

1-1-1991

# The Expected Capacity of Concentrators

Nicholas Pippenger  
*Harvey Mudd College*

---

## Recommended Citation

Nicholas Pippenger. "The Expected Capacity of Concentrators", *Society for Industrial and Applied Mathematics Journal of Discrete Mathematics*, 4, 121 (1991).

This Article is brought to you for free and open access by the HMC Faculty Scholarship at Scholarship @ Claremont. It has been accepted for inclusion in All HMC Faculty Publications and Research by an authorized administrator of Scholarship @ Claremont. For more information, please contact [scholarship@cuc.claremont.edu](mailto:scholarship@cuc.claremont.edu).

## THE EXPECTED CAPACITY OF CONCENTRATORS\*

NICHOLAS PIPPENGER†

**Abstract.** The *expected capacity* of a class of sparse concentrators called *modular concentrators* is determined. In these concentrators, each input is connected to exactly two outputs, each output is connected to exactly three inputs, and the *girth* (the length of the shortest cycle in the connexion graph) is large. Two definitions of expected capacity are considered. For the first (which is due to Masson and Morris), it is assumed that a batch of customers arrive at a random set of inputs and that a maximum matching of these customers to servers at the outputs is found. The number of unsatisfied requests is negligible if customers arrive at fewer than one-half of the inputs, and it grows quite gracefully even beyond this threshold. The situation in which customers arrive sequentially is considered, and the decision as to how to serve each is made randomly, without knowledge of future arrivals. In this case, the number of unsatisfied requests is larger but still quite modest.

**Key words.** communication network, maximum matching, branching process, random packing

**AMS(MOS) subject classifications.** 68E10, 94C15

**1. Batch arrivals.** For the purposes of this paper, a *concentrator* is a bipartite graph  $G = (A, B, E)$  comprising a set  $A$  of *inputs*, a set  $B$  of *outputs*, and a set  $E \subseteq A \times B$  of *edges*. The intended interpretation is that the inputs correspond to “customers,” the outputs correspond to “servers,” and the edges correspond to “channels” or “switches,” each capable of providing direct access by a given customer to a given server.

We consider two modes of operation for a concentrator. In the first mode, the operation of the concentrator takes place in “cycles,” each of which has two “phases.” During the first phase, a subset  $X \subseteq A$  of the inputs, called the *requesting inputs*, is chosen. This represents the arrival of a “batch” of customers. During the second phase, a maximum matching  $M \subseteq E \cap (X \times B)$  between the requesting inputs and the outputs is chosen. This represents the action of a controller granting access to servers to as many customers as possible. The cardinality  $\#X$  is called the *offered traffic*;  $\#M$  is called the *carried traffic*; and  $\#X - \#M$  is called the *lost traffic*.

The *actual capacity* of a concentrator is the largest  $k$  such that the carried traffic is  $k$  for all  $X \subseteq A$  such that  $\#X = k$ . The *expected capacity* of a concentrator (which is a function of the offered traffic  $k$ ) is the expected carried traffic when the requesting inputs are chosen at random, with all sets  $X \subseteq A$  such that  $\#X = k$  is equally likely.

These definitions of actual and expected capacity were given by Masson and Morris [MM], who investigated their values for “binomial” concentrators. In this paper we study their values for a new class of concentrators that we call “modular” concentrators. The asymptotic behaviour of the expected capacity for modular concentrators can be estimated quite sharply, and it appears quite attractive in view of the sparsity of these concentrators. In particular, the lost traffic is negligible when the offered traffic is less than one-half the number of inputs, and it grows quite gracefully even beyond this threshold. Scheinerman [S] has used the methods of this paper to show that even “random concentrators” have performance only slightly worse than that of modular concentrators.

In the second mode, customers arrive sequentially, and the decision as to how to serve each is made randomly, without knowledge of or dependence on future arrivals. We define this mode of operation in more detail in § 7.

---

\* Received by the editors October 10, 1989; accepted for publication (in revised form) May 8, 1990.

† Department of Computer Science, The University of British Columbia, Vancouver, British Columbia, V6T 1W5 Canada.

**2. Modular concentrators.** We deal with a class of concentrators for which each input meets exactly two edges and each output meets exactly three edges. For such a concentrator there is a natural number  $n$  such that  $\#A = 3n$ ,  $\#B = 2n$ , and  $\#E = 6n$ . These concentrators will be called  $(3 : 2)$ -concentrators.

We begin with the observation that the actual capacity of a  $(3 : 2)$ -concentrator is always rather small. By the *cyclomatic number* of a graph with  $v$  vertices and  $w$  edges, we mean the number  $w - v + 1$ . If the cyclomatic number of a graph is at most one, then it contains at most one simple cycle, and thus it has at most two independent paths between any two vertices.

**THEOREM 2.1.** *The actual capacity of a  $(3 : 2)$ -concentrator is  $O(\log n)$ .*

*Proof.* Given the  $(3 : 2)$ -concentrator  $G = (A, B, E)$ , construct the graph  $G^* = (B, E^*)$  with vertices  $B$  corresponding to the outputs of  $G$  and edges  $E^* = \{\{b, b'\} : \{a, b\}, \{a, b'\} \in E \text{ for some } a \in A\}$  corresponding to pairs of outputs that are connected to a common input. Let  $b$  be any vertex of  $G^*$ . Since each vertex meets exactly three edges in  $G^*$ , there are exactly  $3 \cdot 2^{k-1}$  paths of  $k$  steps starting from  $b$ . Thus, if  $k = \lceil \log_2(2n + 1) \rceil$ , there will be three distinct paths starting from  $b$  and ending at a common vertex  $c$ . The union  $U$  of these three paths has at most  $3(k + 1) - 4 = 3k - 1$  vertices, since the beginning and ending vertices are common. But since there are three independent paths from  $b$  to  $c$  in  $U$ , the cyclomatic number of  $U$  must be at least two, and thus  $U$  must contain more edges than vertices. It follows that there is a set of at most  $3k$  inputs in  $G$  that are connected only to a smaller number of outputs; thus the actual capacity of  $G$  is at most  $3k - 1 = O(\log n)$ .  $\square$

We now turn our attention to a class of  $(3 : 2)$ -concentrators for which the expected capacity is much larger than the actual capacity. The *girth* of a graph is the length of the shortest simple cycle in the graph. We construct  $(3 : 2)$ -concentrators with girth  $\Omega(\log n)$ . Our construction follows ideas of Margulis [M1] and Imrich [I].

Let  $PSL(2, \mathbf{Z})$  denote the group of two-by-two integer matrices  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  with determinant one ( $ad - bc = 1$ ), where two matrices are considered the same if their corresponding entries are negatives of each other. This group is generated by the matrices  $S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  and  $R = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}$ . We have  $S^2 = R^3 = -I$ , where  $I$  is the identity matrix. Furthermore, these are the only relations satisfied by  $S$  and  $R$ . Thus  $PSL(2, \mathbf{Z})$  is the free product of  $\mathbf{Z}/(2)$  (generated by  $S$ ) and  $\mathbf{Z}/(3)$  (generated by  $R$ ).

Let  $q \geq 5$  be a prime and let  $PSL(2, \mathbf{Z}/(q))$  be the quotient group of  $PSL(2, \mathbf{Z})$  in which two matrices are considered the same if their corresponding entries differ by multiples of  $q$ . There are  $(q - 1)q(q + 1)/2$  elements in  $PSL(2, \mathbf{Z}/(q))$ . The natural homomorphism  $\pi$  from  $PSL(2, \mathbf{Z})$  to  $PSL(2, \mathbf{Z}/(q))$  reduces entries modulo  $q$ .

A word in  $S$  and  $R$  that is reduced with respect to  $S^2 = I$  and  $R^3 = I$  must consist of occurrences of  $S$  alternating with occurrences of  $R$  or  $R^2 = R^{-1}$ . If such a word is in the kernel of  $\pi$ , it must have norm at least  $q - 1$ . (By the *norm* of a matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , we mean the maximum length of the vector  $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$ , as the vector  $\begin{pmatrix} x \\ y \end{pmatrix}$  varies over the circle  $x^2 + y^2 = 1$ . In particular, the norm of a matrix is at least the maximum of the absolute values of its entries.) It follows that a reduced word in the kernel of  $\pi$  must contain at least  $\log_\beta(q - 1)$  occurrences of  $R$  and  $R^{-1}$ , where  $\beta = (1 + \sqrt{5})/2$ , since the norm of  $S$  is 1, the norms of  $R$  and  $R^{-1}$  are  $\beta$ , and the norm is submultiplicative.

For each prime  $q \geq 5$ , let  $G_q$  denote the  $(3 : 2)$ -concentrator whose edges correspond to the elements of  $PSL(2, \mathbf{Z}/(q))$ , whose inputs correspond to pairs of elements that differ by a factor of  $S$ , and whose outputs correspond to triples of elements that differ by factors of  $R^{\pm 1}$ . Such a concentrator will be called a *modular concentrator*. Clearly,  $n = (q - 1)q(q + 1)/12$ . By the argument of the preceding paragraph, any simple cycle in  $G_q$  must have length at least  $2 \log_\beta(q - 1) = \Omega(\log n)$ . Thus we have proved the following lemma.

LEMMA 2.2. *The girth of a modular concentrator is  $\Omega(\log n)$ .*

The first construction of a  $(3 : 2)$ -concentrator with girth  $\Omega(\log n)$  is due to Gallager ([G, Appendix C]), in the form of the parity-check matrix of a low-density parity-check code with rate  $\frac{1}{3}$  over  $GF(2)$ . (Gallager's construction can be carried out in "polynomial time," but it is not as explicit as the one given above, which can be carried out in "logarithmic space.") We observe that there are more sophisticated constructions that give  $(3 : 2)$ -concentrators with even larger girth than  $G_q$  (roughly  $(8/3) \log_2 n$  rather than  $(2/3) \log_8 n$ ) (see Biggs and Hoare [BH], Weiss [W], Margulis [M2], and Chiu [C]). We also observe that we do not need the full strength of Lemma 2.2. If  $g$  denotes the girth, it is sufficient that  $g \rightarrow \infty$  as  $n \rightarrow \infty$ .

**3. Hypergeometric and binomial capacities.** The expected capacity has been defined *hypergeometrically*, that is, by taking all sets  $X$  of inputs with  $\#X = k$  to be equally likely. We begin by showing that it is possible to deal instead with a set  $X$  of inputs that is defined *binomially*, that is, in which each input appears independently with probability  $p = k/2n$ .

Let  $H(3n, k)$  denote the expected cardinality of a maximum matching when each set of inputs  $X$  with  $\#X = k$  is equally likely. Let  $J(3n, p)$  denote the expected cardinality of a maximum matching when  $X$  contains each input independently with probability  $p$ .

LEMMA 3.1. *For  $0 < p < 1$ ,  $0 < \epsilon < \min\{p, 1 - p\}$  and  $3np$  an integer, we have*

$$J(3n, p - \epsilon) - \epsilon^{-2} \leq H(3n, 3np) \leq J(3n, p + \epsilon) + \epsilon^{-2}.$$

*Proof.* Let  $X'$  be a set in which each input appears independently with probability  $p + \epsilon$ . We have  $\text{Ex}(\#X') = 3n(p + \epsilon)$  and  $\text{Var}(\#X') \leq 3n$ . Thus, by Chebyshev's inequality, we have  $\text{Pr}(\#X' < 3np) \leq 1/3n\epsilon^2$ . If  $\#X' \geq 3np$ , then we may delete  $\#X' - 3np$  inputs from  $X'$  to obtain a set  $X$  with exactly  $3np$  inputs, in such a way that every set of  $3np$  inputs is equally likely. The expected cardinality of a maximum matching for  $X'$  is thus at least  $H(3n, 3np)$  in this case. We thus have  $J(3n, p + \epsilon) \geq (1 - 1/3n\epsilon^2)H(3n, 3np)$ . Since  $H(3n, 3np) \leq 3n$ , we obtain the right-hand assertion of the lemma. A similar argument yields the left-hand assertion.  $\square$

In the following sections we prove the following.

THEOREM 3.2. *For  $0 < p < 1$ , we have*

$$J(3n, p) = 3nh(p) + O(n/(\log n)^{1/2}),$$

where

$$h(p) = \begin{cases} p, & \text{if } 0 < p \leq \frac{1}{2} \\ p - (2p - 1)^3/3p^3, & \text{if } \frac{1}{2} < p < 1. \end{cases}$$

Since  $h(p)$  is continuous in  $p$ , we may apply Lemma 3.1 with  $\epsilon \rightarrow 0$  as  $n \rightarrow \infty$  to obtain the following corollary.

COROLLARY 3.3. *For rational  $0 < p < 1$  and  $n$  such that  $3np$  is integral, we have*

$$H(3n, 3np) = 3nh(p) + O(n/(\log n)^{1/2}).$$

**4. Reduction to small components.** We seek to determine the expected number of pairs in a maximum matching when each input is independently requesting with probability  $p$ . Let  $F(p)$  be the subgraph of  $G$  obtained by deleting each input that is not requesting and each edge meeting such an input. Let  $F^*(p)$  be the corresponding subgraph of  $G^*$ , in which each edge is retained independently with probability  $p$ .

LEMMA 4.1. *In an acyclic connected component of  $F^*(p)$ , all but exactly one of the outputs appear in a maximum matching. In a cyclic connected component of  $F^*(p)$ , all of the outputs appear in a maximum matching.*

*Proof.* If  $F^*(p)$  contains a vertex that meets exactly one edge, we may pair the input corresponding to the edge with the output corresponding to the vertex, then find a maximum matching in the graph that remains after this edge and vertex are deleted. This transformation does not change the cyclomatic number of any component. Since an acyclic component that contains an edge must contain a vertex that meets exactly one edge, repeated application of this transformation to an acyclic component must eventually yield an isolated vertex. This proves the first assertion. Repeated application to a cyclic component must eventually yield a graph  $K^*$  in which every vertex meets at least two edges. In the corresponding bipartite graph  $K$ , every input is connected to exactly two outputs, and every output is connected to at least two inputs, so the marriage theorem ensures the existence of a matching including all of the outputs. This proves the second assertion.  $\square$

Let  $Z(p)$  denote the expected number of acyclic component in  $F^*(p)$ . Lemma 4.1 implies that

$$(4.1) \quad J(3n, p) = 2n - Z(p).$$

Let  $g$  denote the girth of  $G$ , and let  $Y(p)$  denote the expected number of components of  $F^*(p)$  that contain at most  $g/8$  edges. A component with at most  $g/8$  edges must be acyclic, so  $Y(p) \leq Z(p)$ . On the other hand, there are at most  $6n/(g/8) = 48n/g$  components with more than  $g/8$  edges, so  $Z(p) \leq Y(p) + 48n/g$ . Since  $g = \Omega(\log n)$ , we have

$$(4.2) \quad Z(p) = Y(p) + O(n/\log n).$$

Let  $V(p)$  denote the expected number of vertices in  $F^*(p)$  in components with at most  $g/8$  edges, and let  $W(p)$  denote the expected number of edges in such components. Since these components are all acyclic, we have

$$(4.3) \quad Y(p) = V(p) - W(p).$$

Equations (4.1), (4.2), and (4.3) together give the formula

$$(4.4) \quad J(3n, p) = 2n - V(p) + W(p) + O(n/\log n)$$

for the expected capacity in terms of the expected numbers of vertices and edges in small components of  $F^*(p)$ . In the next section we determine the asymptotic behaviour of these expected numbers.

**5. Analysis of small components.** Let  $I$  be an infinite tree in which each vertex meets exactly three edges. Let  $I(p)$  be a random subgraph of  $I$  in which each edge is independently retained with probability  $p$ . Let  $v_k(p)$  be the probability that a vertex of  $I$  belongs to a component of  $I(p)$  with at most  $k$  edges. Let  $w_k(p)$  be the conditional probability that an edge  $e$  of  $I$  belongs to a component of  $I(p)$  with at most  $k$  edges, given that  $e$  is retained in  $I(p)$ . It is clear that

$$(5.1) \quad V(p) = 2n v_{g/8}(p) \quad \text{and} \quad W(p) = 3np w_{g/8}(p),$$

since a neighbourhood of radius  $g/8$  about any vertex or edge in  $G^*$  is isomorphic to a corresponding neighbourhood in  $I$ , and all quantities in (5.1) are defined in terms of random variables that are independent of events outside these neighbourhoods.

Let  $v(p)$  denote the probability that a vertex in  $I$  belongs to a finite component of  $I(p)$ , and let  $w(p)$  denote the conditional probability that an edge  $e$  of  $I$  belongs to a finite component of  $I(p)$ , given that  $e$  is retained in  $I(p)$ . The theory of branching processes gives the following lemma.

LEMMA 5.1. *We have*

$$v(p) = \begin{cases} 1, & \text{if } 0 < p \leq \frac{1}{2} \\ (1-p)^3/p^3, & \text{if } \frac{1}{2} < p < 1; \end{cases}$$

and

$$w(p) = \begin{cases} 1, & \text{if } 0 < p \leq \frac{1}{2}; \\ (1-p)^4/p^4, & \text{if } \frac{1}{2} < p < 1. \end{cases}$$

*Proof.* Consider a branching process in which the first generation contains a single individual, and each individual in the  $i$ th generation independently contributes to the  $(i + 1)$ st generation a number of offspring that is binomially distributed with generating function  $(1 - p + px)^2$ . According to Harris ([H], Chap. I, Thm. 6.1]), the probability of extinction (that is, the probability that the family generated in this way is finite) is the root  $q(p)$  of equation  $x = (1 - p + px)^2$ , given by

$$q(p) = \begin{cases} 1, & \text{if } 0 < p \leq \frac{1}{2}; \\ (1-p)^2/p^2, & \text{if } \frac{1}{2} < p < 1. \end{cases}$$

The probability that a vertex in  $I(p)$  belongs to a finite component is simply the probability of extinction when the first generation contains a number of individuals distributed with generating function  $(1 - p + px)^3$ , the generating function for the number of edges incident with the given vertex in  $I(p)$ . This extinction probability is  $(1 - p + pq(p))^3$  (which is as given in the statement of the lemma).

Similarly, the conditional probability that an edge  $e$  in  $I$  belongs to a finite component of  $I(p)$ , given that  $e$  is retained in  $I(p)$ , is  $(1 - p + pq(p))^4$  (which is as given in the statement of the lemma), since  $(1 - p + px)^4$  is the conditional generating function for the number of edges incident with  $e$  in  $I(p)$ , given that  $e$  is retained in  $I(p)$ .  $\square$

LEMMA 5.2. *We have*

$$v_k(p) = v(p) + O(k^{-1/2})$$

and

$$w_k(p) = w(p) + O(k^{-1/2}).$$

*Proof.* Clearly,  $v_k(p) \leq v(p)$ . Furthermore,  $v(p) - v_k(p)$  is simply the probability that, in the branching process described in the proof of Lemma 5.1 (with the generating function of the initial distribution being  $(1 - p + px)^3$ ), extinction occurs after the size of the family exceeds  $k$ . According to Harris ([H, Chap. I, Thm 13.1]), the conditional probability that the size of the family is  $j$ , given that extinction occurs, is  $O(j^{-3/2})$ . (The decay is actually much faster than this unless  $p = \frac{1}{2}$ .) Thus the probability that extinction occurs after the size exceeds  $k$  is  $\sum_{j>k} O(j^{-3/2}) = O(k^{-1/2})$ . The proof for  $w_k(p)$  and  $w(p)$  is analogous.  $\square$

Applying Lemmas 2.2 and 5.2 to (5.1) yields

$$V(p) = 2nv(p) + O(n/(\log n)^{1/2})$$

and

$$W(p) = 3npw(p) + O(n/(\log n)^{1/2}).$$

Substitution of these formulae and Lemma 5.1 into (4.4) completes the proof of Theorem 3.2.

**6. Extensions for batch arrivals.** The concentrators that we have considered are one-stage networks; that is, each edge directly connects an input to an output. It is easy to see, however, that the analysis we have given has immediate application to some multistage networks.

Consider for example the “two-stage (9 : 4)-concentrators” constructed in the following way. Let  $q \geq 5$  and  $q' \geq 5$  be primes (equal or distinct), and set  $n = (q - 1)q(q + 1)/12$  and  $n' = (q' - 1)q'(q' + 1)/12$ . Take  $3n'$  disjoint copies of  $G_q$  and  $2n$  disjoint copies of  $G_{q'}$ , and link each output of each copy of  $G_q$  to an input of a copy of  $G_{q'}$ , with exactly one link between each copy of  $G_q$  and each copy of  $G_{q'}$ . If the inputs of the resulting network are independently requesting, and if appropriate random choices of the maximum matchings in the copies of  $G_q$  are made, then the inputs of each copy of  $G_{q'}$  will be independently requesting, and the analysis given above can be applied to each stage in turn. (The traffics offered to the various copies of  $G_{q'}$  will be dependent, but this does not affect the expected capacity.) The expected capacity will again be piecewise rational, now with breakpoints at  $p = \frac{1}{3}$  (the onset of loss in the second stage) and  $p = \frac{1}{2}$ . The extension to three or more stages should be clear.

It is possible to extend the analysis we have given, with hardly any changes in the arguments, to “( $a : 2$ )-concentrators with large girth” (for integer  $a > 2$ ). (The construction of such concentrators can be accomplished by the methods of the papers cited in § 2.) It may also be possible to extend Theorem 3.2 (though not Theorem 2.1) to “( $a : b$ )-concentrators with large girth” (for integers  $a > b > 1$ ). There seems to be nothing as simple as Lemma 4.1 in this case, but the success of Karp and Sipser [KS] in treating the problem of maximum matchings in sparse random graphs gives hope. For  $b = 2$  we prove (and for  $b > 2$  it is natural to conjecture) that  $v(p)$  is replaced by  $q(p)^a$  and  $w(p)$  is replaced by  $q(p)^{(a-1)b}$ , where  $q(p)$  is now the appropriate root of the equation  $x = (1 - p + px^{b-1})^{a-1}$ .

**7. Sequential arrivals with random hunting.** We now turn to a second mode of operation for concentrators. Consider a concentrator  $G = (A, B, E)$ . Associate with each input  $a \in A$  an *arrival time*  $\tau_a$ , uniformly distributed in the interval  $[0, 1]$ , and independent of all other arrival times. The intended interpretation is that the customer corresponding to input  $a$  arrives at time  $\tau_a$ .

Next associate with each input  $a \in A$  a *hunting order*  $\beta_a$ , uniformly distributed over the total orders among the outputs connected to  $a$ , independent of the hunting orders of other inputs and independent of the arrival times of all inputs. The intended interpretation is that when the customer arrives at input  $a$  (at time  $\tau_a$ ), it examines the outputs connected to  $a$  in the order prescribed by  $\beta_a$  until it finds one that has not been engaged previously (that is, at a time less than  $\tau_a$ ). If it finds such an output, the output is engaged at time  $\tau_a$ . If it finds no such output, no action is taken, and the customer remains unserved.

Some comments about this mode of operation are in order. First, the assumption of uniformly distributed arrival times will facilitate calculations, but other independent and identically distributed arrival times would also result in all possible orders of arrival being equally likely, and in the number of arrivals before time  $t$  being binomially distributed. (The choice of the arrival-time distribution may be regarded as a choice of the parametrisation of time. An exponential distribution, corresponding to Poisson arrivals, seems the most natural physically.) Second, results concerning the expected number of customers served for this “binomial” arrival process can easily be translated (by the argument given in § 3) into results for the “hypergeometric” arrival process, in which some number  $k$  of customers arrive at distinct inputs, with all possible sets of  $k$  inputs, as well as all possible orders of arrival, being equally likely.

**8. Sequential arrivals for trees.** We begin our analysis by looking at some concentrators that are trees. Let  $C_0$  denote the concentrator with a single input that is connected to two outputs, one of which is called the *root* and the other of which is called the *leaf*. For some  $k \geq 1$ , suppose that  $C_{k-1}$  has been defined. Let  $C_k$  denote the concentrator obtained by identifying the leaf of a copy of  $C_0$  with the roots of two copies of  $C_{k-1}$  to form an internal output (neither a root nor a leaf); the root of the copy of  $C_0$  becomes the root of  $C_k$ , and the leaves of the copies of  $C_{k-1}$  (of which there are  $2^k$ ) become the leaves of  $C_k$ .

For  $k \geq 0$  and  $0 \leq t \leq 1$ , let  $Q_k(t)$  denote the probability that the root of  $C_k$  is engaged at time  $t$ .

LEMMA 8.1. *We have*

$$Q_0(t) = t/2$$

and

$$(8.1) \quad Q_k(t) = \int_0^t 1 - \frac{1}{2}(1 - Q_{k-1}(s))^2 ds.$$

*Proof.* For the root of  $C_0$  to be engaged at time  $t$ , the customer must arrive by time  $t$ , which happens with probability  $t$ , and must choose the root before the leaf in the hunting order, which happens independently with probability  $\frac{1}{2}$ . This proves the first assertion. For the root of  $C_k$  to be engaged at time  $t$ , the customer must again arrive by time  $t$ . If the customer arrives at time  $s$ , then it will engage the root unless it chooses the leaf of  $C_0$  before the root in the hunting order, and the leaf of  $C_0$  is not engaged by time  $s$ . This leaf will be engaged by time  $s$  if and only if the root of one of the copies of  $C_{k-1}$  would be engaged by time  $s$  (with the same arrival times and hunting orders in the copies). These events depend on arrival times and hunting orders for disjoint sets of inputs, so they are independent. This proves the second assertion.  $\square$

We now show that the transformation  $Q_{k-1} \mapsto Q_k$  has a fixed point; that is, a solution  $Q$  of the integral equation

$$(8.2) \quad Q(t) = \int_0^t 1 - \frac{1}{2}(1 - Q(s))^2 ds.$$

To do this, we differentiate (8.2) with respect to  $t$  to obtain the differential equation

$$(8.3) \quad Q'(t) = 1 - \frac{1}{2}(1 - Q(t))^2,$$

with the initial condition  $Q(0) = 0$ . Since (8.3) does not involve  $t$  explicitly, it can be solved by quadratures:

$$(8.4) \quad \int_0^{Q(t)} \frac{dx}{1 - (1-x)^2/2} = t,$$

where the lower limit of integration has been chosen to satisfy the initial condition. The substitution  $y = (1 - x)/\sqrt{2}$  reduces the integral to

$$\sqrt{2} \int_{(1-Q(t))/\sqrt{2}}^{1/\sqrt{2}} \frac{dy}{1-y^2} = \sqrt{2} \tanh^{-1} \frac{1}{\sqrt{2}} - \sqrt{2} \tanh^{-1} \frac{1-Q(t)}{\sqrt{2}}.$$

Thus

$$Q(t) = 1 - \sqrt{2} \tanh \left( \ln(1 + \sqrt{2}) - \frac{t}{\sqrt{2}} \right),$$

since  $\tanh^{-1} 1/\sqrt{2} = \ln(1 + \sqrt{2})$ .



LEMMA 8.2. *We have  $Q_k(t) \rightarrow Q(t)$  uniformly in  $t$  as  $k \rightarrow \infty$ .*

*Proof.* Set  $\Delta_k(t) = Q_k(t) - Q(t)$ . Since  $0 \leq Q_0(t), Q(t) \leq 1$ , we have  $|\Delta_0(t)| \leq 1$ . Furthermore, (8.1) and (8.2) imply

$$|\Delta_k(t)| = \left| \int_0^t \frac{\Delta_{k-1}(s)(2 - Q_{k-1}(s) - Q(s))}{2} ds \right|.$$

Since  $0 \leq Q_{k-1}(s), Q(s) \leq 1$ , we have  $|2 - Q_{k-1}(s) - Q(s)| \leq 2$ , so that

$$|\Delta_k(t)| \leq \int_0^t |\Delta_{k-1}(s)| ds.$$

Thus by induction on  $k$  we obtain

$$|\Delta_k(t)| \leq t^k/k!.$$

This completes the proof.  $\square$

For  $k \geq 0$ , let  $D_k$  denote the concentrator obtained by identifying the roots of three copies of  $C_k$  to form the root of  $D_k$ ; the leaves of the copies of  $C_k$  (of which there are  $3 \cdot 2^k$ ) are the leaves of  $D_k$ . Letting  $R_k(t)$  denote the probability that the root of  $D_k$  is engaged at time  $t$ , we clearly have  $R_k(t) = 1 - (1 - Q_k(t))^3$ . Finally, putting  $R(t) = 1 - (1 - Q(t))^3$ , we see that  $R_k(t) \rightarrow R(t)$  uniformly in  $t$  as  $k \rightarrow \infty$ . Thus we have proved the following proposition.

PROPOSITION 8.3. *As  $k \rightarrow \infty$ , the probability  $R_k(t)$  that the root of  $D_k$  is engaged at time  $t$  tends to*

$$R(t) = 1 - \left( \sqrt{2} \tanh \left( \ln(1 + \sqrt{2}) - \frac{t}{\sqrt{2}} \right) \right)^3,$$

*uniformly in  $t$ .*

**9. Sequential arrivals for modular concentrators.** Now consider the concentrator  $G_q$  and arbitrarily designate one output of this concentrator as the “root.” Let  $N_k$  denote the subgraph of  $G_q$  induced by the inputs of  $G_q$  at distance at most  $2k + 1$  from the root and the outputs of  $G_q$  at distance at most  $2k + 2$  from the root. Call the outputs at distance  $2k + 2$  from the root the “leaves” of  $N_k$ . Set  $k = \lfloor (g - 6)/4 \rfloor$ , where  $g$  is the girth of  $G_q$ . Since  $g = \Omega(\log q)$  (by Part I, Lemma 2.2), we have  $k \rightarrow \infty$  as  $q \rightarrow \infty$ . Furthermore, since  $4k + 4$  is less than the girth of  $G_q$ ,  $N_k$  is a tree isomorphic to  $D_k$ , with root corresponding to root, and leaves corresponding to leaves. Let  $S_q(t)$  denote the probability that the root of  $G_q$  is engaged at time  $t$ .

LEMMA 9.1. *We have  $S_q(t) \sim R_k(t)$  uniformly in  $t$  as  $q \rightarrow \infty$  and hence  $k \rightarrow \infty$ .*

*Proof.* Suppose we wish to determine whether the root of  $G_q$  is engaged at some time  $t$ . This is determined by the arrival times and hunting orders of the inputs in  $N_1$ , unless some input at distance three from the root has an earlier arrival time than the intermediate vertex at distance one; that is, unless there is a path of decreasing arrival times from the root to some leaf of  $N_2$ . Even if there is such a path, the engagement of the root is determined by the arrival times and hunting orders of the inputs in  $N_2$ , unless there is a path of decreasing arrival times from the root to a leaf in  $N_3$ . In general, the engagement of the root is determined by the arrival times and hunting orders of the inputs in  $N_k$ , unless there is a path of decreasing arrival times from the root to a leaf in  $N_k$ .

Let  $X_k$  denote the event “there is a path of decreasing arrival times from the root to a leaf in  $N_k$ .” We have  $\Pr(X_k) \leq 3 \cdot 2^k/k!$ , since there are  $3 \cdot 2^k$  paths from the root to a leaf in  $N_k$ , and the probability that the arrival times along some such path are decreasing is  $1/k!$  (since all  $k!$  orders of arrival are equally likely). Furthermore, unless

$X_k$  occurs, the root is engaged in  $G_q$  when and only when it is engaged in  $N_k$ . Thus we have  $|S_q(t) - R_k(t)| \leq 3 \cdot 2^k / k!$ .  $\square$

Combining this with Proposition 8.3, we have proved the following theorem.

**THEOREM 9.2.** *As  $q \rightarrow \infty$ , the probability  $S_q(t)$  that an output in  $G_q$  is engaged at time  $t$  tends to*

$$R(t) = 1 - \left( \sqrt{2} \tanh \left( \ln(1 + \sqrt{2}) - \frac{t}{\sqrt{2}} \right) \right)^3,$$

*uniformly in  $t$ . In particular, the probability that an output is never engaged tends to*

$$\left( \sqrt{2} \tanh \left( \ln(1 + \sqrt{2}) - \frac{1}{\sqrt{2}} \right) \right)^3 = 0.0145 \dots$$

**10. Extensions for sequential arrivals.** The extensions we have described for batch arrivals all apply to sequential arrivals as well. In particular, for “ $(a : b)$ -concentrators with large girth,” we obtain integral equations that can still be solved by quadratures, though not in general in terms of elementary functions. It is easy, however, to carry out the quadratures numerically and to obtain the fraction of unused servers as a function of time.

When the concentration ratio  $a/b$  is an integer, a new possibility arises that does not occur for  $(3 : 2)$ -concentrators. In this case, it is possible to assign fixed hunting orders to the inputs in such a way that each output is the first choice for  $a/b$  inputs, the second choice for another  $a/b$ , and so forth. For such an assignment, there can be no unused servers after all customers have arrived. The analysis of this mode of operation leads to differential equations (or systems of differential equations) that cannot be solved by quadratures. It is easy, however, to integrate them numerically, and to obtain the fractions of requests that are served by their first choice, their second choice, and so forth.

#### REFERENCES

- [BH] N. L. BIGGS AND M. J. HOARE, *The sextet construction for cubic graphs*, *Combinatorica*, 3 (1983), pp. 153–165.
- [C] P. CHIU, *Cubic Ramanujan graphs*, *Combinatorica*, to appear.
- [G] R. G. GALLAGER, *Low-Density Parity-Check Codes*, MIT Press, Cambridge, MA, 1963.
- [H] T. E. HARRIS, *The Theory of Branching Processes*, Springer-Verlag, New York, 1963.
- [I] W. IMRICH, *Explicit constructions of regular graphs without small cycles*, *Combinatorica*, 4 (1984), pp. 53–59.
- [KS] R. M. KARP AND M. SIPSER, *Maximum matchings in sparse random graphs*, *IEEE Symp. on Foundations of Computer Science*, 22 (1981), pp. 364–375.
- [M1] G. A. MARGULIS, *Explicit constructions of graphs without short cycles and low density codes*, *Combinatorica*, 2 (1982), pp. 71–78.
- [M2] ———, *Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators*, *Problems Inform. Transmission*, 24 (1988), pp. 39–46.
- [MM] G. M. MASSON AND S. B. MORRIS, *Expected capacity of  $(\frac{m}{2})$ -networks*, *IEEE Trans. Comput.*, 32 (1983), pp. 649–656.
- [P] N. PIPPENGER, *Random sequential adsorption on graphs*, *SIAM J. Discrete Math.*, 2 (1989), pp. 393–401.
- [S] E. R. SCHEINERMAN, *On the expected capacity of binomial and random concentrators*, *SIAM J. Comput.*, 19 (1990), pp. 156–163.
- [W] A. WEISS, *Girths of bipartite sextet graphs*, *Combinatorica*, 4 (1984), pp. 241–245.