Spectral Bounds on the Convergence of Classical and Quantum Markov Chains

Oleg Szehr,[1] David Reeb,[1] and Michael M. Wolf[1]

[1] *Zentrum Mathematik, Technische Universität München, 85748 Garching, Germany*

(Dated: October 5, 2012)

We introduce a new framework that yields spectral bounds on norms of functions of transition maps for finite, homogeneous Markov chains. The employed techniques work for classical as well as for quantum Markov chains and they do not require additional assumptions like detailed balance, irreducibility or aperiodicity. We use the method in order to derive convergence bounds which improve significantly upon the known spectral bounds. The core technical observation is that power-boundedness of transition maps of Markov chains enables a Wiener algebra functional calculus in order to upper bound any norm of any holomorphic function of the transition map.

I. INTRODUCTION

Across scientific disciplines, Markov chains are ubiquitous in algorithms as well as in models for time evolutions. In many cases one is interested in when their limit behaviour is setting in. For algorithms this is often necessary in order to extract the right information and for time evolutions of physical systems this is the time scale on which relaxation or equilibration takes place. Some of the most widespread tools for bounding this time scale are based on the spectrum of the transition map. For time-homogeneous Markov chains with finite state space, the transition map is a stochastic matrix in the classical context and a completely positive trace-preserving map in the quantum case. Since these maps have spectral radius equal to 1, it is somehow clear that only eigenvalues of magnitude one survive the limit, that the largest subdominant eigenvalue governs the speed of convergence, and that the rest of the spectrum only matters on shorter time scales. When making this mathematically rigorous, one typically employs additional properties of the Markov chain such as detailed balance, irreducibility, aperiodicity, uniqueness of the fixed point, Gibbs distribution of the stationary state, etc.. Clearly, these assumptions are not always fulfilled—in particular in the quantum context, especially in the framework of dissipative quantum computing [11] and dissipative state preparation [2, 3, 11], they typically fail to hold.

The goal of the present work is to derive non-trivial upper bounds which depend exclusively on the spectrum of the transition map. Such bounds are of general interest for the theory of Markov chains, especially they are related to the sensitivity of the chain to perturbations [4, 5, 10], are used to study "cut-off" phenomena [1] and random walks on groups [9]. Before coming to our approach, a brief discussion of two more traditional, linear algebraic, approaches is in order. If $\mathcal{T}$ and $\mathcal{T}_\infty$ are the transition map and its asymptotic part, respectively, then a Jordan decomposition of the difference of the two maps yields a bound of the form

$$\|\mathcal{T}^n - \mathcal{T}_\infty^n\| \leq C\mu^{n-k}n^k \tag{1}$$

after n time steps. Here μ is the magnitude of the largest eigenvalue inside the open unit disc and $k+1$ is the size of the largest corresponding Jordan block. C is constant w.r.t. n, but it depends on $\mathcal{T}$ as it is essentially the condition number of the similarity transformation appearing in the Jordan decomposition. Unfortunately, there is no a priori bound on this condition number. An alternative way is to use Schur's instead of Jordan's normal form. This leads indeed to a similar expression as in Eq.(1) where C can be bounded independent of $\mathcal{T}$, albeit not of n, and we obtain something like $C \sim (Dn)^D$, where D is the dimension of the underlying vector space. Needless to say, this "constant" seems to be far from optimal and motivates a more elaborate analysis and the use of a new toolbox.

The method which we employ appears to be new to the theory of Markov processes and it enables us in principle to derive spectral bounds on norms of arbitrary functions of transition maps. When applied to power functions, we basically obtain the sought convergence bounds.

II. THE METHOD

We consider the following basic task. Given an arbitrary norm $\|\cdot\|$ and some sufficiently regular, say holomorphic, function f, obtain an upper bound for $\|f(\mathcal{T})\|$ as a function of the spectrum of $\mathcal{T}$.

The first, simple but crucial observation on this way is that transition maps are *power bounded operators*, meaning that there is a $C \in \mathbb{R}$ so that for all $\mathcal{T}$ and $n \in \mathbb{N}$ we have $\|\mathcal{T}^n\| \leq C$. One way to see this is that by the Russo-Dye theorem it holds true for the diamond norm $\|\mathcal{T}^n\|_\diamond = 1$ and thus, by the equivalence of norms in finite dimensional Banach spaces, also for any other norm for some $C \in \mathbb{R}$. General power-bounded operators are extensively studied in the mathematical literature [6–8] but, to the best of our knowledge, this hasn't found its way into the analysis of convergence properties of Markov chains.

The power boundedness condition can now be exploited in order shift our problem from spaces of operators to function spaces, which then offer a plethora of powerful tools to conduct the analysis. The basic idea is the following: instead of trying to estimate $\|f(\mathcal{T})\|$ directly, we consider a normed space of holomorphic functions, $(A, \|\cdot\|_A)$ and a map $\mathcal{J}_\mathcal{T}$ that relates the original Banach space of transition maps $(\mathfrak{T}, \|\cdot\|)$ to $(A, \|\cdot\|_A)$. The map $\mathcal{J}_\mathcal{T} : A \to \mathfrak{T}$ with $f \mapsto f(\mathcal{T})$ then corresponds to plugging the operator $\mathcal{T}$ into the function f. Suppose that for all transition maps $\|\mathcal{J}_\mathcal{T}\| := \sup_{f\in A} \frac{\|\mathcal{J}_\mathcal{T}(f)\|}{\|f\|_A} \leq C$ holds. It follows immediately that $\|f(\mathcal{T})\| \leq C\,\|f\|_A$. In the mathematical literature A is referred to as a *functional algebra* and $\mathcal{J}_\mathcal{T}$ as a *functional calculus.*

In our context, it is natural to start with the *Wiener algebra* W of absolutely convergent holomorphic functions on the open unit disc,

$$W := \Big\{ f = \sum_{k\geq 0} \hat{f}(k) z^k \mid \|f\|_W = \sum_{k\geq 0} |\hat{f}(k)| < \infty \Big\},$$

where the $\hat{f}(k)$'s are now the Taylor coefficient of the holomorphic function f. The basic reason for this is that for any function $f \in W$ we can apply the triangle inequality and bound

$$\|f(\mathcal{T})\| \leq \sum_{k\geq 0} |\hat{f}(k)| \left\|\mathcal{T}^k\right\| \leq \sum_{k\geq 0} C|\hat{f}(k)| = C\,\|f\|_W\,. \tag{2}$$

At first glance this appears to be of little use since the right hand side no longer depends on $\mathcal{T}$, if we invoke standard functional calculus. We can, however, tailor $\mathcal{J}_\mathcal{T}$ to $\mathcal{T}$ and exploit spectral properties of $\mathcal{T}$ to significantly strengthen the inequality (2). To see this, let $\{\lambda_i\}$ be the spectrum of $\mathcal{T}$ and consider natural numbers k_i and a polynomial $p_\mathcal{T}(z) = \prod_i (z-\lambda_i)^{k_i}$ which by construction has the property that $p_\mathcal{T}(\mathcal{T}) = 0$. Classical examples of polynomials that enjoy this property are, by the Cayley-Hamilton theorem, the characteristic polynomial and the *minimal polynomial* $m_\mathcal{T}$. Since by definition $m_\mathcal{T}$ is the minimal degree polynomial with $m_\mathcal{T}(\mathcal{T}) = 0$, the most general case is to consider $m_\mathcal{T} g$ with $g \in W$ arbitrary. For any $f, g \in W$ we have then that $\|(f + m_\mathcal{T} g)(\mathcal{T})\| = \|f(\mathcal{T}) + m_\mathcal{T}(\mathcal{T}) g(\mathcal{T})\| = \|f(\mathcal{T})\|$ and an application of (2) reveals that for all $g \in W$ we have $\|f(\mathcal{T})\| \leq \|f + m_\mathcal{T} g\|_W$. This leads us to the following simple but crucial observation:

Lemma 1. *Classical and quantum Markovian maps obey a Wiener algebra functional calculus: Let* $\|\cdot\|$ *be any norm such that for every transition map* $\mathcal{T} \in \mathfrak{T}$ *we have that* $\|\mathcal{T}\| \leq C$. *Then*

$$\|f(\mathcal{T})\| \leq C \, \|f\|_{W/m_{\mathcal{T}}W}$$

holds for any function $f \in W$, *where* $\|f\|_{W/m_{\mathcal{T}}W} := \inf\{\|f + m_{\mathcal{T}}g\|_W \mid g \in W\}$.

By going to the quotient space $W/m_{\mathcal{T}}W$ this makes use of the entire spectrum of $\mathcal{T}$. Of course, all the technical work is still ahead and goes into computing or bounding $\|f\|_{W/m_{\mathcal{T}}W}$. To this end, however, an extensive literature exists [6–8] which also offers smart choice for g.

III. A PURELY SPECTRAL BOUND ON THE LIMIT BEHAVIOUR

We now use the above observation in order to derive a spectral bound on $\|\mathcal{T}^n - \mathcal{T}_\infty^n\|$. To this end, note that $\mathcal{T}^n - \mathcal{T}_\infty^n = (\mathcal{T} - \mathcal{T}_\infty)^n$ and that the latter is power bounded for instance with $C = 2$ for the diamond norm.

Theorem 2. *Let* $\mathcal{T} \in \mathfrak{T}$ *be the transition map of a classical or quantum Markov chain and let* $\mathcal{T}_\infty$ *be the map describing its limit behaviour. We write* $m = m_{\mathcal{T}-\mathcal{T}_\infty}$ *for the minimal polynomial,* $\sigma(\mathcal{T} - \mathcal{T}_\infty) = \{\lambda_1, ..., \lambda_D\}$ *for the spectrum and* $\mu = |\lambda_D|$ *for the spectral radius of* $\mathcal{T} - \mathcal{T}_\infty$. *Finally, let* $\|\cdot\|$ *be any norm such that* $\|\mathcal{T}\| \leq C$ *for all* $\mathcal{T} \in \mathfrak{T}$. *Then for* $n > \frac{\mu}{1-\mu}$ *we have*

$$\|\mathcal{T}^n - \mathcal{T}_\infty^n\| \;\leq\; 4Ce^2\sqrt{|m|}(|m|+1)\,\frac{\mu^n}{\left(1-(1+\frac{1}{n})\mu\right)^{3/2}}\,\mathcal{B}(m,n), \tag{3}$$

$$\mathcal{B}(m,n) \;:=\; \prod_{m/(z-\lambda_D)} \frac{1-(1+\frac{1}{n})\mu|\lambda_i|}{\mu - |\lambda_i| + \frac{\mu}{n}}. \tag{4}$$

Here, the product is taken over all i *such that the corresponding linear factor* $(z - \lambda_i)$ *occurs in a prime factorization of* $m/(z - \lambda_D)$ respecting *multiplicities.*

Note that the bound is, as expected, asymptotically dominated by the factor μ^n and that the dimension-dependent pre-factor is at most $\sim D^{3/2}$ (when $|m| = D$) and in this sense benign. The (inverse) *Blaschke produkt* $\mathcal{B}$ takes care of possibly appearing Jordan blocks or perturbations thereof. There is evidence, albeit at this point on more heuristic footing, that for purely spectral bounds, there is not much room for improvement. A closer analysis, together with a detailed proof, can be found in the technical part of the submission.

Needles to say, there are only few cases where the entire spectrum of the transition map is known. These cases include random walks on groups [9]. However, even if the information about the spectrum is not very detailed, one can again bound the above expression and one still arrives at an improvement on the bound based on the Schur decomposition. Admittedly, if additional assumptions are fulfilled, in particular if detailed balance holds, the chain is irreducible, aperiodic *and* has a Gibbsian stationary state, then, the traditional bounds which are enabled by these properties, typically outperform the above bound if the same μ is chosen—these properties (if valid) can provably not be inferred from the spectrum alone.

---

[1] P. Diaconis. The cutoff phenomenon in finite markov chains. *Proc. Natl. Acad. Sci.*, 93:1659–1664, 1996.

[2] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. Büchler, and P. Zoller. Quantum states and phases in driven open quantum systems with cold atoms. *Nature Physics*, 4:878–883, 2008.

[3] B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller. Preparation of entangled states by quantum Markov processes. *Phys. Rev. A*, 78:042307, Oct 2008.

[4] A. Mitrophanv. Stability and exponential convergence of continuous-time markov chains. *J. Appl. Prob.*, 40:970–979, 2003.

[5] A. Mitrophanv. Sensitivity and convergence of uniformly ergodic markov chains. *J. Appl. Prob.*, 42:1003–1014, 2005.

[6] N. Nikolski. *Treatise on the Shift Operator*. Springer: Grundlehren der mathematischen Wissenschaft, 1986.

[7] N. Nikolski. Condition numbers of large matrices and analytic capacities. *St. Petersburg Math. J.*, 17:641–682, 2006.

[8] N. Nikolski. *Operators, Functions and Systems: An Easy Reading.* AMS: Mathematical Surveys and Monographs: 93, 2009.

[9] J. Rosenthal. Convergence rates of Markov chains. *SIAM*, 37(1995):387–405, 1995.

[10] O. Szehr and M. Wolf. Perturbation bounds for quantum Markov processes and their fixed points . 2011. arXiv: 1210.1171.

[11] F. Versraete, M. Wolf, and I. Cirac. Quantum computation and quantum-state engineering driven by dissipation. *Nat. Phys.*, 5 No.9:633–636, 2009.