On the Fault Tolerance of the Butterfly

Anna R. Karlin* Greg Nelson † Hisao Tamaki‡

Abstract

We study the robustness of the butterfly network against random static faults. Suppose that each edge of the butterfly is present independently of other edges with probability p. Our main result is that there is a 0-1 law on the existence of a linear-sized component. More formally, there is a critical probability $p*$ such that for p above $p*$, the faulted butterfly almost surely contains a linear-sized component, whereas for p below $p*$, the faulted butterfly almost surely does not contain a linear-sized component.

1 Introduction

Given a graph G, let G/p denote the random subgraph obtained by considering each edge independently and including it in the subgraph with probability p, excluding it with probability $1-p$. [1] We call G/p a *faulted version* of G. A long list of theorems illustrate the basic fact that small changes in p can lead to dramatic changes in the connectivity of G/p. In this paper we add to this list a new theorem that shows the degree of fault-tolerance of the butterfly network: we prove that there is a critical probability for the existence of a linear-sized component in the faulted butterfly, and give a characterization of the critical probability from which it can be estimated.

We begin with a precise definition of the critical probability. Let $\Gamma(G)$ be the fraction of the nodes of G contained in the largest connected component of G.

Let G be a family of graphs, indexed by the natural numbers, so that G_d is a graph for each $d \geq 0$. We say that $p*$ (which may be a function of d) is the *critical probability for the existence of a linear-sized component* if for every constant $\epsilon > 0$, (1) for every $p > (1+\epsilon)p*$, there exists $c > 0$, such that

$$\lim_{d\to\infty} \Pr(\Gamma(G_d/p) > c) = 1$$

and (2) for all $c > 0$ and for all $p < (1-\epsilon)p*$,

$$\lim_{d\to\infty} \Pr(\Gamma(G_d/p) > c) = 0.$$

It is not at all obvious that critical probabilities exist. However, several well-known theorems can be interpreted as giving the critical probabilities for various classes of graphs. In particular:

- Let K_d denote the complete graph on d nodes. Then K has critical probability $1/d$. This is a central result of the fundamental paper on random graph theory, by Erdős and Rényi in 1960 [ER60].
- Let G_d denote the d by d two-dimensional grid. Then G has critical probability $1/2$. This was proved by Kesten in 1980, and is the best-known result of percolation theory [Kes80].
- Let C_d denote the hypercube of dimension d. Then C has critical probability $1/d$. This was proved by Ajtai, Komlós, and Szemerédi in 1981 [ES79, AKS82].

*DEC Systems Research Center, 130 Lytton Ave., Palo Alto, CA 94301.

†DEC Systems Research Center, 130 Lytton Ave., Palo Alto, CA 94301.

‡IBM T.J. Watson Research Center, P.O.Box 218 Yorktown Heights, NY 10598; part of this work was done while this author was at the University of Toronto

[1] All results in this paper can be carried over mutatis mutandis to the case where nodes fail instead of edges.

We have altered the usual statement of these theorems to fit them into our mold. For example, the result for the grid is usually stated as the critical probability for the existence of an infinite component in the faulted infinite grid. But it is not difficult to prove that the infinite component appears in the infinite grid at the same critical probability that a linear-sized component appears in the family of faulted finite grids.

Notice that in the case of the complete graph and the hypercube, the critical probability is essentially the reciprocal of the degree; so that the linear-sized component appears as soon as the expected degree in the faulted graph exceeds one. In this sense, the critical probability of the grid is higher.

Note also that in the case of the complete graph and the hypercube, there is another critical probability, above which all nodes of the faulted graph will be connected. But there is no analog of this for bounded-degree networks such as the grid and the butterfly; a linear-sized component is the best we can hope for if p is bounded away from 1. For example, for any $p < 1$, G_d/p will be disconnected, since any particular node will be isolated with probability $(1-p)^4$.

The main theorem of this paper is:

Theorem 1.1 *Let B_d denote the d-dimensional butterfly graph. Then B has critical probability $p*$, where $0.337 \leq p* \leq 0.436$.*

This theorem implies that there is a sense in which the butterfly is far more fault-tolerant to random faults than previous results have suggested. Although it was known that there exists a constant $p < 1$ such that B_d/p contains a linear-sized component[Tam92], the value of p readily extractable from the proof therein is astronomically close to 1. In comparison to the 2-dimensional grid which has critical probability 0.5, our theorem shows that the butterfly, with the same degree 4 as the grid, withstands significantly more faults. Although the existence of a linear-sized component is not the only measure of fault-tolerance, it certainly is one of the most important measures: without a linear-sized component, the computation on the faulted version is bound to be slower than the original network by more than a constant factor. Interestingly, for the 2-dimensional grid, the critical probability for the existence of a linear-sized component coincides with the critical probability for other properties also related to the efficiency of computation, such as the ability of message routing and the ability of embedding the fault-free version[Rag89, KKL$^+$90, Mat92]. Whether or not the same phenomena occur on the butterfly is an interesting question, towards which we make a small amount of progress as described in the following paragraphs.

A *fault-avoiding self-embedding* maps nodes of G to nodes of G/p and edges of G to unfaulted paths of G/p, such that if a path corresponds to an edge, its endpoints correspond to the edge's endpoints. The *load* of the embedding is the maximum number of nodes of G mapped onto the same node of G/p. The *dilation* of the embedding is the maximum of the lengths of the paths in G/p corresponding to the edges in G. The *congestion* of the embedding is the maximum number of edges in G whose corresponding paths share a particular edge of G/p. A self-embedding with small load, dilation, and congestion implies an efficient simulation of the fault-free network by the faulty network, and thus can be considered evidence of the robustness of the network structure. (In fact, the slowdown factor of the simulation is O(load + dilation + congestion) due to the result of Leighton, Maggs, and Rao [LMR] on off-line routing.)

Efficient fault-avoiding self-embeddings have been constructed for the hypercube (by Hastad, Leighton and Newman[HLN87, HLN89] and Aiello and Leighton [AL91]) and the grid (by a mammoth collaboration [KKL$^+$90]). But the butterfly has proved harder to deal with, since it does not have the high connectivity of the hypercube, and it does not have the simple geometric structure of the grid, which can be exploited to circumnavigate faults. Recently one of the authors (Tamaki [Tam92]) has constructed fault-avoiding self-embeddings of the butterfly, such that the load is $O(1)$ and the congestion and dilation are both polyloglog in the number of nodes. However, the allowable probability of edge failure extracted from the proof of [Tam92] is unrealistically small. Using the techniques in the proof of theorem 1.1 we can improve this result to

obtain the following.

Theorem 1.2 *For edge presence probability exceeding 0.586, a faulted n-node butterfly can embed a fault-free n-node butterfly with $O(1)$ load, $O((\log\log n)^{2.6})$ dilation and $O((\log\log n)^c)$ congestion with probability at least $1-n^{-1}$, where c is a constant that depends on p. There is a polynomial time algorithm for constructing the embedding.*

In related work, Leighton, Maggs and Sitaraman [LMS92] have shown that for small, but constant edge-failure probability, [2] an n-node butterfly can emulate a fault-free version of itself with a slowdown of $2^{O(\log^* n)}$. Their emulation uses redundant computation. It would be very nice to be able to prove results of this kind for any $p > p*$.

We also prove two additional facts about the largest connected component of B_d/p, for constant $p > p*$. The first is that the linear-sized component is uniform, in that it contains a constant fraction of the nodes on each level of the butterfly. The second is that the linear-sized component is unique; the sizes of all other components are sublinear.

The rest of the paper is organized as follows. Section 2 describes our notation and gives the characterization of the critical probability. Section 3 proves that above the critical probability, the faulted butterfly has a linear sized component. Section 4 proves that below the critical probability, the faulted butterfly does not have a linear sized component. Taken together, this proves the main theorem. Section 5 presents the above mentioned two extensions to our main theorem and explains how we apply our result to obtain Theorem 1.2. Section 6 presents some upper and lower bounds on $p*$.

2 Notation and Preliminaries

Let B_d denote the butterfly of dimension d: each node is defined by its *pattern*, a sequence of bits of length d, and its *level*, a number in $\{0, 1, \ldots, d\}$. Two nodes are connected by an edge if their levels differ by exactly one and their patterns differ at most at the bit whose index is the lower of the two levels. We sometimes imagine that the lower levels are on the left and the higher levels are on the right. We denote the number of nodes of B_d by $n_d = (d+1)2^d$.

A fact that will come in handy is that the subgraph of B_d induced by the set of nodes whose patterns have a given j-bit prefix and a given k-bit suffix and whose levels are between j and $d-k$ is isomorphic to B_{d-j-k}.

We give the following characterization of the critical probability of B_d. Fix a node v in level 0 of B_d, and let $X_d(p)$ be the random variable denoting the number of nodes in level d of B_d that are connected to v in B_d/p. (Because any two patterns in B_d can be exchanged by a level-preserving isomorphism, the distribution of $X_d(p)$ does not depend on the choice of the node v in level 0.) Define $p*$ by

$$p* = \inf\{p \mid E(X_d(p)) > 1 \text{ for some } d\}.$$

Our main theorem states that $p*$ is the critical probability for the existence of a linear-sized component in B_d. See section 6 for a discussion of the explicit value of $p*$.

3 Above the critical probability

The first part of our theorem states that for constant $p > p*$, B_d/p has a linear sized component with high probability.

Theorem 3.1 *For any constant $p > p*$, there is a constant $c > 0$ such that*

$$\lim_{d\to\infty} \Pr(\Gamma(B_d/p) > c) = 1.$$

Proof:

The basic technique for showing the presence of a large connected component in a faulted graph is to compute an upper bound on the probability that it is separated into two smaller components, by multiplying the number of separations by the maximum probability that any separation *survives*. A separation survives if none of the edges of G/p bridges the two components of the separation. If many edges of G bridge the separation, then the probability that all of them are excluded from G/p will be small. This approach is difficult to apply to the butterfly, since there are separations that

[2] These results are actually stated for node-failure probability, but results of one type extend to the other type and vice versa.

are bridged by too few edges of B_d. For example, a subbutterfly of size m is connected to the main butterfly by only $O(m/\log m)$ edges.

To get around this problem, we imagine the randomization performed in two stages. That is, we first choose p_1 and p_2 with $p* < p_1$ and $0 < p_2$, so that $(1-p_1)(1-p_2) = 1-p$. Note that $B_d/p_1 \bigcup B_d/p_2 = B_d/p$, since the probability that an edge is not present in either B_d/p_1 or B_d/p_2 is $(1-p_1)(1-p_2) = 1-p$. In the first part of the proof we show that in B_d/p_1, a constant fraction of the nodes are in connected components of size at least $(n_d)^\delta$. (We call these components *atoms*.) In the second part of the proof we show that the additional edges in B_d/p_2 connect many of the atoms, enough so that the resulting component is of linear size.

Because the first randomization leaves a constant fraction of the nodes in atoms, the number of possible separations that need to be considered in the second part of the proof is only $2^{(n_d)^{(1-\delta)}}$ instead of 2^{n_d}. This is small enough to allow us to show that with high probability, no separation survives the second randomization. However, to prove that no separation survives, we must consider bridging the separation by paths, not just edges. To keep the length of the paths small enough, we need to prove another fact about the first randomization: that every node is reasonably close to an atom. This turns out to follow from essentially the same argument that shows that a constant fraction of the nodes are in atoms.

Although the details are different, the idea of using two stages of randomization, and branching process theory in order to construct connected components in the first randomization can be found in the hypercube result of [AKS82]. The next two subsections give the details of the proof.

3.1 Part 1

The proof of part 1 relies on a result from the theory of branching processes. Recall that a discrete-time branching process consists of an initial population. In our case, the initial population has size 1. Each individual in the population has a number of immediate descendents given by a random variable Z_1. The nth generation of the process is composed of descendents of the $(n-1)$st generation, each of whose members independently produces a number of progeny, distributed according to Z_1. We denote by Z_n the population size of the nth generation. It is easy to see that $E(Z_n) = (E(Z_1))^n$.

It is well known that as long as $E(Z_1) > 1$, there is positive probability that the population will never go extinct. Furthermore, if the line continues forever, then with high probability the number of offspring in generation n, Z_n, is very likely to be close to its expected value. We formalize this statement in the following lemma. [3]

Lemma 3.2 *Suppose $E(Z_1) = \mu > 1$. Then for any constant $0 < \alpha < 1$,*

$$\liminf_{n\to\infty} \Pr(Z_n > \mu^{\alpha n}) > 0.$$

The following lemma is the basis for our proof.

Lemma 3.3 *Let v be any node on level 0 of the butterfly B_d. Suppose $p > p*$. Then there exist positive constants γ and ϵ such that, for every sufficiently large d, the probability is at least γ that node v belongs to a connected component of B_d/p with at least $(n_d)^\epsilon$ nodes.*

Proof: Let d_0 be a constant such that $\mu = E(X_{d_0}(p)) > 1$. Assume d much bigger than d_0 and consider the following branching process. The generation 0 individual is the given node v on level 0 of B_d/p. In general, generation r individuals will be nodes on level rd_0 of B_d/p. The process is defined up to the kth generation, for $k = \lfloor d/d_0 \rfloor$. Suppose that a node w at level rd_0 is a generation r descendant. Then, we consider the (unique) d_0–dimensional subbutterfly that contains w on its leftmost level. The immediate descendants (in generation $r+1$) of w in the branching process are the nodes on the rightmost level of this subbutterfly that are connected in B_d/p to w through a path confined to this subbutterfly. Note that for any two distinct nodes in the same generation, the subbutterflies defining their immediate descendants do not overlap, since their patterns have different prefixes. Thus, we have a branching process with $Z_1 = X_{d_0}(p)$. Since we chose d_0

[3] A stronger version of this lemma, for the special case Z_1 is binomial, is given in [Pip92].

so that $E(X_{d_0}(p)) > 1$, Lemma 3.2 applies. Let γ be the value of the limit in the statement of the lemma. Then, for any $\alpha < 1$, there is a constant probability γ that the process survives with at least $\mu^{\alpha k}$ descendants in generation k. But k is $\Omega(d) = \Omega(\log(n_d))$, and $\mu > 1$. Therefore, with probability at least γ, v belongs to a connected component of B_d/p with at least $\mu^{\alpha k} \geq (n_d)^{\epsilon}$ nodes, for some $\epsilon > 0$. ∎

Let $0 < \beta < 1/4$ be a constant whose value will be determined in part 2 of the proof. Let $\delta = \epsilon/4$, where ϵ is the constant given in the above lemma depending on p, and define an *atom* to be a maximal connected component in B_d/p that has size at least $(n_d)^{\delta}$.

Lemma 3.4 *For any constant $p > p*$, there exist constants $\nu > 0$ and $c' > 0$ such that*

1. *The probability that fewer than $c'n_d$ nodes in B_d are in atoms is at most $2^{-(n_d)^{\nu}}$.*
2. *The probability that there exists a node in B_d that is not within distance βd (in B_d) of an atom is at most $2^{-(n_d)^{\nu}}$.*

Proof: Consider a level l of B_d where $l \leq 3d/4$. From each $d/4$–dimensional subbutterfly of B_d whose leftmost level is on level l of B_d, pick exactly one node on the leftmost level and collect them into a set S of $2^{3d/4}$ nodes. For each $w \in S$, let $B(w)$ denote the subbutterfly from which w is picked. By lemma 3.3 applied to $B(w)$, each $w \in S$ has probability at least γ of being in an atom (note $(n_{d/4})^{\epsilon} \geq (n_d)^{\delta}$), and these events are independent. Therefore, setting $c' = \frac{\gamma}{2}$, the probability that fewer than $c'|S|$ nodes in S are in atoms is $2^{-\Omega(|S|)} = 2^{-\Omega(2^{3d/4})}$, by a standard tail estimate. Now, the set of nodes on each level l can be partitioned into $2^{d/4}$ disjoint subsets each of which are constructed as S above (flip left/right for $l > 3d/4$). Applying the above argument to each subset and then summing up the failure probability, we conclude that the probability of fewer than $c'n_d$ nodes of B_d/p being in atoms is still $2^{-\Omega(2^{3d/4})}$.

For the second part, let v be an arbitrary node and assume without loss of generality that v is on the left half of B_d. Consider the βd–dimensional subbutterfly whose leftmost level contains v and let T be the set of nodes on the rightmost level of this subbutterfly. This set T has the same property as S above except that the its size is $2^{\beta d}$. By a similar argument, the probability that no node of T is in an atom is at most $2^{-\Omega(2^{\beta d})}$. Since the graph distance between v and any node in T is βd, this bounds the probability that a given node violates the second statement of the lemma. Sum this failure probability over all nodes in the graph. ∎

3.2 Part 2

Choose p_1 and p_2 with $p* < p_1$ and $0 < p_2$, so that $(1-p_1)(1-p_2) = 1-p$. In part 1, we showed that with high probability, $c'n_d$ of the nodes in B_d/p_1 are in atoms, connected components of size at least $(n_d)^{\delta}$, and that with high probability, each node in B_d/p_1 is within distance βd of an atom. We now outline the proof that the edges of B_d/p_2 are likely to connect the atoms in B_d/p_1 into a linear sized component. The edges of B_d/p_2 connect the atoms into *clumps*.

Observation 3.5 *If no clump has size at least $c'n_d/3$, the clumps can be partitioned into two groups, each of size at least $c'n_d/3$, such that no edge of B_d/p_2 connects the two groups.*

A pair of sets of nodes (A, B) is a *separation* if A and B are disjoint, each atom is contained in either A or B, and each of A and B has size at least $c'n_d/3$. The separation *survives* if no path of B_d/p_2 connects A and B. Then the probability that some separation survives is bounded by the number of separations times the maximum probability that a separation survives.

Since atoms have size at least $(n_d)^{\delta}$, there are at most $(n_d)^{1-\delta}$ atoms and hence at most $2^{(n_d)^{1-\delta}}$ possible separations. We will show that the probability that a separation survives is at most $2^{-(n_d)^{\alpha}}$, where $\alpha > 1 - \delta$, which will complete the proof of the theorem.

Let (A, B) be a separation. Define A^* as the set of nodes whose distance from A is at most βd, and B^* similarly. By definition, $A \subseteq A^*$ and $B \subseteq B^*$.

The following fact is well-known.

Fact 3.6 *For any set S of nodes in B_d such that both S and its complement are linear-sized, $|\partial(S)| \geq \Omega(|S|/\log|S|)$, where $\partial(S)$ is the set of edges with one endpoint in S and one endpoint in the complement of S.*

Let $E = \partial(A^* \setminus B)$. Since A and B are linear-sized, and $|A| \leq |A^* \setminus B| \leq n_d - |B|$, Fact 3.6 implies that $|E| \geq \Omega(n_d/d)$.

Each edge in E has one endpoint in A^* and the other endpoint in B^*, since $A^* \bigcup B^*$ contains all the nodes. By definition, the endpoint in A^* (resp. B^*) is at distance at most βd from a node in A (resp. B). Therefore, associated to each edge in E, there is a path of at most $2\beta d + 1$ edges that can connect A and B. Since each edge can be part of at most $(2\beta d+1)\cdot 3^{2\beta d}$ such paths, we can greedily choose at least $|E|/(2\beta d+1)^2 \cdot 3^{2\beta d} \geq (n_d)^{1-4\beta-o(1)}$ edge-disjoint paths from the set. Therefore the probability that the separation survives is at most

$$\left(1 - p_2^{2\beta d+1}\right)^{(n_d)^{1-4\beta-o(1)}} \leq 2^{-(n_d)^{\alpha}},$$

for β sufficiently small.

This completes the proof that for $p > p*$, B_d/p contains a linear-sized component.

∎

4 Below the critical probability

The second part of our theorem states that if $p < p*$, then B_d/p is very unlikely to have a linear sized component.

Theorem 4.1 *For any constant $p < p*$, and for any constant $c > 0$,*

$$\lim_{d\to\infty} \Pr(\Gamma(B_d/p) > c) = 0.$$

Proof: We prove the contrapositive. That is, we assume that for some positive c and γ there are infinitely many values of d such that $\Pr(\Gamma(B_d/p) > c) \geq \gamma$, and use that to show that $E(X_d(p))$ exceeds 1 for sufficiently large d.

We need the following lemma.

Lemma 4.2 *For any positive c and γ, there exist positive α and δ such that for any d, if $\Pr(\Gamma(B_d/p) > c) \geq \gamma$ then there exist two nodes u and v in B_d at least αd levels apart such that the probability that u and v are connected in B_d/p is at least δ.*

Proof: We refer to the largest connected component of B_d/p (selected uniformly at random if there is more than one) as the *primary* component of B_d/p. For each pair of nodes u and v of B_d, let $Y_{u,v}$ denote the indicator random variable that is 1 if both u and v belong to the primary component of B_d/p and 0 otherwise. Then the sum of $E(Y_{u,v})$ over all ordered pairs of nodes of B_d is the expected number of ordered pairs in the primary component. Hence,

$$\sum_{u,v} E(Y_{u,v}) \geq \gamma(cn_d)^2,$$

where the summation is over all ordered pairs of nodes of B_d. Therefore, there exist some positive constants α and δ independent of d such that at least for $2\alpha(n_d)^2$ pairs (u, v) the expectation $E(Y_{u,v})$ is at least δ. At least one such pair (u, v) must be αd levels apart, because the number of pairs less than αd levels apart is at most $n_d(2\alpha d - 1)n_d/d < 2\alpha(n_d)^2$. Such a pair (u, v) satisfies the required condition, because the probability that u and v are connected in B_d/p is at least $E(Y_{u,v})$. ∎

By our contrapositive assumption, there exist positive constants c and γ for which the assumption of Lemma 4.2 is satisfied by infinitely many values of d. Among such values of d, let d_0 be a constant to be determined. Furthermore, let l and $l+m$ be the levels of u and v in B_{d_0} given by lemma 4.2. Note that $m > \alpha d_0$.

Now consider B_d for $d > d_0$. We find a maximal sequence of nodes $w_0, w_1, \ldots, w_k$, that satisfies the following conditions.

1. Each w_i is on level $l + i \cdot m$, for $0 \leq i \leq k$.
2. For each i, $1 \leq i \leq k$, there is a d_0-dimensional subbutterfly such that w_{i-1} and w_i occupy the positions of u and v in that subbutterfly, under some isomorphism.

Let C_i, $1 \leq i \leq k$, denote the event that w_{i-1} and w_i are connected in B_d/p. We have $\Pr(C_i) \geq \delta$ for each i by the choice of d_0, condition 2 above and lemma 4.2.

Let w_0' be the node on level 0 with the same pattern as w_0 and w_k' be the node on level d with the same pattern as w_k. Let D denote the event that

node w_0' is connected to w_0 in B_d/p and E the event that w_k is connected to w_k' in B_d/p. By considering only the simplest paths, we have $\Pr(D) \geq p^l$, and $\Pr(E) \geq p^{d_0-l}$.

Clearly,

$$\Pr(w_0' \text{ and } w_k' \text{ connected in } B_d/p)$$
$$\geq \Pr(D \& C_1 \& \ldots \& C_k \& E).$$

However, events $D, E, C_1, \ldots, C_k$ are monotone increasing, in the sense that throwing in more edges never turns the events off. Therefore, we can use the FKG inequality (see, for example, [AS92]) to obtain

$$\begin{aligned} & \Pr(w_0' \text{ and } w_k' \text{ connected in } B_d/p) \\ \geq\ & \Pr(D \& C_1 \& \ldots \& C_k \& E) \\ \geq\ & \Pr(D)\Pr(C_1)\ldots\Pr(C_k)\Pr(E) \\ \geq\ & p^{d_0}\delta^k. \end{aligned}$$

We now use symmetry considerations to infer that this bound on connection probability applies to any pair of nodes u on level 0 and v on level d. Therefore,

$$E(X_d(p)) \geq p^{d_0}\delta^k 2^d.$$

Since k is roughly $\frac{d}{\alpha d_0}$, for a judicious choice of d_0, and sufficiently large d, we obtain

$$E(X_d(p)) > 1.$$

This completes the proof that for $p < p*$, B_d/p has no linear-sized component.

5 Uniformity and uniqueness of the linear-sized component

In this section, we give two extensions to our main theorem. The first extension states that the linear-sized component of B_d/p for $p > p*$ is uniform in the sense that it contains a constant fraction of nodes from each level of B_d. The second extension states that the linear-sized component is unique in the sense that the number of nodes in the second largest component is $o(n_d)$.

Theorem 5.1 *For any constant $p > p*$, there exists a constant $c > 0$ such that with probability tending to 1 (as $d \to \infty$) B_d/p has a connected component that contains at least $c2^d$ nodes in each level of B_d.*

The proof is a refinement of the proof of Theorem 3.1. We briefly describe the idea. From the proof of Lemma 3.4, we already know that c'–fraction of the nodes on each level of B_d/p belong to atoms with high probability. To show that the edges of B_d/p_2 are likely to combine these atoms into a clump that contains a constant fraction of nodes from each level, what we need is the following observation analogous to Observation 3.5.

Observation 5.2 *If no clump has at least $c'2^d/3$ nodes from each level, the clumps can be partitioned into two groups such that one group contains at least $c'2^d/3$ nodes from some level i and the other contains at least $c'2^d/3$ nodes from some level j.*

Proof: Under the assumption, we obtain the required partition (X, Y) of the clumps in the following procedure. Initially, set X to be the entire set of clumps and $Y = \{\}$. We move clumps from X to Y one by one until the clumps of Y contains at least $c'2^d/3$ nodes at some level i for the first time. Before the last move, no level had $c'2^d/3$ of its nodes in the clumps of Y and therefore every level had at least $2c'2^d/3$ nodes in X. Let j be the level such that the last clump moved does not have $c'2^d/3$ nodes of level j. Then the clumps of X must have $c'2^d/3$ nodes of level j after the last move. ∎

Thus the following fact which generalizes Fact 3.6 allows us to proceed in the manner similar to the proof of Theorem 3.1.

Fact 5.3 *For any set S of nodes in B_d such that S has a constant fraction of nodes of some level and its complement also has a constant fraction of nodes of some level, $|\partial(S)| \geq \Omega(|S|/\log|S|)$,*

Theorem 5.1 can be applied to the problem of embedding the butterfly into a faulted butterfly. For $p > 0.586$, we can show that the constant c in this theorem is greater than $1/2$. This allows us to show that the base case in the inductive proof of the embedding result in [Tam92] essentially holds

for $p > 0.586$ yielding Theorem 1.2. The details can be found in [Tam93].

Now we turn to the uniqueness of the linear-sized component.

Theorem 5.4 *For any constants $p > p*$ and $\epsilon > 0$, the probability tends to 0 (as $d \to \infty$) that the second largest connected component of B_d/p contains at least ϵn_d nodes.*

Proof Sketch: Let v be an arbitrary node on level d of the $2d$–dimensional butterfly B_{2d}, and let $\theta_d(p)$ denote the probability that v is connected in B_{2d}/p to some node on level 0 or level $2d$. For fixed p, $\theta_d(p)$ is decreasing in d. Moreover, by Theorem 5.1, for any $p > p*$ there exists a constant $\gamma > 0$ such that with probability at least γ, v belongs to a component of B_{2d}/p that contains some nodes on the outermost levels. This means that $\theta_d(p) \geq \gamma$ for every d. Therefore, $\theta_d(p)$ for any fixed $p > p*$ has limit $\theta(p) > 0$ as $d \to \infty$. Roughly speaking, our proof consists of showing for any $\epsilon > 0$ that:

1. the size of the largest connected component of B_d/p is at least $(\theta(p) - \epsilon/4)n_d$ with probability $1 - o(1)$;
2. there exists a constant K such that with probability $1 - o(1)$, at most $(\theta(p) + \epsilon/4)n_d$ nodes of B_d/p belong to connected components that have size larger than K.

Thus, there is no room for a second largest component with more than $(\epsilon/2)n_d$ nodes.

The proof of the first claim is again a refinement of the proof of Theorem 3.1. To carry out this refinement, we need to show (1) that the constant c' of Lemma 3.4, i.e. the fraction of the nodes belonging to atoms, can be made arbitrarily close to $\theta(p)$ and (2) that the constant c of Theorem 3.1 can be made arbitrarily close to c'. We omit the details of (1); the intuition is that we expect a fraction $\theta(p)$ of the nodes to be roots of non-extinct branching processes and, if the branching process does not die out, it is likely to be boosted at some point to form an atom. For (2), we simply notice that the two occurrences of constant $1/3$ in Observation 3.5 can be generalized to a pair of constants $1 - 2\delta$ and δ in this order, where δ is an arbitrarily small positive constant.

To prove the second claim, let d_0 be a large enough constant so that $\theta_{d_0}(p) < \theta(p) + \epsilon/8$. Let d be much larger than d_0, and for each node v of B_d, let A_v denote the event that v is connected in B_d/p to some node at least d_0 levels apart from v. Set $K = n_{2d_0}$ and note that v can belong to a component of B_d/p with more than K nodes only if A_v occurs. Thus, it suffices to bound the number of nodes v for which A_v occurs. By the definition of $\theta_d(p)$, we have $\Pr(A_v) \leq \theta_{d_0}(p) < \theta(p) + \epsilon/8$. Although events A_v are not totally independent, we can none-the-less apply the standard tail estimate, by partitioning the set of nodes of B_d into subsets so that events A_v are mutually independent within each subset. Thus, with probability $1 - o(1)$, A_v occurs to at most $(\theta(p) + \epsilon/4)n_d$ nodes. ∎

6 Estimate of the critical probability

Our characterization of the critical probability gives an algorithm to compute it. For $d \geq 1$, let p_d denote the probability such that $E(X_d(p_d)) = 1$. Then, the sequence $p_1, p_2, \ldots$, is decreasing and converges to $p*$, and each p_d can be computed to arbitrary precision, at least in principle. It is easy to see that $p_1 < 1/2$, and the computation of p_2 is not excessively time consuming; we have $p_2 \sim 0.435$. A Monte-Carlo simulation using a pseudorandom generator suggests $p_{10} \sim 0.39$.

For lower bounds, let λ_d denote the number of simple paths of length d in B_d starting from a particular node v at level 0. Then, $\lambda_d p^d$ is the expected number of simple paths of length d in B_d/p starting from v. If $p > p*$ then this expectation must be $\Omega(1)$ because there is a $\Omega(1)$ probability that v belongs to a uniform linear-sized component; such a component contains a node on level d and thus must contain a path of length d starting from v. Therefore, $\lambda_d p^d = \Omega(1)$ for every $p > p*$. The trivial bound $\lambda_d \leq 3^d$ implies $p* \geq 1/3$. By observing that one out of 27 possible ways of continuing a path from each intermediate node u in three steps leads to the predecessor of u forming a cycle, we get $\lambda_d \leq 26^{d/3}$, hence $p* \geq 1/\sqrt[3]{26} \simeq 0.337$.

Acknowledgments

We thank Nick Pippenger who for his helpful comments and suggestions.

References

[AKS82] M. Ajtai, J. Komlós, and E. Szemerédi. Largest random component of a k-cube. *Combinatorica*, 2(1):1–7, 1982.

[AL91] B. Aiello and F.T. Leighton. Coding theory, hypercube embeddings, and fault tolerance. In *Proceedings of the 3rd ACM Annual Symposium on Parallel Algorithms and Architectures*, 1991.

[AS92] N. Alon and J. H. Spencer. *The probabilistic method.* John Wiley & Sons, New York, 1992.

[ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960. known from [Bol85].

[ES79] P. Erdős and J. Spencer. Evolution of the n-cube. *Computers and Math. with Applications*, 5:33–40, 1979.

[HLN87] J. Hastad, F.T. Leighton, and M. Newman. Reconfiguring a hypercube in the presence of faults. In *Proceedings of the 19th ACM Annual Symposium on Theory of Computing*, pages 274–284, 1987.

[HLN89] J. Hastad, F.T. Leighton, and M. Newman. Fast computation using faulty hypercubes. In *Proceedings of the 21st ACM Annual Symposium on Theory of Computing*, pages 251–263, 1989.

[Kes80] H. Kesten. The critical probability of bond percolation on the square lattice equals 1/2. *Communication in Mathematical Physics*, 74:41–59, 1980.

[KKL+90] C. Kaklamanis, A.R. Karlin, F.T. Leighton, V. Milenkovic, P. Raghavan, S. Rao, C. Thomborson, and A. Tsantilas. Asymptotically tight bounds for computing with faulty arrays of processors. In *Proceedings of the 31st Annual Symposium on Foundations of Computer Science*, pages 285–296. IEEE, 1990.

[LMR] F.T. Leighton, B.M. Maggs, and S.B. Rao. Packet routing and job-shop scheduling in O(congestion + dilation) steps. *Combinatorica.* to appear.

[LMS92] F. T. Leighton, B. Maggs, and R. Sitaraman. On the unexpected fault-tolerance of some popular bounded-degree networks. In *Proceedings of the 33d Annual Symposium on Foundations of Computer Science*, pages 542–552. IEEE, 1992.

[Mat92] T.R. Mathies. Percolation theory and computing with faulty arrays of processors. In *Proceedings of the 3rd ACM-SIAM Symposium on Discrete Algorithms*, pages 100–103, 1992.

[Pip92] N. Pippenger. The asymptotic optimality of spider-web networks. *Discrete applied mathematics*, 37/38:437–450, 1992.

[Rag89] P. Raghavan. Robust algorithms for packet routing in a mesh. In *Proceedings of the 1st ACM Annual Symposium on Parallel Algorithms and Architectures*, pages 344–350, 1989.

[Tam92] H. Tamaki. Efficient self-embedding of butterfly networks with random faults. In *Proceedings of the 33rd Annual Symposium on Foundations of Computer Science*, pages 533–541, 1992.

[Tam93] H. Tamaki. Structural robustness of the butterfly and related networks against random faults. PhD Thesis, Department of Computer Science, University of Toronto, 1993.