Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Stability of Quantized Control Systems under Dynamic Bit Assignment

Qiang Ling and Michael D. Lemmon

***Abstract*— This paper studies the stability of a quantized feedback control system. In recent years, there have been a number of papers characterizing the minimum information rate required to assure closed loop stability. An inherent constraint on these quantized control systems is that the number of quantization levels must be an integer. In our opinion, insufficient attention has been paid to this constraint. This paper takes this integral constraints into account to establish a lower bound on the number of quantization levels required for closed loop stability. We then introduce a novel dynamic bit assignment policy that achieves this bound.**

I. Introduction

In recent years there has been a considerable amount of work studying the stability of feedback control systems under quantized feedback [1] [2] [3] [4] [5] [6] [7] [8] [9] [10]. These papers study the following linear discrete time system

$$\left\{ \begin{array}{rcl} x[k+1] & = & Ax[k] + Bu[k] \\ u[k] & = & Kx^q[k] \end{array} \right. \quad (1)$$

where (A, B) is controllable, $x[k] \in R^N$ is the state, $u[k]$ is the control, $x^q[k]$ is the quantized version of $x[k]$. The eigenvalues of A are denoted as λ_i $(i = 1, \cdots, N)$.

Among the aforementioned quantization papers, [10] considered stochastic moment stability, i.e. some moment of x_k converges to 0. The other papers study deterministic stability, i.e. x_k converges to 0 in the sense of a certain norm, for example the 2-norm. This paper focuses on deterministic stability. The papers on deterministic stability can be further classified into two groups based on their assumptions about the quantization range; static and dynamic quantization policies.

Static quantization policies [1] [4] presume that the quantization policy for the data is fixed for all time. In [1], it is proven that a finite number of quantization levels cannot achieve aymptotic stability. An infinite number of quantization levels is studied in [4] where it is proven that the least dense quantizer achieving asymptotic stability has a logarithmic distribution. In [4], finite quantization levels are shown to be able to achieve practical stability, i.e. the states will converge into a bounded set.

Dynamic quantization policies [5] [6] [8] [9] [7] that $x[k]$ presume that the state at time instant k, lies inside a predetermined set $U[k]$ called the *uncertainty set*. If $U[k]$ converges to 0, i.e. every point in $U[k]$ converges to 0, then the traditional stability of the system in eq. 1 can be guaranteed. $U[k]$ is usually assumed to be rectangular. Many different approaches have been proposed to guarantee the convergence of $U[k]$. These approaches are reviewed below.

In [5] [6] every side of the uncertainty set, $U[k]$, is equally partitioned into M parts, i.e. $U[k]$ is partitioned into N^M small rectangles. Denote the small rectangles as $U_i[k]$ $(i = 0, 1, \cdots, N^M - 1)$. Suppose $x[k] \in U_j[k]$. The index j will be sent. $x^q[k]$ is estimated by the center of $U_j[k]$. $U[k+1]$ is a rectangle bounding $AU_j[k] + Bu[k]$, where $AU_j[k]$ means $AU_j[k] = \{z | z = Ay, y \in U_j[k]\}$. Sufficent conditions for $U[k]$ to converge to 0 are provided.

In [8] [9], a framework similar to that used in [5] is followed. This work may be distinguished from [5] in the choice of quantization policy. In particular, the quantization policy in [9] has the i^{th} side of $U[k]$ equally partitioned into 2^{R_i} parts, i.e. R_i bits are assigned to the i^{th} dimension. It is required that $2^{R_i} > |\lambda_i|$. The total number of bits is $R = \sum_{i=1}^{N} R_i$, i.e. there are $Q = 2^R$ quantization levels. It is asserted in [9] that the system in eq. 1 is asymptotically stable if and only if

$$2^R > \prod_{i=1}^{N} max(1, |\lambda_i|) \quad (2)$$

The lower bound in eq. 2 is derived under the assumption that R_i can be any real number. *This assumption is not meaningful because R_i should be an integer.* A similar problem concerning the assumption of non-integral quantization also exists in [7].

This paper examines the stability of quantized feedback control systems in which the 2^{R_i} are required to be integers. Following the derivation in [9], we can obtain the following bound

$$2^R \geq \prod_{i=1}^{N} max(1, \lceil |\lambda_i| \rceil) \quad (3)$$

where $\lceil \cdot \rceil$ means $\lceil x \rceil = min\{n | n > x, n \in \mathcal{N}\}$. This is similar to the inequality in equation 2, but it is not the greatest lower bound. This paper shows that

$$2^R \geq \left\lceil \prod_{i=1}^{N} \max(1, |\lambda_i|) \right\rceil \quad (4)$$

provides a tighter lower bound than that in equation 3. We then go on to introduce a **dynamic bit assignment policy**

The authors are with Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556; *Email:* qling,lemmon@nd.edu; *Phone:* 574-631-8309; *Fax:* 574-631-4393. The authors gratefully acknowledge the partial financial support of the Army Research Office (DAAG19-01-0743), the National Science Foundation (NSF-CCR02-8537, NSF-ECS02-25265)

that actually achieves this bound. This bit assignment is done as follows. Suppose there are $Q = 2^R$ quantization levels and Q is integer. At every step, only the "longest" (in a certain sense) side of $U[k]$ is equally partitioned into Q parts; the other sides aren't partitioned. Because no side is always the longest, the bit assigning is dynamic rather than static. The main contribution of this paper is to prove that this policy achieves the lower bound in equation 4.

This paper is organized as follows. Section II formally defines the quantized feedback control system, introduces some notation, and lists the frequently used technical lemmas. Section III states and proves the main result conerning the lower bound on the minimum number of quantization levels stabilizing the system. This section also introduces the dynamic bit assignment policy. The proof that this policy achieves the minimum number of quantization levels will be found in the appendix, section V.

II. Quantized Feedback Control System

This paper focuses on a quantized feedback control system with dropouts, which is shown in figure 1. The plant

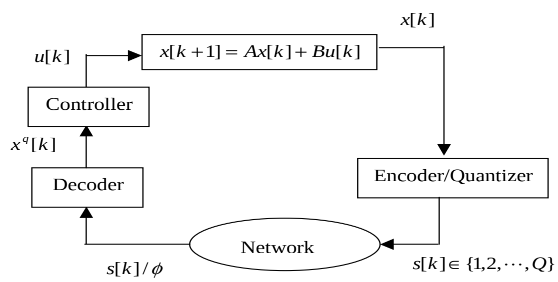

Fig. 1. Quantized networked control system

is a discrete-time linear system whose equations were given in 1. The state $x[k] \in R^n$ is quantized and encoded into a symbol $s[k]$ from a discrete set $\{1, 2, \cdots, Q\}$. Throughout this paper, the terms "quantizer" and "encoder" are used interchangeably. It is important to mention that the coding policy may be time-varying, for example $s[k] = 1$ and $s[k+1] = 1$ may mean quite different signals. $s[k]$ is sent to the decoder over the network. Because the network is non-deterministic, a portion of the transmitted symbols may be dropped. The dropouts are denoted by receiving ϕ at the decoder end. The decoder uses the received symbols to estimate $x[k]$ with $x^q[k]$. The controller computes $u[k]$ with $x^q[k]$ as shown in eq. 1.

We are interested in the following notion of deterministic stability,

$$\lim_{k\to\infty} \|x[k]\|_2 = 0, \forall x[0] \in R^N. \tag{5}$$

where $\|\cdot\|_2$ denotes the Euclidean 2-norm.

A. Assumptions

The following assumptions will be made in this paper. A more detailed discussion justifying these assumptions will be provided below.

1) (A, B) is controllable. $A = diag(J_1, J_2, \cdots, J_p)$ where J_i is an $n_i \times n_i$ real matrix with a single real eigenvalue λ_i or a pair of conjugate eigenvalues λ_i and λ_i^*. All eigenvalues λ_i are assumed to be unstable, i.e. $|\lambda_i| > 1$.
2) The initial condition $x[0]$ lies in a super-parallelogram $P[0]$.
3) Transmitted symbols, $s[k]$, are dropped at the rate of ε symbols per transmission. The precise definition of ε will be found in equation 7. We assume that the encoder and decoder both know whether a dropout has occurred.
4) Both the encoder and the decoder know the system matrices (A and B), the coding-decoding policy and the control law. They also agree upon the initial uncertainty set, i.e. the super-parallelogram which $x[0]$ lies in.

We take the matrix, A, (assumption 1) to be in its real Jordan canonical form [11]. Since any system may be reduced to this form through a similarity transformation, we may therefore assume $A = diag(J_1, J_2, \cdots, J_p)$. When $|\lambda_i| < 1$, the subsystem corresponding to J_i is stable. We can exclude the stable subsystem and consider only the lower dimensional system. This paper therefore assumes that $|\lambda_i| > 1$ $(i = 1, \cdots, p)$.

Assumption 2 requires that the initial state is known to lie within a specified superparallelogram $P[0]$. This set may be written as

$$P[0] = x^q[0] + U[0]$$

where $x^q[0]$ is the center of $P[0]$ and $U[0]$ is a super-parallelogram centered at the origin and defined in equations 9-10.

Assumption 3 comes from the non-determinism of the network. We introduce a dropout indicator $d[k]$,

$$d[k] = \begin{cases} 1, & \text{the symbol at time } k \text{ is dropped} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

We assume that the dropout model satisfies

$$\varepsilon = \lim_{L\to\infty} \frac{\sum_{i=1}^{L} d[i + k_0]}{L}, \text{ for } \forall k_0 \geq 0 \tag{7}$$

where ε is the "average" dropout rate and the convergence in eq. 7 is uniform with respect to k_0.

Assumption 4 requires that the coder and the decoder deal with the same initial uncertainty, and share the same coding-decoding policy and control law so that the symbol produced by the encoder can be correctly interpreted by the decoder. This is a strong assumption for it requires that the encoder and decoder are "synchronized". Maintaining such synchronization in a fault-tolerant manner requires further study, but that study is not done in this paper.

B. Mathematical Preliminaries

This subsection defines some notation and lists frequently used lemmas. Most of lemmas are well known, so their proofs are omitted.

For the matrix A in assumption 1, define

$$\rho(A) = \prod_{i=1}^{p} (max(1, |\lambda_i|))^{n_i} \quad (8)$$

We assume all eigenvalues of A are unstable. So $\rho(A) = |det(A)|$, where $det(\cdot)$ is the determinant of a matrix.

This paper quantizes the state $x[k]$ at time k with respect to a parallelogram representing the quantization "uncertainty". These uncertainty sets are generally represented as

$$P[k] = x^q[k] + U[k]$$

where $x^q[k] \in \Re^n$ is the center of $P[k]$ and $U[k]$ is a super-parallelogram with its center at the origin. The super-parallelogram $U[k]$ is formally represented by a set of vectors $\mathcal{V}[k] = \{v_{i,j}[k] \in \Re^{n_i}\}$ where $i = 1, \dots, p$ and $j = 1, \dots, n_i$. The "side" of the parallelogram associated with the ith Jordan block in A is denoted as the convex hull

$$S_i[k] = \mathrm{Co}\left\{ v \ : \ v = \sum_{j=1}^{n_i} (\pm\frac{1}{2}) v_{i,j}[k] \right\} \quad (9)$$

The entire super-parallelogram, $U[k]$, may therefore be expressed as the Cartesian product of the sides, S_i. In other words

$$U[k] = \bigotimes_{i=1}^{P} S_i[k] \quad (10)$$

U is a measurable set in Euclidean N-space, R^N. Its volume is defined as

$$vol(U) = \int_{x \in U} 1 \cdot dx \quad (11)$$

We define the diameter of U as

$$d_{max}(U) = sup_{x,y \in U} \|x - y\|_2 \quad (12)$$

where $\|\cdot\|_2$ denotes Euclidean 2-norm of a vector.

We define the operation of a matrix T on a set U by

$$T : U \rightarrow \{y : \exists x \in U \ such \ that \ y = Tx\} \quad (13)$$

By the above definition, T can be viewed as a linear mapping. We denote the image set of U under the operation of T as TU. TU may be shown to have the following properties.

Lemma 2.1: $vol(TU) = |det(T)| vol(U)$.

Lemma 2.2: When $T = diag(J_1, J_2, \cdots, J_p)$ (J_i is defined in assumption 1) and $U = c + \bigotimes_{i=1}^{p} S_i$ where S_i is given in equation 9, then

$$TU = c' + \bigotimes_{i=1}^{p} S_i' \quad (14)$$

where $c' = Tc$, and

$$S_i' = \mathrm{Co}\left\{ v \ : \ v = \sum_{j=1}^{n_i} (\pm\frac{1}{2}) J_i v_{i,j} \right\}$$

For J_i defined in assumption 1, we have the following growth rate lemma.

Lemma 2.3: For any non-zero $v_i \in R^{n_i}$,

$$\begin{aligned} \lim_{k\to\infty} \|J_i^k v_i\|_2 &= \infty \\ \lim_{k\to\infty} \frac{\|J_i^{k+1} v_i\|_2}{\|J_i^k v_i\|_2} &= |\lambda_i| \end{aligned}$$

The quantization error of the feedback control system (Eq. 1) is

$$e[k] = x[k] - x^q[k] \quad (15)$$

Denote the set which $e[k]$ lies in as $U[k]$, where $U[k]$ is a set with the origin as its center. $U[k]$ can be measured by its diameter $d_{max}(U[k])$. The relationship between the stability of the control system and $d_{max}(U[k])$ is established by lemma 2.4.

Lemma 2.4: The system in eq. 1 is asymptotically stable if and only if

$$\lim_{k\to\infty} d_{max}(U[k]) = 0. \quad (16)$$

Lemma 2.4 can be proven similarly as Lemma 3.5.1 in [9]. The proof is omitted here. By Lemma 2.4, we focus on the convergence of quantization error, which will guarantee the asymptotic stability of the quantized control system.

III. Main results

This section states a lower bound (theorem 3.1) on the number of quantization levels required to stabilize the feedback control system. The proof of theorem 3.1 follows the approach in [7], so we only sketch the proof. We then present the *dynamic bit assignment policy* (algorithm 3.1) and state a theorem (theorem 3.2) asserting that the lower bound is achieved by our bit assignment policy. The proof of theorem 3.2 will be found in the appendix, section V.

Theorem 3.1: Under assumptions 1 - 4, if the quantized feedback system in eq. 1 can be asymptotically stabilized, then the number of quantization levels, Q, satisfies

$$Q \geq Q_{min} = \left\lceil \rho(A)^{\frac{1}{1-\varepsilon}} \right\rceil \quad (17)$$

Sketch of Proof: The proof is based on the analysis of volume evolution of uncertainty sets and is similar to the proof of Theorem 1 in [7].

The volume of $U[k]$ is updated by

$$vol(U[k+1]) \left\{ \begin{array}{ll} \geq \frac{|det(A)|}{Q} vol(U[k]), & d[k] = 0 \\ = |det(A)| vol(U[k]), & d[k] = 1 \end{array} \right.$$

Because of asymptotic stability, we know, by lemma 2.4, $\lim_{k\to\infty} vol(U[k]) = 0$. This volume limit, together with the dropout rate of ε, yields

$$\frac{|det(A)|}{Q^{1-\varepsilon}} < 1 \quad (18)$$

Because $\rho(A) = |det(A)|$ and Q is an integer, we obtain the lower bound in eq. 17. $\diamondsuit$

The following algorithm dynamically quantizes the state $x[k]$ for the feedback system in eq. 1 under assumptions 1-

4. The algorithm updates a parallelogram, $P[k]$ containing the state at time k. This parallelogram, $P[k]$, is characterized by, $x^q[k]$, the center of the parallelogram, and $U[k]$, the uncertainy set. The uncertainty set $U[k]$ is formed from a set of vectors $\{v_{i,j}[k] \in \Re^{n_i}\}$ ($i = 1, \ldots, p$, $j = 1, \ldots, n_i$) according to equation 9-10. The uncertainty set $U^{(I,J)}[k]$ is a modification of $U[k]$ that is formed from the vectors $\{v'_{i,j}[k]\}$ where $v'_{i,j} = v_{i,j}$ if $(i,j) \neq (I,J)$ and $v'_{i,j} = v_{i,j}/Q$ if $(i,j) = (I,J)$. The basic variables updated by this algorithm are therefore the collection of vectors $\{v_{i,j}[k]\}$ and $x^q[k]$, a one-step ahead prediction of the quantized state at time k given observations up to and including time $k-1$. The quantized signal that is sent between the encoder and decoder at time k is denoted at $s[k]$. This quantized signal is equal to one of Q discrete symbols. The following algorithm description describes two tasks that are executed concurrently, the *encoder* and *decoder* algorithms. Each task's first step starts its execution at the same time instant.

Algorithm 3.1: **Dynamic Bit Assignment:**

Encoder/Decoder initialization:

Initialize $x^q[0]$ and $\{v_{i,j}[0]\}$ so that $x[0] \in x^q[0] + U[0]$ and set $k = 0$.

Encoder Algorithm:

1) **Select** the indices (I, J) that are associated with the longest vector $v_{i,j}[k]$. In other words,

$$(I, J) = \arg\max_{i,j} \|J_i v_{i,j}[k]\|_2$$

2) **Quantize** the state $x[k]$ by setting $s[k] = s$ if and only if

$$x[k] \in x^q[k] + x_s^{(I,J)} + U^{(I,J)}[k]$$

where

$$x_s^{(I,J)} = \begin{bmatrix} 0 & \cdots & 0 & v^T & 0 & \cdots & 0 \end{bmatrix}^T \quad (19)$$

and $v = \frac{-Q+(2s-1)}{2Q} v_{I,J}[k]$ for $s = 1, \ldots, Q$.

3) **Transmit** the quantized symbol $s[k]$ and wait for acknowledgement

4) **Update** the variables

$$\begin{aligned} v_{i,j}[k+1] &= J_i v_{i,j}[k] \\ x^q[k+1] &= (A + BK) x^q[k] \end{aligned}$$

5) **If decoder ack received:**

$$\begin{aligned} v_{I,J}[k+1] &= \frac{1}{Q} v_{I,J}[k+1] \\ x^q[k+1] &= x^q[k+1] + A x_{s[k]}^{(I,J)} \end{aligned}$$

where $x_{s[k]}^{(I,J)}$ is defined in equation 19.

6) Update time, $k = k + 1$ and return to step 1.

Decoder Algorithm:

1) **Update** the variables

$$\begin{aligned} v_{i,j}[k+1] &= J_i v_{i,j}[k] \\ x^q[k+1] &= (A + BK) x^q[k] \end{aligned}$$

2) **Wait** for quantized data, $s[k]$, from encoder.

3) **If data received:**

$$\begin{aligned} v_{I,J}[k+1] &= \frac{1}{Q} v_{I,J}[k+1] \\ x^q[k+1] &= x^q[k+1] + A x_{s[k]}^{(I,J)} \end{aligned}$$

where $x_{s[k]}^{(I,J)}$ is defined in equation 19. Then send *ack* back to the encoder.

4) Update time index, $k = k + 1$, and return to step 1.

Remark: This algorithm assumes the variables $\{v_{i,j}[k]\}$ and $x^q[k]$ are "synchronized" at the beginning of the kth time interval. Furthermore, we assume the "ack" from decoder to the encoder is reliably transmitted.

Remark: The decision in step 1 of the encoder algorithm is made on the uncertainty set at time $k+1$, rather than k. This was motivated by preliminary studies that using the kth uncertainty set may perform poorly when some of the λ_i are large. This observation motivated the one-step ahead decision rule used here.

Theorem 3.2: Let

$$Q = \left\lceil \rho(A)^{\frac{1}{1-\varepsilon}} \right\rceil .$$

The feedback system in eq. 1 is asymptotically stable under the quantizer in algorithm 3.1. Furthermore $U[k]$ converges to zero according to

$$d_{max}(U[k]) \leq \lambda_0 \eta^k \quad (20)$$

where λ_0 is a constant, $\eta = \left(\frac{\rho(A)}{Q^{1-\varepsilon}}\right)^{\frac{1}{N}}$

In order to improve readability, we move the proof of theorem 3.2 to the appendix, section V.

Remark: Now we compare the bounds in eq. 3 and eq. 17 ($\varepsilon = 0$). It is the case that

$$\left\lceil \rho(A)^{\frac{1}{1-\varepsilon}} \right\rceil \leq \prod_{i=1}^{N} max(1, \lceil |\lambda_i| \rceil) \quad (21)$$

The strict inequality in eq. 21 usually holds. So the bound in eq. 17 is better than the bound in eq. 3. We take an example to show the difference. $A = \begin{bmatrix} 1.8 & 0 \\ 0 & 1.1 \end{bmatrix}$. The bound in eq. 3 is $Q \geq 4$. The bound in eq. 17 is $Q \geq 2$. So the latter bound is better. Now we try to give some intuitive explanation for the difference. The quantization policy in [9] seperately deals with two subsystems

$$\begin{aligned} x_1[k+1] &= 1.8 x_1[k] + b_1 u[k] \quad (22) \\ x_2[k+1] &= 1.1 x_2[k] + b_2 u[k] \quad (23) \end{aligned}$$

Every subsystem is unstable and needs at least 2 levels. Total number of quantization levels is at least $2 \times 2 = 4$, as shown by eq. 3.

It can be seen that although the two subsystems are unstable, they, especially the subsystem in eq. 23, are not too unstable. When we assign 2 quantization levels to every subsystem, there exists excess stability margin because $\frac{1.8}{2} < 1$ and $\frac{1.1}{2} < 1$. This paper's dynamic bit assignment policy considers the two subsystems as a whole. The stability margins can be combined together so that fewer quantization levels are needed. Figure 2 clearly shows that the 1-bit quantized system converges to zero as expected.

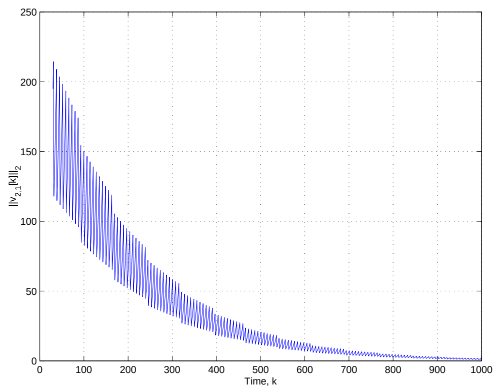

Fig. 2. Response of Quantized System

IV. Conclusions

This paper derived a lower bound on the minimum number of the quantization levels required to stabilize the closed loop system. We showed there exists a dynamic bit assignment policy that achieves this lower bound. The bound therefore appears to be the greatest lower bound on the stabilizing number of quantization levels.

References

[1] D. Delchamps, "Stabilizing a linear system with quantized state feedback," *IEEE Transactions on Automatic Control*, vol. 35(8), pp. 916–924, 1990.

[2] W. Wong and R. Brockett, "Systems with finite communication bandwidth constraints- part i: state estimation problem," *IEEE Transactions on Automatic Control*, vol. 42(9), pp. 1294–1299, 1997.

[3] ——, "Systems with finite communication bandwidth constraints-part ii: stabilization with limited information feedback," *IEEE Transactions on Automatic Control*, vol. 44(5), pp. 1049–1053, 1999.

[4] N. Elia and S. Mitter, "Stabilization of linear systems with limited information," *IEEE Transactions on Automatic Control*, vol. 46(9), pp. 1384 –1400, 2001.

[5] R. Brockett and D. Liberzon, "Quantized feedback stabilization of linear systems," *IEEE Transactions on Automatic Control*, vol. 45(7), pp. 1279 –1289, 2000.

[6] D. Liberzon, "On stabilization of linear systems with limited information," *IEEE Transactions on Automatic Control*, vol. 48(2), pp. 304–307, 2003.

[7] J. Hespanha, A. Ortega, and L. Vasudevan, "Towards the control of linear systems with minimum bit-rate," in *Proc. of the Int. Symp. on the Mathematical Theory of Networks and Systems*, 2002.

[8] S. Tatikonda and S. Mitter, "Control under communication constraints," in *the 38th Annual Allerton Conference on Communications and Control Computation*, 2000, pp. 182–190.

[9] S. Tatikonda, "Control under communication constraints," Ph.D. dissertation, M.I.T., 2000.

[10] G. Nair and R. Evans, "Exponential stabilisability of finite dimensional linear systems with limited data rates," *Automatica*, vol. 39, pp. 585–593, 2003.

[11] R. Horn and C. Johnson, *Matrix analysis*. Cambridge University Press, 1985.

V. Appendix

Define $p[k] = \prod_{i=1}^{P}\prod_{j=1}^{n_i}\|v_{i,j}[k]\|_2$. We first prove $\{p[k]\}$ is eventually bounded by an exponentially decreasing sequence. Based on that, we derive an eventual upper bound on $d_{max}(U[k])$ and show that $d_{max}(U[k])$ converges to 0.

Eventual bound on $p[k]$: Define the following average dropout rate

$$\bar{\varepsilon}_{l,k} = \frac{\sum_{i=0}^{l-1} d[k+i]}{l} \tag{24}$$

By the dropout model in eq. 7, we know for any small $\delta_0 > 0$, there exist $M > 0$ such that

$$\varepsilon - \delta_0 \le \bar{\varepsilon}_{l,k} \le \varepsilon + \delta_0, \forall l \ge M, \forall k \tag{25}$$

By algorithm 3.1, we know $v_{i,j}[k]$ is a scaled version of $J_i^k v_{i,j}[0]$. Lemma 2.3 guarantees for any $\varepsilon_0 > 0$, there exist K_1 such that when $k > K_1$,

$$(1-\varepsilon_0)|\lambda_i| \le \frac{\|J_i v_{i,j}[k]\|_2}{\|v_{i,j}[k]\|_2} \le (1+\varepsilon_0)|\lambda_i|, \forall i, j. \tag{26}$$

We will limit to $l > M$ and $k > K_1$. By the updating rule of $v_{i,j}[k]$, we know

$$\begin{aligned} p[k+l] &= \frac{1}{Q^{1-\bar{\varepsilon}_{l,k}}}\prod_{i=1}^{P}\prod_{j=1}^{n_i}\|J_i^l v_{i,j}[k]\|_2 \\ &\le \eta_{K_1,M}^l p[k] \end{aligned}$$

where $\eta_{K_1,M} = \frac{\rho(A)}{Q^{1-\varepsilon-\delta_0}}(1+\varepsilon_0)^N$. Because $\frac{\rho(A)}{Q^{1-\varepsilon}} < 1$ and ε_0 and δ_0 can be small enough by choosing large K_1 and M,

$$\eta_{K_1,M} < 1 \tag{27}$$

Therefore $p[k+ml]$ is bounded by

$$p[k+ml] \le p[k]\eta_{K_1,M}^{ml} \tag{28}$$

For $n = 1, 2, \cdots, l-1$, we can get the following bound

$$p[k+ml+n] \le \left(\rho(A)(1+\varepsilon_0)^N\right)^n p[k+ml] \tag{29}$$

Combining eq. 28 and 29 yields

$$p[k+l] \le p_0\eta_{K_1,M}^{k+l}, l \ge 0 \tag{30}$$

where p_0 is a constant. By choosing large enough p_0, the above inequality holds for any $k \ge 0$, i.e. $p[k] \le p_0\eta_{K_1,M}^k$. By the definition of $\eta_{K_1,M}$, we know

$$\lim_{K_1\to\infty, M\to\infty}\eta_{K_1,M} = \eta \tag{31}$$

So we rewrite eq. 30 into

$$p[k] \le p_0\eta^k \tag{32}$$

Eventual upper bound on $d_{max}(U[k])$: From eq. 32, we know $p[k]$ converges to 0. But that fact doesn't surely imply the convergence of $\|v_{i,j}[k]\|_2$. We prove that $\frac{\|v_{i_1,j_1}[k]\|_2}{\|v_{i_2,j_2}[k]\|_2}$ will be eventually bounded for any i_1, i_2, j_1, j_2, which yields the convergence of $\|v_{i,j}[k]\|_2$ and furthermore the convergence of $d_{max}(U[k])$.

By eq. 26, we obtain

$$\frac{\|v_{i_1,j_1}[k]\|_2}{\|v_{i_2,j_2}[k]\|_2} \leq \alpha \frac{\|J_{i_1} v_{i_1,j_1}[k]\|_2}{\|J_{i_2} v_{i_2,j_2}[k]\|_2} \tag{33}$$

where $\alpha = \frac{1+\varepsilon_0}{1-\varepsilon_0} \max_{i_1,i_2} \frac{|\lambda_{i_1}|}{|\lambda_{i_2}|}$. Thus the boundedness of $\frac{\|v_{i_1,j_1}[k]\|_2}{\|v_{i_2,j_2}[k]\|_2}$ can be guaranteed by that of $\frac{\|J_{i_1} v_{i_1,j_1}[k]\|_2}{\|J_{i_2} v_{i_2,j_2}[k]\|_2}$, which is pursued in the following.

Define $T_{i,j} = \{k : I_k = i, J_k = j\}$. $T_{i,j}$ is the collection of time instants for $\|J_i v_{i,j}[k]\|_2$ to be the largest. $T_{i,j}$ has infinite number of elements for any i, j (otherwise $\|v_{i,j}[k]\|_2$ will converge to ∞). We thus know there exist $K > K_1$ such that

$$count(T_{i,j} \cap \{K_1, K_1+1, \cdots, K\}) \geq M, \forall i, j. \tag{34}$$

where $count(\cdot)$ denotes the number of elements of a set. From now on, we will limit to $k > K$. At time instant k, suppose $(i_0, j_0) = argmin_{i,j} \|J_i v_{i,j}[k]\|_2$. Thus we can know

$$\|J_{i_0} v_{i_0,j_0}[k]\|_2 \leq \|J_i v_{i,j}[k]\|_2, \forall i, j \tag{35}$$

By eq. 34, we know the definition

$$\begin{aligned} \bar{l} \;=\; & min\,\{m : count(\{k-m, k-m+1, \cdots, k\} \\ & \cap T_{i_0,j_0}) = M\}\,, \end{aligned}$$

is meaningful. Therefore we get

$$\|J_{i_0} v_{i_0,j_0}[k-\bar{l}]\|_2 \geq \|J_i v_{i,j}[k-\bar{l}]\|_2, \forall i, j \tag{36}$$

By eq. 34, we know $k - \bar{l} \geq K_1$. So eq. 26 holds for the time instants after $k - \bar{l}$.

Now we study the evolution of $v_{i,j}[k]$ from $k - \bar{l}$ to k. By the updating rule of $v_{i,j}[k]$ and the definition of $\bar{l}$, we know

$$\left\{ \begin{array}{rcl} J_{i_0} v_{i_0,j_0}[k] & = & \frac{1}{Q^M} J_{i_0}^{\bar{l}} \left(J_{i_0} v_{i_0,j_0}[k-\bar{l}]\right) \\ J_i v_{i,j}[k] & = & \frac{1}{Q^{l_{i,j}}} J_i^{\bar{l}} \left(J_i v_{i,j}[k-\bar{l}]\right), \\ & & (i,j) \neq (i_0, j_0) \end{array} \right. \tag{37}$$

where $l_{i,j}((i,j) \neq (i_0, j_0))$ is governed by

$$\left\{ \begin{array}{lcl} \sum_{(i,j)\neq(i_0,j_0)} l_{i,j} & = & \bar{l} - M \\ & \geq & \bar{l}(1 - \varepsilon - \delta_0) - M \\ l_{i,j} & \geq & 0 \end{array} \right. \tag{38}$$

By eq. 26, 35 and 37, we obtain

$$\begin{aligned} & \frac{((1-\varepsilon_0)|\lambda_{i_0}|)^{\bar{l}}}{Q^M} \|J_{i_0} v_{i_0,j_0}[k-\bar{l}]\|_2 \\ \leq \;& \frac{((1+\varepsilon_0)|\lambda_i|)^{\bar{l}}}{Q^{l_{i,j}}} \|J_i v_{i,j}[k-\bar{l}]\|_2, \forall i, j \end{aligned} \tag{39}$$

Combining eq. 39 and 36 yields

$$((1-\varepsilon_0)|\lambda_{i_0}|)^{\bar{l}} \frac{1}{Q^M} \leq ((1+\varepsilon_0)|\lambda_i|)^{\bar{l}} \frac{1}{Q^{l_{i,j}}} \tag{40}$$

Multiplify the above inequality for all possible i, j ($(i,j) \neq (i_0, j_0)$). We obtain

$$\begin{aligned} & ((1-\varepsilon_0)|\lambda_{i_0}|)^{\bar{l}(N-1)} \frac{1}{Q^{(N-1)M}} \\ \leq \;& \left((1+\varepsilon_0)^{N-1} \prod_{(i,j)\neq(i_0,j_0)} |\lambda_i|\right)^{\bar{l}} \frac{1}{Q^{\sum_{(i,j)\neq(i_0,j_0)} l_{i,j}}} \\ \leq \;& \left((1+\varepsilon_0)^{N-1} \prod_{(i,j)\neq(i_0,j_0)} |\lambda_i|\right)^{\bar{l}} \frac{1}{Q^{\bar{l}(1-\varepsilon-\delta_0)-M}} \end{aligned}$$

Solve the above inequality with respect to $\bar{l}$, we will get

$$\bar{l} \leq l_{i_0} \tag{41}$$

where $l_{i_0} = \frac{MNln(Q)}{(N-1)ln\left(\frac{1-\varepsilon_0}{1+\varepsilon_0}\right) + Nln(|\lambda_{i_0}|) + ln\left(\frac{Q^{1-\varepsilon}}{\rho(A)}\right) - \delta_0 ln(Q)}$ (Because ε_0 and δ_0 can be arbitrarily small and $\frac{Q^{1-\varepsilon}}{\rho(A)} > 1$, we can guarantee $l_{i_0} \geq 0$). Define the maximum of l_{i_0} as $l_0 = \max_{i_0} l_{i_0}$. We know

$$\bar{l} \leq l_0 \tag{42}$$

Now we try to bound $\frac{\|J_i v_{i,j}[k]\|_2}{\|J_{i_0} v_{i_0,j_0}\|_2}$ for $(i,j) \neq (i_0, j_0)$.

$$\begin{aligned} \frac{\|J_i v_{i,j}[k]\|_2}{\|J_{i_0} v_{i_0,j_0}[k]\|_2} & \leq \frac{\frac{|\lambda_i|^{\bar{l}}}{Q^{l_{i,j}}} (1+\varepsilon_0)^{\bar{l}} \|J_i v_{i,j}[k-\bar{l}]\|_2}{\frac{|\lambda_{i_0}|^{\bar{l}}}{Q^M} (1-\varepsilon_0)^{\bar{l}} \|J_{i_0} v_{i_0,j_0}[k-\bar{l}]\|_2} \\ & \leq Q^M \left(\frac{|\lambda_i|}{|\lambda_{i_0}|}\right)^{\bar{l}} \left(\frac{1+\varepsilon_0}{1-\varepsilon_0}\right)^{\bar{l}} \\ & \leq r_0 \end{aligned}$$

where $r_0 = Q^M \left(\max_{i_1,i_2} \frac{|\lambda_{i_1}|}{|\lambda_{i_2}|}\right)^{l_0} \left(\frac{1+\varepsilon_0}{1-\varepsilon_0}\right)^{l_0}$. Eq. 36 is used to reach the above second inequality. By eq. 35, we know $\|J_{i_0} v_{i_0,j_0}[k]\|_2$ is the smallest among $\|J_{i_1} v_{i_1,j_1}[k]\|_2$. So

$$\frac{\|J_{i_1} v_{i_1,j_1}[k]\|_2}{\|J_{i_2} v_{i_2,j_2}[k]\|_2} \leq r_0, \forall i_1, i_2, j_1, j_2 \tag{43}$$

By eq. 33 and 43, we obtain

$$\frac{\max_{m,n} \|v_{m,n}[k]\|_2}{\|v_{i,j}[k]\|_2} \leq r \tag{44}$$

where $r = r_0 \alpha$. So

$$\|v_{i,j}[k]\|_2 \geq \frac{1}{r} \max_{m,n} \|v_{m,n}[k]\|_2 \tag{45}$$

By the above relationship, together with the definition of $p[k]$ and the bound in eq. 32, we know

$$\max_{m,n} \|v_{m,n}[k]\|_2 \leq r \sqrt[N]{p_0} \eta^{\frac{k}{N}} \tag{46}$$

Because $d_{max}(U[k]) \leq N \max_{m,n} \|v_{m,n}[k]\|_2$, we know

$$d_{max}(U[k]) \leq \lambda_0 \eta^{\frac{k}{N}} \tag{47}$$

where $\lambda_0 = Nr \sqrt[N]{p_0}$. Although we get eq. 47 for $k > K$, it can be extended to $k = 1, 2, \cdots$ by choosing a large enough λ_0. $\diamond$