbb{C}^N$, without Hamiltonian drive and with the set of L_k performing a QND measurement, equivalent to the continuous-time version of compatible projective measurents. I.e., there exists an orthonormal basis $\{|d_n\rangle\}_{n=1}^N$ in which all L_k write as $L_k = \sum_{n=1}^N \ell_k(n)|d_n\rangle\langle d_n|$, with all $\ell_k(n)$ real. Then the state ρ_t will be restricted to a deterministically evolving m-dimensional manifold, whose equations are*

given in that basis by:

$$\phi_t^{a,b} = \phi_0^{a,b} \text{ where } \phi_t^{a,b} = \text{phase}(\rho_t(a,b)) \quad \text{for all basis indices } a,b = 1,2,...,N \quad ;$$

$$c_t^{a,b} = \exp\left(-\sum_{k=1}^m (1-\eta_k)(\ell_k(a) - \ell_k(b))^2 \, t\right) c_0^{a,b}$$

$$\text{where} \quad c_t^{a,b} = \frac{|\rho_t(a,b)|^2}{\rho_t(a,a)\rho_t(b,b)} \quad \text{for all } a,b = 1,2,...,N \quad ;$$

$$f_t^\alpha = f_0^\alpha - 2\sigma_\alpha^2 \, t$$

$$\text{where } f_t^\alpha = \sum_{b=1}^N \alpha_b \log(\rho_t(b,b)) \quad \text{and} \quad \sigma_\alpha^2 = \sum_{b=1}^N \sum_{k=1}^m \alpha_b \, (\ell_k(b))^2 \, \eta_k$$

$$\text{with any } \alpha \in \mathbb{R}^N \text{ solving } \sum_{b=1}^N \alpha_b = 0 \, ,$$

$$\sum_{b=1}^N \alpha_b \, \ell_k(b) = 0 \text{ for all } k = 1,2,...,m \, ,$$

$$\text{and } \sigma_\alpha^2 > 0 \, .$$

The notation $c^{a,b}$ refers to coherence among eigenstates a and b of the measurement process; we can see that these coherences decay faster if the corresponding eigenvalues of the L_k differ more — a fact not visible when writing a projective measurement — and in particular for equal eigenvalues, there no decay (as the measurement would perform no distinction at all). The f_t^α characterize how all the diagonal elements (eigenstate populations) depend on each other; they are in fact deterministic functions of the *integral* of the measurement output signals.

The proof of Theorem 1 builds on an exclusive investigation of possible measurement settings, that can be found in [50]; the most challenging part is to check the independence of vector fields induced by the superoperators. The proof of Theorem 2 just follows the steps outlined above, with minimal creativity required to find the variables that work. The second case of Theorem 1 is a particular instance of Theorem 2 for the qubit. A paper presenting relevant generalizations of our deterministic manifolds study to quantum systems of higher dimension than a qubit is under preparation.

For such idealized cases, a direct application of the deterministic manifolds is for designing efficient filters, knowing that the measurements will only matter for determining few degrees of freedom. Developing a theory for approximately invariant manifolds, towards managing realistic approximate yet robust systems, should be a following step. Besides, the insight from manifolds could be exploited for data analysis [27], parameter estimation or similar tasks. Last but not least, the related analytic insight may help in reducing the complexity of controller design.

4 Algorithmic procedures on manifolds and for quantum systems

This section regroups several proposals that we have made for designing control strategies. Their common point is to focus on simple, realistically achievable procedures and proving how these allow to improve the state-of-the-art in controlling complex systems. We end with one quantum algorithmic primitive, which testifies of my increasing interest in the interplay

between quantum algorithms (computer science viewpoint) and hardware (dynamical systems viewpoint).

4.1 Integral action on Lie groups

This work has been carried out during the first part of the PhD thesis of my student Zhifei Zhang at Ghent University (second part at ECUST Shanghai; defense pending). It has been published in [68, 66, 67]. We must note that similar results were developed independently by colleagues at the same time [34].

Integral action is a simple feedback mechanism to cancel the effect, on the steady-state value, of unknown biases acting on a system. Consider a linear Single-Input-Single-Output system

$$\tfrac{d}{dt}x = Ax + Bu + d \quad , \quad y = Cx$$

where x is the state, y is the output, u the control input, d a constant bias. Integral action would define

$$u = u_0 + k_I u_I \quad , \quad \tfrac{d}{dt}u_I = y$$

where u_0 is a standard controller and $k_I u_I$ is the integral action. For the system to be in steady-state, we need $\tfrac{d}{dt}u_I = y = 0$. Thus, provided the gains can be tuned to make this system asymptotically stable, we are certain that it will converge towards $y = 0$ exactly, irrespective of the value of d; usual observability / controllability conditions allow to check if such a tuning exists.

Our problem was how to generalize such integral action to nonlinear manifolds. Indeed, if y does not belong to a vector space, its integral cannot be defined. Corresponding applications include in particular robotic control, where the variable of interest lives in a Lie group of rigid body motions. We have provided a solution for integral control on Lie groups as follows.

To define integral control, we propose:

$$\tfrac{d}{dt}u_I = u_0 \quad \text{in the Lie algebra,}$$

where u_0 is the "nominal" control action. The rationale is that:

- controls belong to tangents spaces, and for a Lie group we can map all the tangent spaces canonically onto the Lie algebra, which is a vector space; thus integration of controls is well defined;

- the nominal feedback control u_0 often features a term somehow "opposite" to the configuration error, and such that $u_0(y) = 0$ if and only if y is at the target. We thus obtain a similar characterization of having truly $y = 0$, independently of biases d;

- as the linear integral action says "the longer y is off in the same direction, the stronger we should push", our proposal on a nonlinear manifold would say "the longer we have been pushing in the same direction, the stronger we should push".

The problem for generalizing this approach from Lie groups to general manifolds would be that there is no canonical nor generically easy global mapping between tangent spaces at different

points; only on Lie groups is this operationally well defined, e.g. as expressing controls 'in body frame'.

We have proven that such integral control does have the expected effect for standard systems on Lie groups

Theorem 1 [68]:
(a) *Consider a first-order system with PI control and input bias:*

$$L_{g^{-1}} \frac{d}{dt} g = -k_p \, grad\phi + k_I u_I + d \, , \quad \frac{d}{dt} u_I = -k_p \, grad\phi \, ,$$

where g is the position on Lie group, ϕ is a cost function indicating the nominal control objective with $\phi = 0$ as only local minimum, and $u_0 = -k_p \, grad\phi$ drives the system nominally towards minima of ϕ.
This system converges globally to a set where $grad\phi = 0$, and only the equilibrium point with $u_I = -d/k_I$ and $\phi = 0$ is stable. The basin of attraction of this point can be increased up to the next lowest-lying critical point of ϕ by increasing k_p.
(b) *Consider a second-order system with PID control and input bias:*

$$L_{g^{-1}} \frac{d}{dt} g = \xi, \quad \frac{d}{dt}\xi = -k_p grad\phi - k_D \xi + k_I u_I + d$$
$$\frac{d}{dt} u_I = -k_p grad\phi - k_D \xi \, .$$

This system, under the condition $k_I < k_D$, converges globally to an equilibrium set where $grad\phi = 0$, $\xi = 0$, $u_I = -d/k_I$. The stability, and basin of attraction for large k_p, hold as for point (a).

<u>Proof idea:</u> Use a Lyapunov function which is the weighted sum of the cost function ϕ, the squared velocity for case (b), and a quadratic term for the integral action. □

A first extension of this basic result allows d to depend smoothly on the configuration g, or in fact on the state [66]. Consider for instance that we want to stabilize a pendulum at some precise angle ϕ_0, but we do not know the shape of the local gravitational potential. With integral action, we just need to measure $\phi - \phi_0$ precisely, and u_I will adjust automatically to cancel the gravitational potential appropriately. In the framework of nonlinear state space, this takes particular importance by allowing bias fields which do not need to satisfy any translational invariance; indeed on manifolds the latter may even be hard to define. We in fact conjecture that this "don't-care" attitude could be brought even one step further: I believe that *on general manifolds, defining integral action via any reasonably smooth transport map between tangent spaces at different points, would always lead to a system with bias-canceling convergence properties similar to Theorem 1.*

A second extension, motivated by robotics applications, is to consider *underactuated* or also called *nonholonomic systems*. In this case the control u spans a restricted subset – typically an affine subspace – of the Lie algebra. For instance, for a vehicle moving in the plane, the translation velocity in body frame may be fixed (both in norm – e.g. an airplane keeping its constant speed – and in direction, as most vehicles cannot translate sideways) and only the rotation rate of the heading can be controlled via steering. In such nonholonomic systems, if the bias does not lie exactly in the subset spanned by u, then it cannot just be canceled. However, a smart modified design of our integral controller can allow to nevertheless

achieve *exactly* some objectives adapted to nonholonomic systems, like stabilizing them in a coordinated motion with a reference system. A coordinated motion implies that the Lie group relative position remains constant over time, but not necessarily equal to zero (as the bias may forbid this). We prove in [67] that for steering controlled planar vehicles, subject to arbitrary bias types, a modified integral controller can stabilize them in coordinated circular motion *exactly* and *globally*.

4.2   Quantum consensus and symmetrization

This work has been carried out in collaboration with Francesco Ticozzi from University of Padova, in the framework of the PhD thesis of his student Luca Mazzarella. It has been published in [36, 48, 37]

The basic task of consensus as explained in Section 2.3 requires a set of N subsystems to reach agreement on their states $x_1 = x_2 = ... = x_N$. When the joint state space of N subsystems is richer than the Cartesian product of individual subsystems, the situation can be richer. For instance a probability distribution over the result of throwing N dices can involve all sorts of correlations, which would not be visible in the marginal over a single die. This *tensor product* property readily extends to the state space of several quantum systems. This requires to redefine both the meaning of consensus, and appropriate algorithmic procedures.

While we have discussed several choices in [36], the most fundamental definition appears to be as follows.

Definition: *Consider $\rho(x_1, x_2, ..., x_N)$ a function over N variables of the same sort. We say that ρ is in consensus if it is invariant under any permutation of the variables, i.e.*
$\rho(x_1, x_2, ..., x_N) = \rho(x_{\pi 1}, x_{\pi 2}, ..., x_{\pi N})$ *for any permutation π of the indices $1, 2, ..., N$.*

The notation ρ here hints at the quantum state (a density operator), but other interpretations are possible. The definition also suggests an operational way to achieve such consensus: swap subsystems with a nontrivial probability $\alpha \in (0, 1)$. Formally, for instance in discrete-time, at each step t we would request to perform:

choose m disjoint pairs of connected agents (j_l, k_l), $l = 1, ..., m$;

for $l = 1$ to m do:

$\rho(x_1, x_2, ., x_{j_l}, ..., x_{k_l}, ..., x_N) \ \mapsto$
$\quad (1 - \alpha)\rho(x_1, x_2, ., x_{j_l}, ..., x_{k_l}, ..., x_N) + \alpha\rho(x_1, x_2, ., x_{k_l}, ..., x_{j_l}, ..., x_N)$;

Of course this procedure only applies to systems on which this formal algorithm is operationally meaningful, see below. When the pairs to be swapped are chosen in a sequential or specific random way, convergence properties towards consensus can be proved; we do not detail this here since the main contribution was rather the formalization. The main idea can be understood on two variables: the state $\frac{1}{2}(\rho(x_1, x_2) + \rho(x_2, x_1))$ is obviously invariant under permutation of x_1 and x_2.

For quantum systems, which include classical probability distributions as a special case, the above procedure can be made more explicit in the form of a Kraus map (completely positive trace-preserving map, the most general form of a quantum information channel). It can also be formulated in continuous-time as a Lindblad equation:

$$\frac{d}{dt}\rho(t) = \sum_{k=1}^{E} \ \gamma_k \left(S_k\rho(t)S_k - \rho(t)\right) . \tag{22}$$

Here index k runs over all the E edges of the interconnection graph, S_k is the corresponding unitary and Hermitian operator swapping the states of the subsystems connected by this edge, and γ_k is the swapping rate. This continuous-time viewpoint is further elaborated in [48]. We note that this procedure homogenizes the quantum state by *mixture*, so relevant applications (e.g. in sensing or reservoir engineering) would have to be connected to such procedure.

In a more abstract sense, the above algorithm appears like a *lift* from the actual state space to the group of permutations. This suggests to lift "consensus" algorithms to "symmetrization" algorithms, as we elaborate in [37]. Starting from the trivial permutation, apply a series of probabilistic "generator" permutations in order to drive the expected state to a uniform distribution over all possible permutations. This permutations need not be pairwise, depending on the application context. Next one may replace permutations by other groups, and speak about symmetrization over groups, with various possible operational meanings as in random circuits or in quantum dynamical decoupling sequences.

4.3 Stabilizing a Schrödinger cat of light with a stream of atoms

This work has been a major contribution of my postdoc with Pierre Rouchon, and in collaboration with experimental physicists at Laboratoire Kastler Brossel (LKB, hosted by ENS Paris / Collège de France). I had worked out a procedure apparently known as reservoir engineering, to stabilize the state of a quantum system — here one mode of an electromagnetic resonator – thanks to repeated tailored interaction with auxiliary systems — here a stream of atoms with two particular levels close to the resonator mode frequency. This work has been published in [52, 51].

Consider $\mathcal{H}_c$ the Hilbert space of a harmonic oscillator (the target system) and $\mathcal{H}_k \simeq \mathbb{C}^2$ the Hilbert space of auxiliary qubit number k, for $k = 1, 2, \dots$. In the first envisioned experimental setting, the harmonic oscillator is a cavity-Quantum-Electro-Dynamics (CQED) mode and the auxiliary qubits are constituted by two Rydberg levels of a stream of Rubidium atoms, flying one after the other through the setup in order to control it. As a general principle, the interaction with one atom implies for the cavity state ρ_c:

$$\rho_c \mapsto U(\rho_c \otimes \rho_a(0))U^\dagger \mapsto \mathrm{trace}_k U(\rho_c \otimes \rho_a(0))U^\dagger = M_1 \rho_c M_1^\dagger + M_2 \rho_c M_2^\dagger \;.$$

Here ρ_a is the initial state of each auxiliary atom; U is the unitary expressing the atom-cavity interaction; the partial trace over auxiliary system k expresses that the atom is discarded after interaction; and we finally express the evolution of ρ_c as a Kraus map, where operators M_1 and M_2 on $\mathcal{H}_c$ satisfy $M_1^\dagger M_1 + M_2^\dagger M_2 =$ Identity. Contrary to Hamiltonian actuation, such a Kraus map can induce nonunitary evolution and thus, as we repeatedly apply it for $k = 1, 2, \dots$, it can stabilize ρ_c towards a target state. The challenge is to design interesting Kraus maps M_1, M_2 from experimentally available interactions, and this is called *reservoir engineering* [45].

For the LKB setup just mentioned, the standard interaction model is [29] the Jaynes-Cummings Hamiltonian H_{JC}. In a proper interaction representation:

$$H_{JC} = \frac{\delta}{2}(|e\rangle\langle e| - |g\rangle\langle g|) + i\frac{\Omega(t)}{2}(\,|g\rangle\langle e|\,\mathbf{a}^\dagger - |e\rangle\langle g|\,\mathbf{a}\,) \tag{23}$$

where $|e\rangle$ and $|g\rangle$ are the two atomic levels involved; δ is the detuning between $e-g$ transition and frequency of the cavity mode; $\mathbf{a}$ is the photon annihilation operator in the cavity mode; and $\Omega(t)$ is the atom-cavity coupling strength, varying with time during the atomic transit through the cavity mode. This Hamiltonian expresses that energy of one photon can be transformed into atomic excitation, and vice versa, while detuning makes this process less resonant as it perturbs the integration of this transition over time.

From there we can compute explicitly the unitary evolution operators corresponding to resonant ($\delta = 0$) and strongly dispersive ($|\delta| \gg \Omega$) interactions:

$$
\begin{aligned}
U_r(\Theta) \quad &= |g\rangle\langle g| \, \cos(\tfrac{\Theta\sqrt{\mathbf{N}}}{2}) + |e\rangle\langle e| \, \cos(\tfrac{\Theta\sqrt{\mathbf{N+I}}}{2}) - |e\rangle\langle g| \mathbf{a}\frac{\sin(\tfrac{\Theta\sqrt{\mathbf{N}}}{2})}{\sqrt{\mathbf{N}}} + |g\rangle\langle e| \, \frac{\sin(\tfrac{\Theta\sqrt{\mathbf{N}}}{2})}{\sqrt{\mathbf{N}}} \mathbf{a}^\dagger \, , \\
U_d(\phi_0) \quad &\approx |g\rangle\langle g| \, e^{-i\,\phi_0 \mathbf{N}} + |e\rangle\langle e| \, e^{+i\,\phi_0(\mathbf{N+I})} \, ,
\end{aligned}
\tag{24}
$$

where $\mathbf{N} = a^\dagger a$ in the functions' arguments is the photon number operator and $\mathbf{I}$ the identity operator; $\Theta = \int \Omega(t)\,dt$ and $\phi_0 = -1/(4\,\delta) \int \Omega^2(t)\,dt$.

Resonant interactions U_r with $0 < u \ll 1$ and $\Theta \ll 1$, stabilize a coherent state $|\alpha\rangle$ of amplitude $\alpha \approx 2u/\Theta$, which is the quantum equivalent of a classical oscillator state of complex amplitude α. This can be understood by writing the continuous-time approximation corresponding to this weak interaction-per-atom: it gives the standard Lindbladian for coherent state stabilization. In [51] we investigate this stabilization in more detail, as it forms the basis of the following.

The main contribution of our work was to realize that the setup can stabilize more interesting states when we vary δ during the atom-cavity interaction, making thus a combination of resonant and dispersive interaction. Explicitly, we propose to apply

$$
U = U_c = \ U_d(\phi_0) \, U_r(\Theta) \, U_d(-\phi_0) \, .
$$

Using the commutation relation $\mathbf{a} f(\mathbf{N}) = f(\mathbf{N+I})\,\mathbf{a}$, we observe that

$$
U_c = \exp[-ih_0(\mathbf{N})] \, U_r(\Theta) \, \exp[ih_0(\mathbf{N})] \, ,
$$

with $h_0(\mathbf{N}) = \phi_0 \mathbf{N}(\mathbf{N+I})$. Therefore, as repeated interactions with U_r stabilize a coherent state $|\alpha\rangle$, repeated interactions with this composite U_c stabilize a state $\exp[-ih_0(\mathbf{N})]\,|\alpha\rangle$. This indeed follows exactly, by a change of frame of unitary $\exp[ih_0(\mathbf{N})]$ on the cavity state.

The main point is that h_0 is quadratic in $\mathbf{N}$, a so-called Kerr Hamiltonian. Transforming a coherent state $|\alpha\rangle$ through a Kerr Hamiltonian can lead to a series of interesting states, depending on the length of the interaction. In particular, for $\phi_0 = \pi/2$, we can stabilize up to an irrelevant phase the so-called "Schrödinger cat" state

$$
|C_\alpha\rangle = (|\text{-}i\alpha\rangle + i\,|i\alpha\rangle)/\sqrt{2} \, ,
$$

which is a coherent superposition of two opposite coherent states. For smaller ϕ_0 (shorter dispersive interaction), one can get superpositions of several coherent states, or other quantum states with negative values in the Wigner function, which can be seen as a quantum signature [29]. Some of the states that are stabilized by this scheme are represented on Figure 4. They include all experimentally realistic imperfections with respect to the idealized situation just presented; more details can be found in [52, 51], including a study of performance of this stabilization strategy for countering cavity decay. The latter can again be done analytically
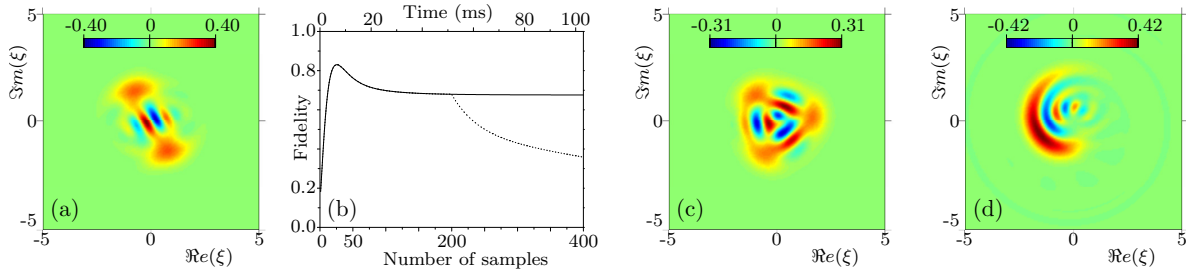
Figure 4: Non-classical states stabilized by the engineered reservoir, including realistic experimental uncertainties and perturbations. (a) Wigner function of the cavity field after 200 steps of reservoir-atom interaction. The state is close to $|C_\alpha\rangle$. (b) Solid line: fidelity of the generated state against the closest $|C_\alpha\rangle$ as a function of number of interactions (bottom axis) or of time (upper axis). Dashed line: reservoir is switched off after 200 interactions. (c) and (d) Wigner functions of stabilized cavity fields close to a superposition of 3 coherent states, and to a 'banana state' respectively. Detailed conditions in [52].

on an approximate continuous-time model. A striking fact, used in later work, is that since $\mathbf{a}|\alpha\rangle = \alpha|\alpha\rangle$, under a photon loss the cat state $|C_\alpha\rangle$ just flips its superposition phase.

In [51] we have extended this scheme to stabilizing entangled states of two cavities, provided the atom can interact first with cavity 1, then with cavity 2, then again with cavity 1; this might be realistic in so-called atomic fountain setups. For our group this research was also an eye-opener on reservoir engineering as a stabilization tool (instead of measurement-based feedback):

- In the meantime we have extended reservoir designs for the LKB setup to stabilizing photon number states, and highly squeezed states thanks to a two-step procedure where atoms enter with alternating initial states; and possibly involving entanglement among atoms.

- In the following years the members of the INRIA/QUANTIC research group have further focused on continuous-time versions of reservoir engineering in order to develop efficient hardware towards actual quantum information processing in circuit-QED setups. This has made further use of the Schrödinger cats of a harmonic oscillator state, and connects us to the next section.

4.4 Remote stabilization of entanglement thanks to flying cat states

This work has been done in collaboration with Mazyar Mirrahimi at INRIA / QUANTIC, and published in [49].

The Schrödinger cat states of a harmonic oscillator, as presented in the previous section, have been realized to possess a particularly nice feature for countering information loss in quantum information processing: under spontaneous photon loss, which is the main loss process on harmonic oscillators, they just flip a sign, and a second photon loss would bring them back to their initial state. In [41], this feature has been further picked up to stabilize a two-dimensional Hilbert subspace, spanned by the two two-legged cats with opposite relative

phase, in order to: (i) encode a logical qubit; (ii) protect it from a whole range of other decoherence channels, thanks to a smart continuous-time reservoir engineering; and (iii) perform operations of quantum logic on this encoded information. This has led to a whole series of paper in the INRIA/QUANTIC groups and by collaborators, as it is currently a particularly promising instance of continuous-variable Quantum Error Correction (QEC), in particular for circuit-QED implementations.

Leaving this line of the work to others, I have nevertheless proposed together with Mazyar Mirrahimi a scheme where a stream of cats are used as auxiliary systems interacting one after the other with the target system, in a setting resembling the one of the previous section but where the final stabilization is achieved by measurement and feedback.

The goal is to stabilize an entangled states of two qubits,

$$|\bar{\psi}\rangle = \frac{|g\rangle|g\rangle + |e\rangle|e\rangle}{\sqrt{2}} = \frac{(|g\rangle+|e\rangle)(|g\rangle+|e\rangle)+(|g\rangle-|e\rangle)(|g\rangle-|e\rangle)}{2\sqrt{2}} , \tag{25}$$

which thus cannot be written as a product of single-qubit states. These entangled states can have useful applications for applying quantum teleportation, towards setting up quantum repeaters for quantum-secure communication protocols [42] or performing modular quantum computing [25]. The challenge is that entanglement cannot be generated without physical interaction, while ideally we want the entangled subsystems to be located at remote places.

Our proposal for stabilizing $|\bar{\psi}\rangle$ is thus to use a stream of auxiliary systems, which each interact consecutively with the two qubits. The challenge remains that these auxiliary systems will have to cover a remote distance along a lossy channel, and our proposal is precisely to use cat states such that this loss has no detrimental effect on the target systems. Focusing on the effect of the loss channel, and in agreement with current experimental state-of-the-art, we assume all local operations to be perfect, or at least acceptable for the application at hand as only the loss channel is distance-dependent.

Concretely, we propose the following measurement setup which is represented on Figure 5:

- At each iteration, we initialize an auxiliary "probe" system p, modeled by the Hilbert space of a harmonic oscillator mode, in the cat state

$$|\psi\rangle_p = |C_\alpha^+\rangle_p := \frac{|\alpha\rangle_p + |-\alpha\rangle_p}{\mathcal{N}_+} ,$$

where $\mathcal{N}_+$ is a normalization constant exponentially close to $\sqrt{2}$ as α increases.

- At the first target qubit A, we apply the unitary interaction:

$$U_A |g\rangle_A \left|C_\alpha^\pm\right\rangle_p = |g\rangle_A \left|C_\alpha^\pm\right\rangle_p, \quad U_A |e\rangle_A \left|C_\alpha^\pm\right\rangle_p = |e\rangle_A \left|C_\alpha^\mp\right\rangle_p ,$$

thus flipping the relative phase of the cat conditional on the logical state of A.

- The losses occurring during transmission of the probe from A to B are modeled as an interaction with the environment:

$$U_{BS}^\eta \left|C_\alpha^\pm\right\rangle_p |0\rangle_{\text{env}} = \frac{1}{N^\pm} \left(|\sqrt{\eta}\,\alpha\rangle_p \left|\sqrt{1\text{-}\eta}\,\alpha\right\rangle_{\text{env}} \pm |\text{-}\sqrt{\eta}\,\alpha\rangle_p \left|\text{-}\sqrt{1\text{-}\eta}\,\alpha\right\rangle_{\text{env}} \right) .$$
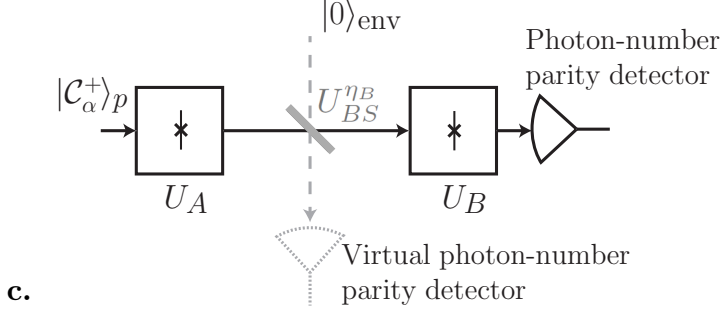
c.

Figure 5: Sketch of our concept involving target qubits $|q_A\rangle, |q_B\rangle$ and a probe field initialized in a coherent superposition of two opposite coherent states ("cat state"), to measure the qubits' parity in a QND and degeneracy-preserving way despite the presence of the losses to the environment along the line connecting the two qubits.

.

- At the second target qubit B, we apply the same procedure as for A, but taking into account that the cat amplitude has been shrinking in expectation, due to photon loss along the channel:

$$U_B \, |g\rangle_B \left|C^{\pm}_{\sqrt{\eta}\alpha}\right\rangle_p = |g\rangle_B \left|C^{\pm}_{\sqrt{\eta}\alpha}\right\rangle_p, \quad U_B \, |e\rangle_A \left|C^{\pm}_{\sqrt{\eta}\alpha}\right\rangle_p = |e\rangle_B \left|C^{\mp}_{\sqrt{\eta}\alpha}\right\rangle_p.$$

- Finally, after interaction with the second qubit, a measurement (e.g. of photon-number parity observable) projects the probe's state onto $\left|C^{+}_{\sqrt{\eta}\alpha}\right\rangle_p$ or $\left|C^{-}_{\sqrt{\eta}\alpha}\right\rangle_p$.

For $\eta = 0$, this scheme performs a Quantum-Non-Demolition measurement of joint parity: the probe output is $qC^{+}_{\sqrt{\eta}\alpha}$ for targets in span$\{|gg\rangle, |ee\rangle\}$ and $qC^{-}_{\sqrt{\eta}\alpha}$ for targets in span$\{|ge\rangle, |eg\rangle\}$. For $\eta \neq 0$, the loss channel can have flipped the cat, which would lead to an erroneous measurement output *but not perturb the target state*. The only perturbation comes from $\mathcal{N}_+ \neq \mathcal{N}_-$, such that a photon loss is not exactly equally likely when A is in state $|e\rangle$ or $|g\rangle$. This effect is however exponentially small in α such that, thanks to the cat encoding, the photon loss channel keeps the Quantum Non-Demolition character of the measurement: it just reduces its contrast such that now, essentially, output qC^{+}_{α} is *more likely* for targets in span$\{|gg\rangle, |ee\rangle\}$, and conversely. The actual parity is then estimated via simple Bayesian reasoning.

Feedback stabilization on the basis of QND measurements with limited contrast is not a big deal. Repeatedly, we apply:

- Measure the parity according to the setup of Figure 5. Update the estimate of state parity according to a Bayes rule.

- If the estimated parity is odd, apply a pulse $|g\rangle \mapsto |e\rangle$, $|e\rangle \mapsto |g\rangle$ to the second qubit, else do nothing.

- Apply a Hadamard gate to both qubits, mapping $|g\rangle \mapsto (|g\rangle + |e\rangle)/\sqrt{2}$, $|e\rangle \mapsto (|g\rangle - |e\rangle)/\sqrt{2}$.
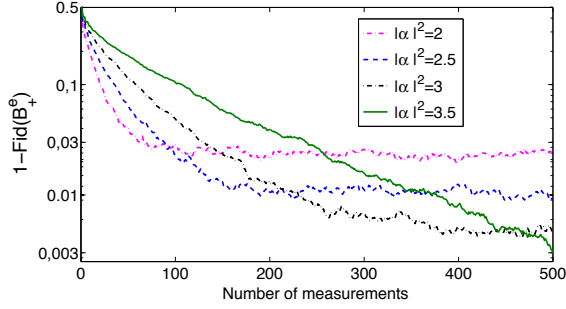
35

Figure 6: Feedback stabilization of the Bell state $\left|B_+^e\right\rangle = |\bar{\psi}\rangle$.

The Hadamard gates ensure that we measure the parity once out of two in the $|e\rangle, |g\rangle$ basis and once in the $(|e\rangle + |g\rangle), (|e\rangle - |g\rangle)$ basis, thereby singling out the entangled state $|\bar{\psi}\rangle$. Standard analysis of QND measurements allows to guarantee convergence to this state at a given rate. Having $\mathcal{N}_+ \neq \mathcal{N}_-$ implies a small perturbation on the QND character of the measurement, in fact inducing a dephasing that must be countered with the feedback action. Taking larger α allows to make $\mathcal{N}_+ \approx \mathcal{N}_-$ exponentially, while the loss in QND measurement contrast due to increased photon loss probability is only polynomial in α; therefore, larger α allows to stabilize $|\bar{\psi}\rangle$ with better accuracy, yet at a slower rate that may be more prone to other disturbance sources.

All these claims are worked out and analyzed in more detail in [49], both semi-analytically and through simulation. It appears that entanglement fidelities above 99%, as would be truly useful for quantum technology, should be within experimental reach (see Figure 6).

4.5 Exponential stabilization of QND eigenstates

This work has been carried out during the Phd thesis of Gerardo Cardona at INRIA/QUANTIC, under the joint supervision of Pierre Rouchon and myself. The corresponding publications are [20, 19].

Compared to the previous section, we now address QND measurement in continuous time, working with stochastic processes as can be obtained when continuously reading an electromagnetic field leaking out of the quantum setup. We consider a quantum system under weak continuous-time measurement, as in Section 3.3:

$$
\begin{aligned}
d\rho_t &= \sum_k (L_k \rho L_k^\dagger - \tfrac{1}{2} L_k^\dagger L_k \rho - \tfrac{1}{2} \rho L_k^\dagger L_k) \, dt \\
&\quad + \sum_k \sqrt{\eta_k}(L_k \rho + \rho L_k^\dagger - \text{trace}(L_k \rho + \rho L_k^\dagger)\rho) \, dW_k \ , \\
dY_{t,k} &= \sqrt{\eta_k}\text{trace}(L_k \rho + \rho L_k^\dagger) \, dt + dW_k \ .
\end{aligned}
\tag{26}
$$

Here ρ is the density operator on Hilbert space $\mathcal{H}$ and $d\rho_t$ its Bayesian update according to measurement results $dY_{t,k}$. The latter involve a measurement efficiency $\eta_k \in [0,1]$ and an operator L_k on $\mathcal{H}$ characterizing the measurement channel. The stochastic nature of quantum measurements is captured by the Wiener processes W_k, appearing both in the measurement and in the related update equation. The latter also contains a deterministic

decay, i.e. the usual Lindblad decoherence, expressing that measuring a quantum system in one basis necessarily induces uncertainty in other bases, whether the measurement is read or not. Interactions with an unread environment then correspond to $\eta_k = 0$.

We focus on the case where the L_k are all Hermitian and mutually commuting. This models a continuous-time version of the standard projective measurements of observables L_k. Such measurement is called Quantum Non-Demolition for the eigenstates or joint eigenspaces of the operators L_k. Indeed, if ρ_0 is an eigenstate of the L_k, then it stays invariant under such measurement; and any initial condition in fact converges exponentially to such a situation, see [20]. The issue is that we converge to eigenstate k with probability equal to its projection on ρ. We aim at adding feedback control that makes the system converge to a predefined target eigenstate $\bar{k}$, for all ρ_0. Thus

$$\rho_{t+dt} = U_t(\rho_t + d\rho_t)U_t^{\dagger} \ , \ \text{where} \ U_t = e^{-iH\,u\,dt} \tag{27}$$

expresses the infinitesimal feedback action. We write it in this way instead of differentially because a higher-order study may be necessary as $u\,dt$ can depend on Wiener processes.

In previous work, continuous-time feedback laws u have been devised as a function of ρ, and eventual convergence of the closed-loop system towards the target with probability 1 has been proved [40]. In our work we have two objectives for improvement. First, we want to prove *exponential* convergence in the sense of the expectation of a Lyapunov function measuring the distance to target; this is both simpler to establish, with Ito calculus, and stronger in terms of system performance. Second, we want to design a feedback law which does not need to run a full quantum filter estimating ρ, but which instead can be deduced from a reduced filter.

The main idea of our approach is to drive the feedback with an independent Brownian noise process dB, whose gain will be state-dependent, i.e.: $u_t dt = g(\rho_t)dB_t$. The rationale is that as the open-loop system converges stochastically to one of the eigenstates, it suffices to shake it back towards a mixture of eigenstates when it tends to converge to a different one than the target $\bar{k}$; eventually, the system will have no choice but converging to the target $\bar{k}$. Denoting Π_k the projector onto a joint eigenstate k of the Hermitian operators L_j and $p_k(t) = texttrace(\Pi_k\rho_t)$ the population on this eigenstate, we thus propose the folmowing. For some $\alpha > \beta$ close to 1 and in $(0,1)$ and some $\bar{g} > 0$:

if $p_k(t) < \beta$ for all $k \neq \bar{k}$ then $g(\rho_t) = 0$;
if $\exists k \neq \bar{k} : p_k(t) > \alpha$ then $g(\rho_t) = \bar{g}$;
else choose g anywhere between 0 and $\bar{g}$.

There remains to prove exponential convergence with this scheme.

Theorem: *Consider the system* (26),(26) *on a finite-dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^N$, with non-degenerate joint eigenstates $k = 1, 2, ..., N$ of the $\{L_j\}$. Consider any target eigenstate $\bar{k}$ and apply the feedback law $u_t dt = g(\rho_t)dB_t$ injecting noise dB with gain $g(\rho_t)$ as just described; and with, for any ρ_0 satisfying $p_k(t) > \alpha$ for some k, Hamiltonian H satisfying $trace(\Pi_{\bar{k}}e^{-iH\tau}\rho_0 e^{iH\tau}) \neq 0$ for some τ (i.e. reachability of nonzero population on target). Then there exist $\alpha_j > 0$, j running over $\{1, 2, ..., N\} \setminus \{\bar{k}\}$ and constants $c, r > 0$ such that the Lyapunov function*

$$V(\rho) = \sum_{j \neq \bar{k}} \sqrt{p_j + c\,(1 - p_{\bar{k}})} + \sum_{j \neq \bar{k}} \alpha_j p_j$$

satisfies $\mathbb{E}(V_t) \leq V_0 \, e^{-rt}$ *for all* $t \geq 0$.

Proof idea: The main challenge was to identify the Lyapunov function. The proof consists in computing the Markov generator $\mathbb{E}(dV|V)$ with Ito calculus, and proving lower bounds on the terms appearing. In open-loop, the second term in the Lyapunov function disappears from the Markov generator. The first term is chosen concave, with a singularity in the derivative at $\rho = \Pi_{\bar{k}}$ only, to get favorable terms in the second-order derivative from the Ito correction; the open-loop convergence then makes this term decrease everywhere, except for *rho* close to $\rho = \Pi_k$. In the latter neighborhoods, for $k \neq \bar{k}$, the second term of the Lyapunov function must provide the exponential decrease when noise is turned on. Here the problem is that for a given fixed H, it may be impossible to *directly* transfer population to $\bar{k}$: maybe from the neighborhood of a level k, the state must first transit through several intermediate levels. The progress of this transit is quantified thanks to the α_j, which are tuned as a function of the transitions between eigenstates implied by H. $\qquad\square$

The result of the above theorem involves no finite-return-time arguments as in [40], and is therefore easy to comprehend. The exponential character gives some additional guarantees. Several further observations can be made around this basis.

- We made no particular effort to optimize the convergence rate in this general setting. On a qubit, we have been able to obtain an exponential convergence rate arbitrarily close to the open-loop convergence rate of the measurement process towards its set of eigenstates.

- We have assumed a quite strict actuation constraint, i.e. a single Hamiltonian whose gain we can adapt. More efficient strategies are possible if we have independent control over several Hamiltonians.

- The Theorem states a case where the measurement channels isolate non-degenerate eigenstates. The same principles apply – up to some basic assumptions on the actuation Hamiltonian – for stabilizing subspaces. In particular, in [19] we show how our strategy can be applied to a continuous-time version of the repetition code used for correcting bit-flips in Quantum Error Correction (QEC).

- The feedback decision only depends on the populations p_k, i.e. the diagonal of ρ in the eigenbasis of the measurement operators. We can thus try to reduce the filter to these variables. If we apply independent noise on the Hamiltonian associated to each transition, then this reduction can be exact. Otherwise, approximate reduced filters must be investigated when the noise is turned on. Furthermore, the p_k are quite directly related to the measurement outputs: in the case of L_k providing binary information, like the syndrome measurements of QEC, they are even directly proportional to the expected measurement outputs. This suggests that elementary filters, performing just e.g. low-pass smoothing of the output signals, might suffice towards efficient feedback.

This last point is a subject that I want to further pursue: developing low-dimensional dynamical controllers for quantum systems, in opposition to the non-scalable ones based on full state observers and the limited static controllers [65]. The QEC repetition code in particular is subject of ongoing work in this direction.

4.6   Quantum Fast Forwarding of Markov chains

This work is the second part of the PhD thesis of Simon Apers under my supervision at Ghent University. It has been published in [7].

We have seen that in principle, for any Quantum Random Walk on a graph, there exists a classical Markov chain which produces the same distribution after the same number of steps. However, in that result, we put aside the complexity that may be necessary for designing the Markov chain: our proof-of-principle in fact computes explicitly what the full distribution should be, and embeds it in a classical construction. This approach may still be relevant in applications to design networks for physical mixing. However, in many algorithmic applications, one uses a Markov chain that is only specified in terms of "simple local rules", to learn something about the induced distribution. In this case, we should instead compare what the "simple local rules" can allow us to do in classical or quantum cases. We have provided a partial answer in this direction, and a constructive one, by showing how a quantum primitive can quadratically accelerate the simulation of a Markov chain.

Consider a symmetric random walk on a graph with transition matrix P. In a typical algorithm, when the walker sits at a given node, it can use knowledge of some short memory and links in its vicinity in order to derive a decision on where to jump next. In the quantum case, it is customary to take over such "local query models", plus assuming that they work in superposition: if the input is in a quantum superposition of two graph nodes, then it gets back the superposition of local information as well.

The objective of our Quantum Fast Forwarding (QFF) algorithm is to provide the distribution of the Markov chain at a given, arbitrary time $t > 0$ when starting at a given, arbitrary node of the graph. This is in contrast to existing results which only focus on the asymptotic behavior for large t, by proving properties of the spectral gap [61, 54]. Examining properties of a Markov chain at intermediate times, thus its transient, can give useful information about a graph, like identifying the presence of bottlenecks or clusters. We do present a full quantum algorithm for such applications, on the basis of the clustering algorithm by Spielman and Teng [55]. Other applications are under investigation.

The result is that we have designed a QFF algorithm (or rather algorithmic primitive) with the following property.

Theorem: *Given any quantum state $|v\rangle$ representing a graph node, time $t \geq 0$ and precision $\epsilon > 0$, the QFF primitive allows to output a quantum state ϵ-close to $|P^t v\rangle / \|P^t v\|$ using*

$$O\left(\|P^t v\|^{-1} \sqrt{t \, \log(1/\epsilon \|P^t v\|)}\right)$$

expected Quantum Walk steps, and using $O(\|P^t v\|)$ reflections around $|v\rangle$.

Here $|P^t v\rangle$ must be understood as the quantum vector use entries, in graph-nodes-basis, are the values of the classical vector $P^t v$. The norm represents the 2-norm, and part of the slow-down just comes from changing the 2-norm of $P^t v$, which is quantum relevant, while the one-norm of course remains conserved (total probability). The essential element is however that we simulate t steps with order $\sqrt{t}$ quantum steps.

The main ideas of the algorithm go as follows.

- We associate a quantum walk to P, whose spectral properties are closely tied to those of P.

- We show that in order to capture the essential features of the walk at time t, it is sufficient to match a truncated Chebyshev expansion of the Markov chain eigenvalue function, which we limit to size $\sqrt{t}$.

- The Chebyshev expansion can be obtained by applying a *linear combination of unitaries* procedure on the quantum walk operator [5]. This provides an accelerated version of the Markov chain, in spectral basis and thus in any basis.

Note that for the second item, it is essential for P to be symmetric. Indeed, the irreversible walks include e.g. a persistent drift along a line; there is no way in which the result of t deterministic steps along a line could be well approximated by $\sqrt{t}$ steps of a quantum random walk respecting graph locality. This points again to a phenomenon already encountered in the first part: apparently, what stands to be gained with local memories or other improvements, is to replace diffusion by more efficient mass transport. As certain walks have a highly diffusive behavior, actually repeating back-and-forth steps to implement a move, it seems possible on the basis of local or partial graph knowledge only, to make more direct effective steps. This picture can be used intuitively on lattices, its validity for general graphs hinges upon our mathematical proofs. After having identified the proper combination of algorithmic components, the remaining mathematical difficulty is to prove the relevant bounds, as everything is based on approximations.

Before concluding, we must acknowledge two facts. First, we prove to do better than all known procedures, but one must be careful about what might be found as alternatives, also in classical settings. Second, one must be especially careful about what power a "quantum query model" carries intrinsically, in addition to just classical encoding of the data. Indeed, for a range of algorithms related to large low-rank matrices, the recent work of [62] has clarified that just sampling from the quantum data and applying a classical algorithm already provides a key speedup.

5 Conclusion

This document summarizes research contributions of three different types.

In the first part, we derive hard bounds on how particular dynamical systems can be tuned to achieve certain tasks: rejecting long-range disturbances based on local measurements; remaining stable at short-range when embedded in a long chain; and accelerating convergence of symmetrizing algorithms (consensus, mixing). The mathematical effort in this part has relied mostly on identifying the proper model and property to prove, in a sufficient generality; the proofs themselves then involve no highly complicated steps. Beyond the precise answers that these bounds give, they are a starting point for exploring which elements of the setting may be changed in order to improve the situation. Sometimes a problem definition may appear to require revision, as for the noise model or the infinity-limit that would make it impossible to remain locally stable when embedded in a large network. Sometimes new elements may have to be taken into account, like design constraints for random walk mixing. These adaptations may be application-dependent. As I am now focusing on quantum computer design, besides the acquired general knowledge I am mostly keeping in mind the last application, which has led to the work on Quantum Fast Forwarding (see last part).

In the second part, we provide a set of analysis tools for quantum systems. From a dynamical systems perspective, we are addressing two aspects. First, continuously measured

quantum systems follow nonlinear stochastic differential equations with a specific structure, and it appears that this particular structure can carry a dominant deterministic component. We have investigated a few cases where this allows exact model reduction; a future step should consider how approximate model reduction can be made along the se lines. This appears essential if we want to have efficient quantum filters, taking decisions about feedback corrections at the nanosecond scale. Second, quantum models for unmeasured systems are rather simple, linear ODEs, yet on high-dimensional spaces. The key point for a more efficient analysis is to exploit specific structure: how the network is composed of components, how physical "channels" act individually on the system. Maintaining such quantum structure also allows to inspire new guidelines for design. The contribution on adiabatic elimination – although its summary in this report is rather brief – has been an essential technique repeatedly used in our group and in others, to investigate design options and effects of approximations prior to purely numerical investigation. It is bound to keep playing a role as more complex networks and systems are being built, and we are indeed keeping developing it.

In the third part, we have reviewed several proposals for dynamical primitives that stabilize interesting things. Those proposals are based on a careful analysis of essential mathematical ingredients in the dynamical system. In some sense these proposals were meant to build a deeper understanding of what is possible to achieve with dynamical systems, in terms of robust information-related objectives like rejecting biases, distributing entanglement, or speeding up algorithms. Our work has also mostly been picked up in this way, i.e. not as finished products but as giving ideas for other designs and proposals. The last item, about quantum algorithms, has initiated a track that i would like to pursue if possible in the coming years: trying to distill essential ingredients of quantum dynamics which "explain" the quantum speedup. The concrete objectives would be (i) to identify what is truly essential in the hardware dynamical systems in order to achieve quantum(-like) speedup and (ii) to develop some consolidated intuition on how to design algorithms for quantum computers. If truly impactful contributions along this second line may appear very ambitious, at least it can only help to better understand the fundaments of algorithms in order to pursue the hardware design.

Many systems among those presented in this report are related to interaction among dynamical components. All of them are related to the central theme of interplay between information and physical systems. The driving force of my current research is indeed to understand how dynamical systems can process information in efficient ways, and the quantum computer is a paradigmatic example of this. As the key is to understand how coupling imperfect systems allows to essentially obtain an interesting answer, dynamical systems and control theory, together with probabilistic tools, have an essential role to play in this endeavor.

A Background: quantum dynamical systems

We here summarize some essentials of quantum dynamical systems modeling, without delving into their interpretation. We provide the basic building blocks and notation, from which the results in the main text are built.

A.1 Closed quantum systems

By closed, we mean that the quantum system sends information to no outside device: neither some environment, nor a measurement device. The only "outside interaction" may be that we would tune some system parameters over time, which is modeled as a classical control input on the parameters.

Like in classical systems, a quantum state is meant to represent all the knowledge we have about the system and predicting its future evolution. For a closed quantum system, the state $|psi\rangle$ is a complex function on the system Hilbert space $\mathcal{H}$; for finite-dimensional $\mathcal{H} \simeq \mathbb{C}^n$ it is thus a complex vector of length n. The adjoint of $|\psi\rangle$ is denoted $|\psi\rangle^\dagger = \langle\psi|$ and the scalar product $\langle\psi_1|\psi_2\rangle$. This so-called wave-function must satisfy $\|\psi\|^2 = \langle\psi|\psi\rangle = 1$, because it represents a probability distribution – in fact, a whole set of probability distributions. To maintain this property, the discrete-time and continuous-time evolutions must be respectively

$$|\psi\rangle_{t+1} = U_t|\psi\rangle_t \quad \text{and} \quad \tfrac{d}{dt}|\psi\rangle = -iH(t)|\psi\rangle \ ,$$

where U is a unitary operator and $H(t)$ is a Hermitian operator, namely the system Hamiltonian. On the right one recognizes the standard Schrödinger equation. One may also consider stochastic differential equations for $|\psi\rangle_t$, we will come back to this later. Having a classical control input here means that we can modulate U_t or $H(t)$, thus e.g. writing $H(u(t))$.

When a quantum system is composed of several subsystems with respective Hilbert spaces $\mathcal{H}_1, \mathcal{H}_2$ for instance, its wave-function is a complex function over the tensor product space $\mathcal{H}_1 \otimes \mathcal{H}_2$. More concretely, if $|\psi_k\rangle$ is written as a function of x_k, then $|\psi\rangle$ is a joint function of x_1, x_2. For a finite-dimensional system, think of x_1 as the index of the vector ψ_1, i.e. $x_1 \in \{1, 2, 3, 4\}$ for a four-level system $\mathcal{H}_1 \simeq \mathbb{C}^4$); and assume for instance the same dimension for $\mathcal{H}_2$; then ψ is a function over all possible index combinations (1,1), (1,2),... (4,4); the tensor product amounts to the Kronecker product of the vectors. The dimension of $\mathcal{H}$ is thus the product of the dimensions of the components. This is reminiscent of joint probability distributions over several subsystems. States $|\psi\rangle \neq |\psi_1\rangle \otimes |\psi_2\rangle$ i.e. which cannot be factorized, are called entangled. The particular correlation created by entanglement plays a central role in quantum information.

To close this section, we introduce the density operator ρ. Its idea is to be a convex combination of projectors $|\psi(k)\rangle\langle\psi(k)|$ related to the possible state $|\psi(k)\rangle$ of the system, in order to model uncertainties about $|\psi\rangle$. Formally, ρ is a Hermitian operator on $\mathcal{H}$, positive semidefinite, trace-class and with $\text{trace}(rho) = 1$. This is consistent with the diagonal of ρ, in various bases, representing various probability distributions. Having ρ a convex combination of projectors and taking into account the Schrödinger equation, we have

$$\rho_{t+1} = U_t\rho_t U_t^\dagger \quad \text{and} \quad \tfrac{d}{dt}\rho = -i[H(t), \rho]$$

where $[A, B] = AB - BA$ denotes the commutator between operators on $\mathcal{H}$.

A.2 Open quantum systems: measurements

The most idealized measurement on a quantum system is a so-called *projective* measurement. This is also the prime theoretical standard, but in terms of actual applications, it is the least realistic. When measuring a quantum system, a measurement basis must be given. This is most concisely expressed via a Hermitian operator Q on $\mathcal{H}$, and the measurement will involve the spectrum of $Q = \sum_k q_k \Pi_k$, with the Π_k a set of orthonormal projectors onto eigenspaces of Q and $q_k \in \mathbb{R}$ the associated eigenvalues. A projective measurement of Q is then a discrete, probabilistic operation yielding:

$$\text{with probability } \|\Pi_k|\psi\rangle\|^2 : \quad \text{output } y = q_k$$
$$\text{state update } |\psi\rangle_+ = \frac{\Pi_k|\psi\rangle}{\|\Pi_k|\psi\rangle\|} \ .$$

With the density operator, this writes:

$$\text{with probability } \text{trace}(\Pi_k\rho) : \quad \text{output } y = q_k$$
$$\text{state update } \rho_+ = \frac{\Pi_k\rho\Pi_k}{\text{trace}(\Pi_k\rho)} \ .$$

Combing the Schrödinger equation and this projective measurement, it appears that ρ can only be rank-one, of the form $\rho = |\psi\rangle\langle\psi|$. However, the rank of ρ will change with the following models, and in this case a modeling on basis of $|\psi\rangle$ alone is not possible anymore.

The above idealized measurement model is not the closest to experiments, for two reasons. First, it assumes that no information whatsoever is lost from the box "system and measurement device". Second, we may actually choose to extract only partial information about Q, in order to also extract partial information about some other observable. The latter becomes particularly relevant if we want to write a model for continuous-time measurements. Both issues can in principle be covered by considering projective measurements not of the system itself, but of the system coupled to an auxiliary system; the effective measurement on the system can then be brought back to one of these weaker forms. For computations, it is often simpler to consider the effective weak(er) measurement directly. In this case, the mod A general weak measurement, without information loss, is defined in discrete time by a so-called POVM (positive operator-valued measurement) instrument. This generalization of the projective measurement replaces the set Π_k by a set of operators M_k, not necessarily Hermitian nor orthogonal, which just have to satisfy $\sum_k M_k^\dagger M_k = \text{Identity}$. The measurement operation yields:

$$\text{with probability } \text{trace}(M_k\rho M_k^\dagger) : \quad \text{output } y = k \text{ or equivalently } y = M_k^\dagger M_k \text{ (the POV)}$$
$$\text{state update } \rho_+ = \frac{M_k\rho M_k^\dagger}{\text{trace}(M_k\rho M_k^\dagger)} \ .$$

The constraint on the M_k ensures that total probability equals 1. Through the update equation, the rank of ρ can still not increase, i.e. if we start with a so-called "pure state" $\rho = |\psi\rangle\langle\psi|$, then the measurement process without information loss, although weak, maintains us in a pure state. This is consistent with higher rank describing information loss, as we will see in the next subsection, while here we just retrieve information more weakly.

It is sometimes more natural to view k as being actually composed of several indices $(k_1, k_2, ...)$ corresponding to "different measurement channels", with $M_k = M_{k_1} M_{k_2}... $. A

generalization to $k_1, k_2, ...$ belonging to a continuum is of course possible. In particular it makes sense to take the joint limit with time becoming continuous. This yields a Stochastic Differential Equation in which the rank of ρ never increases. Without entering into the details of the derivation, this SDE for continuous weak measurement takes the form:

with $dW_{\ell,t}$ following the probabilities of independent Wiener processes :

$$\text{output } dy_{\ell,t} = \text{trace}((L_\ell + L_\ell^\dagger)\rho_t)\, dt + dW_{\ell,t}$$

$$\text{state update } d\rho_t = \sum_\ell \left(L_\ell \rho L_\ell^\dagger - \frac{L_\ell^\dagger L_\ell \rho_t + \rho_t L_\ell^\dagger L_\ell}{2} \right) dt$$

$$+ \left(L_\ell \rho_t + \rho_t L_\ell^\dagger - \text{trace}((L_\ell + L_\ell^\dagger)\rho_t)\rho_t \right) dW_{\ell,t} \ .$$

Here the whole object $(L_\ell, dW_\ell \in [k, k + dk])$ is essentially an infinitesimal and continuous version of M_{k_ℓ}, and the equation must be understood in the Ito sense. This measurement equation can be combined with a Hamiltonian evolution.

A.3 Open quantum systems: the environment

In the above, we always assume that we perfectly measure the overall system, i.e. the target system and the auxiliary system associated to the measurement in case of a POVM, even when we describe the effect on the target system only. When we do not have infinite confidence about measurement results, or in other words when information that was present in the system is lost during the measurement process, the state ρ describes the evolution by taking the expectation over the "missing information". This is consistent with a Bayesian viewpoint, and ρ being used like an a posteriori probability describing all possible future measurements via $\text{trace}(M_k \rho M_k^\dagger)$.

The associated evolution equation in discrete time can be written by using a double index, the first index being perfectly measured while information about the second index is completely lost:

$$\text{with probability } \sum_\ell \text{trace}(M_{k,\ell} \rho M_{k,\ell}^\dagger) :$$

$$\text{output } y = k$$

$$\text{state update } \rho_+ = \frac{\sum_\ell M_{k,\ell} \rho M_{k,\ell}^\dagger}{\text{trace}(\sum_\ell M_{k,\ell} \rho M_{k,\ell}^\dagger)} \, ,$$

where thus the $M_{k,\ell}$ satisfy $\sum_{k,\ell} M_{k,\ell}^\dagger M_{k,\ell} = \text{Identity}$. (Note that $M_{k,\ell}$ here *need not* factorize as $M_k M_\ell$; the latter was just a particular example to introduce "different measurement channels".) In the most extreme case where no measurement result is retrieved, we obtain a *deterministic and linear* equation for a quantum system open to a Markovian environment, known as a Kraus map or CPTP (completely positive trace-preserving) linear map:

$$\rho_+ = \sum_\ell M_\ell \rho M_\ell^\dagger \ .$$

There is no sum over k since there is no different possible measurement results, and then since $\sum_\ell M_\ell^\dagger M_\ell = \text{Identity}$ the denominator equals 1 automatically for all ρ. In this derivation

of the Kraus map, Markovianity is hidden in our assumption that the environment is re-initialized as a device that performs a new unread measurement at each time step. This Kraus map can modify the rank of ρ. It can express dissipation or mixing of information. From a converse viewpoint, a well-designed Kraus map can be used to stabilize ρ towards a target value (while a unitary evolution, as implied by the Schrödinger equation, cannot). Indeed, consider that we actually do have access to the measurement result $y = k$, but we perform a unitary evolution V_k conditioned on the measurement result (feedback) and *after this* we throw away our knowledge of k. The expected evolution can be described as

$$\rho_+ = \sum_k V_k M_k \rho M_k^\dagger V_k^\dagger = \sum_k \tilde{M}_k \rho \tilde{M}_k^\dagger$$

and with unitary V_k we still have $\sum_k \tilde{M}_k^\dagger \tilde{M}_k = $ Identity. If $M_k = |\psi_k\rangle\langle\psi_k|$ expresses a projective measurement, then the state after measurement of k is $|\psi_k\rangle$; and nothing forbids to take V_k satisfying $V_k|\psi_k\rangle = |\psi_{\text{target}}\rangle$ for all k. Thus, with this procedure the Kraus map would stabilize the state to $|\psi_{\text{target}}\rangle$ in one deadbeat step. Of course this is just a most extreme example, not compatible with realistic experimental constraints. It is just meant to illustrate the Kraus map model on a simple case.

In continuous time, by far the most standard model for information loss involves a measurement efficiency $\eta_\ell \in [0,1]$ for each measurement channel and the SDE becomes:

with $dW_{\ell,t}$ following the probabilities of independent Wiener processes :

output $dy_{\ell,t} = \sqrt{\eta_\ell}\text{trace}((L_\ell + L_\ell^\dagger)\rho_t)\,dt + dW_{\ell,t}$

state update $d\rho_t = \sum_\ell \left(L_\ell\rho L_\ell^\dagger - \frac{L_\ell^\dagger L_\ell \rho_t + \rho_t L_\ell^\dagger L_\ell}{2} \right) dt$

$$+ \sqrt{\eta_\ell} \left(L_\ell\rho_t + \rho_t L_\ell^\dagger - \text{trace}((L_\ell + L_\ell^\dagger)\rho_t)\rho \right) dW_{\ell,t} .$$

For $\eta_\ell = 1$, no measurement information is lost. For $\eta_\ell = 0$, all the measurement information is lost and the equation becomes deterministic. This deterministic equation is the continuous-time analog of the Kraus map, and its most common name is Lindblad equation:

$$\frac{d}{dt}\rho = \sum_\ell L_\ell\rho L_\ell^\dagger - \frac{L_\ell^\dagger L_\ell \rho_t + \rho_t L_\ell^\dagger L_\ell}{2} .$$

We could add to this a Hamiltonian evolution. Like its discrete-time analog and unlike its information-preserving counterpart, the Lindblad equation can describe dissipation, inducing contraction among various trajectories or even asymptotic convergence of ρ. Depending on the goal and setting, this contraction will be viewed as information loss ("decoherence" that we want to fight) or help stabilize a target state.

References

[1] A.Farnam and A.Sarlette. "About string stability of a vehicle chain with unidirectional controller". In: *European Journal of Control* in press (2019), doi:10.1016/j.ejcon.2019.03.002.

[2] A.Farnam and A.Sarlette. "Towards a comprehensive impossibility result for string stability". In: *IEEE Transactions on Automatic Control* in press (2019), arXiv:1804.04858.

[3] Dorit Aharonov et al. "Quantum walks on graphs". In: *Proceedings of the thirty-third annual ACM symposium on Theory of computing*. ACM. 2001, pp. 50–59.

[4] Andris Ambainis et al. "One-dimensional quantum walks". In: *Proceedings of the thirty-third annual ACM symposium on Theory of computing*. ACM. 2001, pp. 37–49.

[5] A.M.Childs and N.Wiebe. "Hamiltonian simulation using linear combinations of unitary operations". In: *Quantum Information and Computation* 12 (2012), pp. 901–924.

[6] S. Apers and A. Sarlette. "Accelerating consensus by spectral clustering and polynomial filters". In: *IEEE Transactions on Control of Network Systems* 4.3 (2017), pp. 544–554.

[7] S. Apers and A. Sarlette. "Quantum Fast-Forwarding: Markov chains and graph property testing". In: *Quantum Information & Computation* 19.3-4 (2019), pp. 181–213.

[8] S. Apers, A. Sarlette, and F. Ticozzi. "Bounding the Convergence Time of Local Probabilistic Evolution". In: *Springer Lecture Notes: International Conference on Geometric Science of Information* (2017), pp. 754–762.

[9] S. Apers, A. Sarlette, and F. Ticozzi. "Simulation of quantum walks and fast mixing with classical processes". In: *Physical Review A* 98.3 (2018), p. 032115.

[10] S. Apers, F. Ticozzi, and A. Sarlette. "Lifting Markov Chains To Mix Faster: Limits and Opportunities". In: *under revision for IEEE Transactions on Information Theory* arXiv:1705.08253 (2019).

[11] A.Sarlette and R.Sepulchre. "Control limitations from distributed sensing: Theory and Extremely Large Telescope application". In: *Automatica* 50.2 (2014), pp. 421–430.

[12] Stephane Attal et al. "Open quantum random walks". In: *Journal of Statistical Physics* 147.4 (2012), pp. 832–852.

[13] (a) R. Azouit, A. Sarlette, and P. Rouchon. "Adiabatic elimination for open quantum systems with effective Lindblad master equations". In: *Proc. IEEE Conference on Decision and Control* (2016), pp. 4559–4565; (b) R. Azouit et al. "Structure-preserving adiabatic elimination for open bipartite quantum systems". In: *Proc. IFAC World Congress* 50 (2017), pp. 13026–13031; (c) R. Azouit et al. "Towards generic adiabatic elimination for bipartite open quantum systems". In: *Quantum Science and Technology* 2.4 (2017), p. 044011; (d) P. Forni et al. "Adiabatic elimination for multi-partite open quantum systems with non-trivial zero-order dynamics". In: *Proc. IEEE Conference on Decision and Control* (2018), pp. 6614–6619; (e) R. Azouit, A. Sarlette, and P. Rouchon. "Convergence and adiabatic elimination for a driven dissipative quantum harmonic oscillator". In: *Proc. IEEE Conference on Decision and Control* (2015), pp. 6447–6453; (f) P. Forni et al. "High-accuracy adiabatic elimination: formulas and examples". In: *in preparation for Physical Review A* (2019); (g) P. Forni and A. Sarlette. "Minimizing decoherence on target in bipartite open quantum systems". In: *submitted to IEEE Conference on Decision and Control* (2019).

[14] R. Azouit, A. Sarlette, and P. Rouchon. "Well-posedness and convergence of the Lindblad master equation for a quantum harmonic oscillator with multi-photon drive and damping". In: *ESAIM: Control, Optimization and Calculus of Variations* 22.4 (2016), pp. 1353–1369.

[15] B. Bamieh et al. "Coherence in Large-Scale Networks: Dimension Dependent Limitations of Local Feedback". In: *IEEE Transactions on Automatic Control* 57 (2012), pp. 2235–2249.

[16] (a) P. Barooah and J.P. Hespanha. "Error amplification and disturbance propagation in vehicle strings with decentralized linear control". In: *Proc. IEEE Conference on Decision and Control* (2005), pp. 4964–4969; (b) P. Seiler, A. Pant, and K. Hedrick. "Disturbance propagation in vehicle strings". In: *IEEE Transactions on Automatic Control* 37 (2004), pp. 1835–1842.

[17] T. Benoist, C. Pellegrini, and F. Ticozzi. "Exponential stability of subspaces for quantum stochastic master equations". In: *Annales Henri Poincaré* 18.6 (2017), pp. 2045–2074.

[18] P. Campagne-Ibarcq et al. "Observing quantum state diffusion by heterodyne detection of fluorescence". In: *Physical Review X* 6.1 (2016), p. 011002.

[19] G. Cardona, A. Sarlette, and P. Rouchon. "Continuous-time Quantum Error Correction with Noise-assisted Quantum Feedback". In: *submitted to IEEE Conference on Decision and Control* arXiv:1902.00115 (2019).

[20] (a) G. Cardona, A. Sarlette, and P. Rouchon. "Exponential stochastic stabilization of a two-level quantum system via strict Lyapunov control". In: *Proc. IEEE Conference on Decision and Control* (2018), pp. 6591–6596; (b) G. Cardona, A. Sarlette, and P. Rouchon. "Exponential stabilization of a QND measurement eigenstate". In: *in preparation* (2019).

[21] Fang Chen, László Lovász, and Igor Pak. "Lifting Markov chains to speed up mixing". In: *Proceedings of the thirty-first annual ACM symposium on Theory of computing*. ACM. 1999, pp. 275–281.

[22] (a) E.B. Davies. "Quantum dynamical semigroups and the neutron diffusion equation". In: *Reports on Mathematical Physics* 11.2 (1977), pp. 169–188; (b) E.B. Davies. "Quantum theory of open systems". In: *IAM* (1976).

[23] Persi Diaconis, Susan Holmes, and Radford M Neal. "Analysis of a nonreversible Markov chain sampler". In: *Annals of Applied Probability* (2000), pp. 726–752.

[24] Persi Diaconis and Laurent Miclo. "On the spectral analysis of second-order Markov chains". In: *Annales de la Faculté des Sciences de Toulouse. Mathématiques. Série 6* 22.3 (2013), pp. 573–621.

[25] L.-M. Duan et al. "Long-distance quantum communication with atomic ensembles and linear optics". In: *Nature* 414 (2001), pp. 413–418.

[26] (a) A. Farnam and A. Sarlette. "About strong string stability of a vehicle chain with time-headway control". In: *Proc. IEEE Conference on Decision and Control* (2018), pp. 5077–5083; (b) A. Farnam and A. Sarlette. "String Stability towards Leader thanks to Asymmetric Bidirectional Controller". In: *Proc. IFAC World Congress* (2017).

[27] (a) S. Gammelmark, B. Julsgaard, and K. Moelmer. "Past quantum states of a monitored system". In: *Physical Review Letters* 111.16 (2013), p. 160401; (b) D. Tan et al. "Homodyne monitoring of postselected decay". In: *Physical Review A* 96.2 (2017), p. 022104.

[28] (a) H. Hao and P. Barooah. "Stability and robustness of large platoons of vehicles with double-integrator models and nearest neighbor interaction". In: *International Journal of Nonlinear Control* 23 (2013), pp. 2097–2122; (b) I. Herman, S. Knorn, and A. Ahlén. "Disturbance scaling in bidirectional vehicle platoons with different asymmetry in position and velocity coupling". In: *Automatica* 82 (2017), pp. 13–20.

[29] S. Haroche and J.M. Raimond. "Exploring the quantum". In: *Oxford University Press* (2006).

[30] (a) J. Hendrickx et al. "Graph diameter, eigenvalues, and minimum-time consensus". In: *Automatica* 50.2 (2014), pp. 635–640; (b) L. Georgopoulos. "Definitive consensus for distributed data inference". In: *Ph.D. dissertation, Lab. Nonlinear Syst., EPFL Lausanne* (2011).

[31] Viv Kendon. "Decoherence in quantum walks–a review". In: *Mathematical Structures in Computer Science* 17.6 (2007), pp. 1169–1220.

[32] S. Klinge and R.H. Middleton. "Time headway requirements for string stability of homogenous linear unidirectionally connected systems". In: *Proc. IEEE Conference on Decision and Control* (2009), pp. 1992–1997.

[33] Y. Li and B. De Schutter. "Control of a string of identical pools using non-identical feedback controllers". In: *IEEE Transactions on Control Systems Technology* 20 (2012), pp. 1638–1646.

[34] DHS. Maithripala and JM. Berg. "An intrinsic PID controller for mechanical systems on lie groups". In: *Automatica* 54 (2015), pp. 189–200.

[35] FL Marquezino et al. "Mixing times in quantum walks on the hypercube". In: *Physical Review A* 77.4 (2008), p. 042312.

[36] L. Mazzarella, A. Sarlette, and F. Ticozzi. "Consensus for quantum networks: Symmetry from gossip interactions". In: *IEEE Transactions on Automatic Control* 60.1 (2015), pp. 158–172.

[37] L. Mazzarella, F. Ticozzi, and A. Sarlette. "Extending robustness and randomization from consensus to symmetrization algorithms". In: *SIAM Journal on Control and Optimization* 53.4 (2015), pp. 2076–2099.

[38] David A Meyer. "From quantum cellular automata to quantum lattice gases". In: *Journal of Statistical Physics* 85.5-6 (1996), pp. 551–574.

[39] V. Milanés and S.E. Shladover. "Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data". In: *Transportation Research Part C: Emerging Technologies* 48 (2014), pp. 285–300.

[40] (a) M. Mirrahimi and R. Van Handel. "Stabilizing feedback controls for quantum systems". In: *SIAM Journal on Control and Optimization* 46.2 (2007), pp. 445–467; (b) K. Tsumura. "Global stabilization at arbitrary eigenstates of n-dimensional quantum spin systems via continuous feedback". In: *Proc. American Control Conference* (2008), pp. 4148–4153.

[41] M. Mirrahimi et al. "Dynamically protected cat-qubits: a new paradigm for universal quantum computation". In: *New Journal of Physics* 16.4 (2014), p. 045014.

[42] C. Monroe et al. "Large-scale modular quantum-computer architecture with atomic memory and photonic interconnects". In: *Physical Review A* 89 (2014), p. 022317.

[43] S. Öncü et al. "Cooperative adaptive cruise control: network-aware analysis of string stability". In: *IEEE Transactions on Intelligent Transportation Systems* 15 (2014), pp. 1527–1537.

[44] D. Petz. "Monotone metrics on matrix spaces". In: *Linear Algebra Appl.* 244 (1996), pp. 81–96.

[45] J.F. Poyatos, J.I. Cirac, and P. Zoller. "Quantum Reservoir Engineering with Laser Cooled Trapped Ions". In: *Physical Review Letters* 77 (1996), p. 4728.

[46] P. Rouchon and A. Sarlette. "Contraction and stability analysis of steady-states for open quantum systems described by Lindblad differential equations". In: *Proc. IEEE Conference on Decision and Control* (2013), pp. 6568–6573.

[47] A. Sarlette. "Adding a single state memory optimally accelerates symmetric linear maps". In: *IEEE Transactions on Automatic Control* 61.11 (2016), pp. 3533–3538.

[48] A. Sarlette, L. Mazzarella, and F. Ticozzi. "Symmetrization for Quantum Networks: a Continuous-Time Approach." In: *International Symposium on Mathematical Theory of Network and Systems* (2014), pp. 1685–1690.

[49] A. Sarlette and M. Mirrahimi. "Loss-tolerant parity measurement for distant quantum bits". In: *Physical Review A* 95.3 (2017), p. 032329.

[50] A. Sarlette and P. Rouchon. "Deterministic submanifolds and analytic solution of the quantum stochastic differential master equation describing a monitored qubit". In: *Journal of Mathematical Physics* 58.6 (2017), p. 062106.

[51] A. Sarlette et al. "Stabilization of nonclassical states of one-and two-mode radiation fields by reservoir engineering". In: *Physical Review A* 86.1 (2012), p. 012114.

[52] A. Sarlette et al. "Stabilization of Nonclassical States of the Radiation Field in a Cavity by Reservoir Engineering". In: *Physical Review Letters* 107.1 (2011), p. 10402.

[53] (a) S.Muthukrishnan, B. Ghosh, and M.H. Schultz. "First- and second-order diffusive methods for rapid, coarse, distributed load balancing". In: *Theory of Computing Systems* 31 (1998), pp. 331–354; (b) B. Johansson and M. Johansson. "Faster linear iterations for distributed averaging". In: *Proc. IFAC World Congress* (2008), pp. 2861–2866; (c) E. Ghadimi, I. Shames, and M. Johansson. "Accelerated gradient methods for networked optimization". In: *Proc. American Control Conference* (2011); (d) J. Liu et al. "Analysis of accelerated gossip algorithms". In: *Automatica* 49 (2013), pp. 873–883.

[54] R.D. Somma et al. "Quantum simulations of classical annealing processes". In: *Physical Review Letters* 101.13 (2008), p. 130504.

[55] Daniel A. Spielman and Shang-Hua Teng. "A local clustering algorithm for massive graphs and its application to nearly linear time graph partitioning". In: *SIAM Journal on Computing* 42.1 (2013), pp. 1–26.

[56] (a) DW. Stroock and SRS. Varadhan. "On the support of diffusion processes with applications to the strong maximum principle". In: *Proc. 6th Berkeley Symposium on Mathematical Statistics and Probability* 3 (1972), pp. 333–359; (b) M. Chaleyat-Maurel and D.Michel. "The support of the density of a filter in the uncorrelated case". In: *Stochastic Partial Differential Equations and Applications II, Springer Lecture Notes in Mathematics* 1390 (1989), pp. 33–41.

[57] S. Stüdli, M.M. Seron, and R.H. Middleton. "From vehicular platoons to general networked systems: String stability and related concepts". In: *Annual Reviews in Control* 44 (2017), pp. 157–172.

[58] S. Sundaram and C. Hadjicostis. "Finite-time distributed consensus in graphs with time-invariant topologies". In: *Proc. American Control Conference* (2007), pp. 711–716.

[59] D. Swaroop and J. Hedrick. "String stability of interconnected systems". In: *IEEE Transactions on Automatic Control* 41 (1996), pp. 349–357.

[60] Mario Szegedy. "Quantum speed-up of Markov chain based algorithms". In: *Foundations of Computer Science, 2004. Proceedings. 45th Annual IEEE Symposium on.* IEEE. 2004, pp. 32–41.

[61] Mario Szegedy. "Quantum speed-up of markov chain based algorithms". In: *Proc. Annual IEEE Symposium on Foundations of Computer Science* (2004), pp. 32–41.

[62] (a) E. Tang. "A quantum-inspired classical algorithm for recommendation systems". In: *arXiv:1807.04271* (2018); (b) E. Tang. "Quantum-inspired classical algorithms for principal component analysis and supervised clustering". In: *arXiv:1811.00414* (2018).

[63] Kristan Temme et al. "The χ 2-divergence and mixing times of quantum Markov processes". In: *Journal of Mathematical Physics* 51.12 (2010), p. 122201.

[64] L. Viola, E. Knill, and S. Lloyd. "Dynamical decoupling of open quantum systems". In: *Physical Review Letters* 82.12 (1999), p. 2417.

[65] HM. Wiseman and GJ. Milburn. "Quantum theory of optical feedback via homodyne detection". In: *Physical Review Letters* 70 (1993), p. 548.

[66] Z. Zhang, Z. Ling, and A. Sarlette. "Integral control on nonlinear spaces: two extensions". In: *Proc. IFAC Conference on Nonlinear Control Systems (NOLCOS)* 49.18 (2016), pp. 1000–1005.

[67] Z. Zhang, Z. Ling, and A. Sarlette. "Modified Integral Control Globally Counters Symmetry-Breaking Biases". In: *Symmetry* 11.5 (2019), p. 639.

[68] Z. Zhang, A. Sarlette, and Z. Ling. "Integral control on Lie groups". In: *Systems & Control Letters* 80 (2015), pp. 9–15.