Appeared in ***31st ACM Symposium on Principles of Distributed Computing (PODC '12)***

On the Price of Equivocation in Byzantine Agreement

Alexander Jaffe
University of Washington
ajaffe@cs.washington.edu

Thomas Moscibroda
Microsoft Research Asia
moscitho@microsoft.com

Siddhartha Sen
Princeton University
sssix@cs.princeton.edu

ABSTRACT

In the Byzantine agreement problem, a set of n processors, any f of whom may be arbitrarily faulty, must reach agreement on a value proposed by one of the correct processors. It is a celebrated result that unless $n > 3f$, Byzantine agreement is impossible in a variety of computation and communication models. This is due to the fact that faulty processors can *equivocate*, that is, say different things to different processors. If this ability is mitigated, for example by assuming a global broadcast channel, then $n > 2f$ is sufficient. With very few exceptions, the literature on Byzantine agreement has been confined to the $n > 2f$ and $n > 3f$ paradigms.

We bridge the gap between these two paradigms by assuming partial broadcast channels among sets of three processors, observing that equivocation is fundamentally an act involving three parties: a faulty processor that lies (inconsistently) to two correct processors. We characterize the conditions under which Byzantine agreement is possible for all $n = 2f + h$, h an integer in $[1..f]$, by giving asymptotically tight bounds on the number of necessary and sufficient partial broadcast channels. We prove these bounds by a reduction to a problem in extremal combinatorics, which itself is a natural generalization of a well-studied hypergraph coloring problem. Algorithmically, we show that deciding whether a given set of broadcast channels enables Byzantine agreement is co-NP-complete. Although partial broadcast channels have been studied in prior work, the bounds obtained on the number of required channels were sub-optimal by up to a factor of $\Theta(n^2)$. Moreover, this work has been confined to the synchronous model. In contrast, we apply our results to several distinct models and provide stronger motivation for using partial broadcast channels in practice, drawing from recent work in the systems community.

Categories and Subject Descriptors

G.2.2 [**Discrete Mathematics**]: Graph Theory—*graph labeling, hypergraphs*; F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems—*computations on discrete structures*

Keywords

Byzantine agreement, partial broadcast, hypergraph coloring, expander graphs

PODC'12, July 16–18, 2012, Madeira, Portugal.
Copyright 2012 ACM 978-1-4503-1450-3/12/07 ...$10.00.

1. INTRODUCTION

One of the most celebrated results in distributed computing theory is the bound on the redundancy required to solve the *Byzantine Agreement* problem. In this problem, a set of n processors, each with an initial value and any f of whom may be arbitrarily faulty, must reach agreement on a value proposed by one of the correct processors.[1] Lamport, Shostak, and Pease showed in 1982 that in the standard communication model of a complete synchronous network of pairwise authenticated channels, Byzantine agreement is possible if and only if $n > 3f$ [31], implying a significant amount of redundancy. Moreover, the $n > 3f$ bound is remarkably robust to changes in the underlying communication and computation model [9, 18, 28].

A closer inspection of these results reveals that the fundamental reason for requiring $n > 3f$ processors is that faulty processors can *equivocate*, *i.e.*, say different things to different processors. For instance, in a synchronous system with 3 processors, a single faulty processor can consistently send different messages to the two correct processors and make them agree to different values [31]. In asynchronous systems, the adversary's ability to delay messages (in addition to its ability to equivocate) can confound the correct processors even if cryptography is used [18]. Thus, several researchers have considered using stronger communication primitives such as broadcast channels. Broadcast channels mitigate equivocation by ensuring that a message appears identically at all recipients on the channel. In the synchronous model, Rabin and Ben-Or [37] introduced a global broadcast channel and achieved Byzantine agreement (in fact, any multiparty protocol) if and only if $n > 2f$. Fitzi and Maurer [20] added $\Theta(n^3)$ *partial broadcast* channels among every set of three processors to achieve Byzantine agreement if and only if $n > 2f$. Ravikant et al. [38] reduced the number of 3-processor channels required for $n = 2f + h$, where h is an integer in $[1..f]$, assuming sufficient connectivity in the underlying network. However, they get asymptotically tight results only for the same case $h = 1$ as Fitzi and Maurer do. Finally, Considine et al. [14] generalized partial broadcast to sets of $b > 3$ processors, achieving reliable broadcast (but not consensus) when $n > f(b+1)/(b-1)$.

The above overview of previous work illustrates that the gap between $n > 3f$ and $n > 2f$ represents a fundamental *price of equivocation*. In this paper, we give a complete and tight characterization of this price, by studying the relationship between the fraction of processor 3-tuples that are prevented from equivocating, modeled as 3-processor partial broadcast channels, and the fault resilience required to solve Byzantine agreement. Specifically, we study the use, application, and algorithmic implications of 3-processor broadcast channels to Byzantine agreement. We view these channels as

[1] This variant of the problem is called *consensus*. The variant where only a single processor has an initial value is called *reliable broadcast*. Consensus implies reliable broadcast, but the reverse is only true if faulty processors are in the minority, *i.e.* $n > 2f$.

3-hyperedges in the processor graph. Whereas prior work has almost exclusively focused on resilience $n > 2f$, we show that the range $n \geq 2f+h$, where $h \in [1..f]$ allows for significantly more efficient constructions by requiring asymptotically fewer than $\binom{n}{3} = \Theta(n^3)$ 3-hyperedges, assuming standard requirements on the connectivity of the underlying graph [34, 38]. We derive asymptotically tight bounds for all $h \in [1..f]$. Interestingly, this problem turns out to be a natural generalization of a well-studied *hypergraph coloring problem*. Although prior work has been limited to the synchronous model, we show how to apply our results to various other models as well.

The motivation for studying 3-hyperedges is two-fold. First, a 3-hyperedge (x,y,z) represents the fundamental unit of equivocation: a faulty processor x says different things to correct processors y and z. Second, while it is possible (and likely more efficient) to create a single or larger broadcast channel with x and the processors it has 3-hyperedges with, the expressiveness of 3-hyperedges may be useful in practice. For example, if hyperedges (x,y,z) and (x,y,w) exist but not (x,z,w), then a protocol may require x to use partial broadcast when sending a message to y and z or to y and w, but not when sending a message to z and w.

1.1 Results and outline

Let H be a 3-uniform hypergraph on n vertices (representing processors), where each 3-hyperedge represents a partial broadcast channel. Our main result is an asymptotically tight characterization of the necessary and sufficient number of 3-hyperedges required to achieve Byzantine agreement despite f faulty processors, for all $n \geq 2f+h$, h a positive integer in $[1..f]$. h is thus the parameter that interpolates between the well-studied cases $n > 2f$ and $n > 3f$. As in prior work [38], we assume the underlying graph is at least $(2f+1)$-connected. Let $T_n(h)$ denote the minimum m such that there exists an H with m 3-hyperedges that achieves Byzantine agreement. We show:

- $T_n(h(n)) = \Theta\left(\frac{n^3}{h(n)^2}\right)$

For comparison, the only other existing work that gives results on this trade-off is by Ravikant et al. [38], who obtain an upper bound of $T_n(h(n)) = \mathrm{O}((f-h(n)+1)f^2) = \mathrm{O}((n-3h(n)+1)n^2)$, which is up to a factor of $\Theta(n^2)$ off the correct bound, depending on $h(n)$. They also give a near-tight bound for the case $n = 2f+1$, but this result is asymptotically identical to the trivial solution of including all 3-hyperedges. Their constructions are elementary and use a clever and simple recursive structure, though the analysis is nontrivial. We improve on their results by using the power of expander graphs, building hypergraphs out of existing expander constructions in order to exploit their high connectivity. Although we can prove our bound on $T_n(h(n))$ using a simple probabilistic method argument, in this work we are concerned with *explicit constructions*. A strong motivation for this goal is given by the following result.

Theorem 1. *Given a 3-uniform hypergraph $H = (V,E)$ with $|V| = n$, it is* co-NP*-complete to decide, for any $n = 2f+h$, $h \in [1..f]$, whether Byzantine agreement is possible in H despite f faulty processors.*

The proof of this theorem, a reduction from balanced bipartite independent set, can be found in Section 6. Since it is intractable to detect the possibility of Byzantine agreement in general, it is possible that explicit construction is the only reliable means of exploiting the efficiency gains of sparse fault-tolerant hypergraphs.

Our final result gives an exact bound for $U_n(h(n))$, the minimum m such that any H with m hyperedges achieves Byzantine agreement:

- $U_n(h(n)) = \binom{n}{3} - \frac{n-h(n)}{2} \cdot h(n)^2 + 1$

Section 1.2 discusses the application of our results to upper and lower bounds for Byzantine agreement in various models. Section 2 formally defines the problem and proves an equivalence to a more natural combinatorial problem, which we use in the remainder of the paper. Section 3 proves the lower bound on $T_n(h(n))$ using a graph projection and counting arguments. Section 4 describes explicit constructions of H that match the upper bound on $T_n(h(n))$; we rely on a "lifting" procedure that converts existing constructions of Ramanujan graphs into hypergraphs with expander-like properties. Section 5 proves the bound on $U_n(h(n))$ using multivariate minimization techniques. Section 7 concludes with some open problems. Due to space constraints, some proofs are omitted and deferred to the full version of this paper.

1.2 Algorithms and applications

The results of this paper naturally give rise to new upper and lower bounds for Byzantine agreement in various models (augmented with partial broadcast channels). Specifically, our results apply to any proof that relies on the following intersection property between quorums of processors (made precise in Section 2), which we call *f-tolerance*: Consider a quorum of size $n-f$; although $n-f$ is the number of correct processors, a quorum of this size may still contain up to f faulty processors. Given two such quorums, the correct processors of one quorum may disagree with those of the other because faulty processors common to both may equivocate, unless their intersection contains at least one correct processor: $2(n-f)-n > f \implies n > 3f$. The key insight is that we get the same guarantee by replacing the correct processor x with a 3-hyperedge (x,y,z) such that y and z are correct processors in distinct quorums. Even if x is faulty, hyperedge (x,y,z) prevents it from equivocating to y and z and making their quorums agree on inconsistent values.

Ravikant et al. [38] prove that Byzantine agreement is possible in the synchronous model if and only if a set of conditions which includes f-tolerance holds. The remaining conditions are the standard connectivity requirements of the underlying graph, which we also assume in our setting. Thus Theorem 1 in their paper implies lower and upper bounds for Byzantine agreement in our setting as well. However, the hypergraphs they construct are suboptimal: they provide a tight bound only for $n = 2f+1$ [38, Sections 5.2 and 5.3] and loose upper bounds for any $n = 2f+h, h \in [1..f]$ [38, Section 5.1]. By replacing their construction with ours from Section 4, we reduce the number of 3-hyperedges and the message complexity of their protocol by up to a factor of $\Theta(n^2)$.

In the asynchronous model, we can adapt the lower bound of Bracha and Toueg [9] to show that f-tolerance is necessary for Byzantine agreement. The proof is essentially the same as Theorem 3 in their paper, except that instead of any two $(n-f)$-sized quorums, a pair of quorums that violates f-tolerance must be used. We can obtain upper bounds for any $n = 2f+h$ by modifying the protocols of Bracha and Toueg [9, Figure 2] or Bracha [8], for example. We do this by overlaying our hypergraph construction onto the processor graph and requiring x to use hyperedge (x,y,z), if it exists, when sending a message to both y and z. Although this reduces the efficiency of the protocol, the relevant correctness proofs ([9, Theorem 4] and [8, Sections 3.2 and 6]) are readily adapted because they rely on f-tolerance. Similarly, there has been tremendous recent interest in the systems community on designing efficient Byzantine agreement protocols in a partially synchronous model with cryptography (*e.g.*, [4, 12, 15, 27, 29]). This model is subject to Bracha and Toueg's lower bound [9], and essentially all upper bounds are derivatives of Castro and Liskov's protocol [11], which relies on f-tolerance for correctness [11, Invariants A.1.4 and A.1.5]. Therefore we can improve the resiliency of these protocols as above. This reduces their replication costs, which is often cited as an obstacle to practical deployment [27, 39, 43, 44].

There are several ways to implement 3-hyperedges in practice. One way is to use multicast groups; another is to use a shared cyptographic key; another is to use trusted primitives like an *append-only log* [12] or *trusted incrementer* [32]. Depending on the implementation, it may not always be possible to force a processor x to use a 3-hyperedge *a priori*, but it is always possible for y and z to generate a proof of misbehavior (POM) [2] against x *a posteriori* if x violates the protocol. Several systems [2, 26, 29] use POMs in this manner. Finally, if each 3-hyperedge is allowed to fail with some probability [20], our hypergraph construction reduces the probability of failure because there are fewer 3-hyperedges to union bound over.

1.3 Other related work

If cryptography is used, consensus is possible in the synchronous model when $n > 2f$ [18, 31]. Whereas consensus cannot be solved in the asynchronous model [19], the $n > 3f$ bound applies in this model when using randomized algorithms that terminate almost surely [9] or with probability $1-\varepsilon$ for fixed $\varepsilon > 0$ [28], even if cryptography is used [18]. The same bound also applies in partially synchronous models [18].

Partial broadcast was first considered by Franklin, Wright, and Yung [22, 23] in the context of secure point-to-point communication over an incomplete network. Stronger primitives based on trusted subsystems and cryptography have also been used, such as *weak sequenced broadcast* [1] to solve weak Byzantine agreement in the asynchronous model, and *append-only log* [12] and *trusted incrementer* [32] to solve Byzantine agreement in a partially synchronous model. These primitives achieve resilience $n > 2f$.

We mention two other lines of work that are related to ours. The first considers hybrid fault models that combine Byzantine and crash failures (*e.g.*, [25, 30]), in which optimal bounds on resilience [25, 30] depend on the number of faults of each type. The second considers a non-threshold adversary characterized by an *adversary structure* (*e.g.*, [5, 21]), or a monotone set of subsets of processors any one of which may be faulty. It is known [21] that Byzantine agreement is possible if and only if no three sets cover all processors.

2. PROBLEM DEFINITION AND AN EQUIVALENCE

We model a system of n processors as a 3-uniform, n-vertex hypergraph $H = (V,E)$ where each edge $(x,y,z) \in E$ represents a partial broadcast channel. For a fixed integer f, we analyze the conditions under which Byzantine agreement is possible in H, when up to f processors are faulty. As in prior work [38], we assume the underlying graph is at least $(2f+1)$-connected (*e.g.*, via a complete set of 2-hyperedges (edges) connecting the processors). As explained in Section 1.2, this problem is equivalent to ensuring that in the intersection of any two size-$(n-f)$ quorums S and T, there exists a node z that cannot equivocate between correct nodes $x \in S, y \in T$. We assume w.l.o.g. that $S \cap T$ contains only faulty nodes, because a correct node in $S \cap T$ would prevent equivocation.

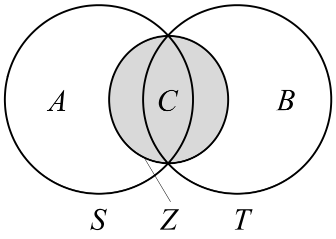

Figure 1: The relationship between sets S,T,Z of f-tolerance and sets A,B,C of h-disjointness. Shaded regions contain faulty nodes. The diagram above shows $C \subset Z$, but in general $C \subseteq Z$.

This reduces the possibility of Byzantine agreement to the following property of H.

Definition 1. *A 3-uniform hypergraph $H = (V,E)$ with $|V| = n$ vertices is f-*tolerant *if $\forall\, S,T,Z \subset V$ and $S \neq T$ satisfying the conditions below, $\exists\, x \in S \setminus Z, y \in T \setminus Z, z \in S \cap T$ for which $(x,y,z) \in E$:*

- $|Z| \leq f$,
- $|S|, |T| \geq n-f$,
- $S \cap T \subseteq Z \subset S \cup T$.

The property as defined is somewhat unwieldy, because the sets S, T, and Z can overlap in a variety of ways. To simplify our problem statement, we introduce a new property on disjoint sets and show its equivalence. Consider the sets $A = S \setminus Z$, $B = T \setminus Z$, and $C = S \cap T$; A, B, and C are disjoint because $S \cap T \subseteq Z$. A and B contain the correct nodes in quorums S and T, respectively, and C contains the faulty nodes in their intersection. (There may be other faulty nodes in the two quorums, limiting the size of A and B, but only in sum.) We will shortly redefine f-tolerance in terms of the following notion.

Definition 2. *A 3-uniform hypergraph $H = (V,E)$ with $|V| = n$ vertices is h-*disjoint *if for all disjoint $A,B,C \subset V$ satisfying the conditions below, $\exists\, x \in A, y \in B, z \in C$ for which $(x,y,z) \in E$:*

- $|A|, |B|, |C| \geq h$,
- $|A| + |B| + |C| \geq \frac{n+3h}{2}$.

Figure 1 illustrates the equivalence of f-tolerance and h-disjointness. Note that although A, B, and C are symmetric in the above definition, we will often distinguish them as in Figure 1 for ease of exposition. The following theorem makes this equivalence precise, with proof deferred to the full version:

Theorem 2. *Let H be a 3-uniform hypergraph on n vertices, and integer $f \geq \frac{n}{3}$. Then H is f-tolerant if and only if H is $(n-2f)$-disjoint.*

h-disjointness is equivalent to the notion of $(3,f)$-hyper-$(3f-n+1)$-connectedness in [38]. However, we find our definition to be simpler and more clearly related to the hypergraph coloring literature, discussed below. The remainder of this paper characterizes the hypergraphs that are h-disjoint by deriving tight bounds on the necessary and sufficient number of edges, $T_n(h)$ and $U_n(h)$ respectively. As we observed in Section 1.2, these results imply new upper and lower bounds for Byzantine agreement in different models. We start with $T_n(h)$:

Definition 3. *For positive integers n and h, $T_n(h)$ is the minimum m such that there exists an h-disjoint 3-uniform hypergraph with m edges.*

2.1 Related Combinatorial Problems

h-disjointness can be seen as a generalization of a rich body of work on *mixed hypergraph coloring* and *the upper chromatic number* (see Voloshin's book [42]). We present a single definition that essentially captures all of these concepts. A *k-heterochromatic coloring* of a hypergraph $H = (V,E)$ is a surjection $\chi : V \to [k]$ such that the restriction of χ to some $e \in E$ is injective. In other words, some edge has no repeated color. When H is k-uniform, this is equivalent to some edge being k-chromatic, as in h-disjointness.

A primary line of research in this area sought to analyze $f(n,k)$, the minimum number of edges among k-heterochromatically colorable, k-uniform, n-vertex hypergraphs [6, 7, 10, 17, 41], which was recently resolved up to lower order terms by Bujtás and Tuza [10]. The specific (earlier) result that is relevant to this paper is $f(n,3) = \frac{n(n-2)}{3}$. h-disjointness has immediate connections to $f(n,3)$, but introduces the additional concepts of *balance* and *partiality* in colorings, controlled by h. When $h = 1$, h-disjointness is equivalent to the condition that there is a trichromatic edge for all *small partial* colorings A,B,C, with $|A|,|B|,|C|$ non-empty, but total size only $|A|+|B|+|C| = (n+3)/2$. This condition is strictly stronger than requiring all complete colorings to have a trichromatic edge, because every complete coloring contains a small partial coloring. In contrast, when $h = f = n/3$, h-disjointness is equivalent to the condition that there is a trichromatic edge for all *balanced complete* colorings, with $|A| = |B| = |C| = n/3$. This is strictly weaker than restricting all complete colorings to have a trichromatic edge. For $1 < h < n/3$ the condition will be that all somewhat small, somewhat balanced colorings have a trichromatic edge.

Thus we may already state that $T_n(1) \geq f(n,3) = \frac{n(n-2)}{3}$ and $T_n(n/3) \leq f(n,3) = \frac{n(n-2)}{3}$. As we will see, we can in fact do much better and show a smooth transition $T_n(h) = \Theta\left(\frac{n^3}{h(n)^2}\right)$. In particular, this gives $T_n(1) = \Theta(n^3)$, and $T_n(n/3) = \Theta(n)$, both a factor of n from $f(n,3)$.

3. LOWER BOUNDS

In this section, we will give asymptotically tight bounds on $T_n(h(n))$ for all non-decreasing functions $h(n)$. We will need the following notion of a hypergraph-projection.

Definition 4. *Let $H = (V,E)$ be a 3-uniform hypergraph, and $W \subseteq V$ a subset of its vertices. We define the* projection of H by W *as the* ***graph*** *$H_W = (V_W, E_W)$. The vertex set V_W is defined as $V \setminus W$. The edge set E_W has an edge (u,v) if and only if E has an edge (u,v,w), for some $w \in W$.*

This definition allows us to apply graph-theoretic theorems to hypergraphs, with potentially little loss. We extend the technique of [40], [7], [16], used to prove the aforementioned lower bound on k-heterochromatically colorable hypergraphs. They consider a hypergraph H for which every k-coloring contains a k-chromatic hyperedge, and proceed to lower bound the size of its edge set in two steps. First, they lower bound the number of edges in the projection of H by each $(k-2)$-vertex subset. Then, they upper bound the number of possible edge-projections of each hyperedge, giving a lower bound on the number of original hyperedges.

Our analysis is similar, with an added layer of complexity in lower bounding the number of edges in each projection. Because $h(n)$-disjointness implies that hyperedges cross somewhat balanced partitions of subsets of the vertices, we cannot assume that the projections are connected graphs. For large $h(n)$, we can only assume very weak conditions on the projections' connectivity. For small $h(n)$, we can assume conditions even stronger than connectedness. To address this, we prove a pair of results on the number of edges in a simple graph having appropriate connectivity conditions.

3.1 Linear $h(n)$

We first consider the regime in which $h(n)$ is linear in n. In fact, the bound we derive below holds for all legitimate $h(n)$, but it is asymptotically optimal only for linear $h(n)$. In the subsequent subsection, we will give a bound that holds for a smaller range of $h(n)$, but is asymptotically optimal for sublinear $h(n)$.

Theorem 3. *For any positive $h(n)$ that is bounded above everywhere by $\frac{n}{3}$,*

$$T_n(h(n)) \geq \frac{3}{4}n(1-o(1)).$$

PROOF. Let $H = (V,E)$ be a 3-uniform hypergraph on n vertices. Consider a coloring A,B,C of V, for which $|C| \leq n/3$. To satisfy $h(n)$-disjointness, H must contain a hyperedge that is trichromatic in A,B,C. In particular, for any bisection $(S,\bar{S})$ of $V \setminus C$, there is an edge in H_C crossing $(S,\bar{S})$. We will use the following lemma, with proof deferred to the full version.

Lemma 4. *For* ***graph*** *$G = (V,E)$, $|V| = n$, if all bisections are crossed by at least one edge, then $|E| \geq n/2$.*

To apply Lemma 4, observe that since $|V_C| \geq 2n/3$, we have $|E_C| \geq n/3$. Now we sum over all $|C| = n/3$.

$$\sum_{\substack{C\subset V\\ |C|=n/3}} |E_C| \geq \binom{n}{n/3} n/3. \qquad (1)$$

In order to turn this into a bound on $|E|$, we need to upper bound the extent to which hyperedges in E are overcounted in (1). A hyperedge in E induces a (single) edge in E_C only if one of its vertices is in C, and two of them are not. For a given hyperedge in E, there are three possible vertices that could be in C. Conditioned on that vertex being in C, and the two other vertices being outside C, there are at most $\binom{n-3}{n/3-1}$ ways to choose the remaining vertices of C. Hence each edge in E contributes 1 to $|E_C|$ for at most $3 \cdot \binom{n-3}{n/3-1}$ distinct C.

Dividing out the maximum contribution of each edge gives our desired lowered bound:

$$\begin{aligned}
|E| &\geq \frac{\binom{n}{n/3}n/3}{3\cdot\binom{n-3}{n/3-1}} \\
&= \frac{n!(n)(n/3-1)!(2n/3-2)!}{9(n-3)!(n/3)!(2n/3)!} \\
&= \frac{(n-1)(n-2)}{4n/3-2} \\
&= \frac{3}{4}n(1-o(1)).
\end{aligned}$$

□

3.2 Sublinear $h(n)$

Theorem 5. *For any function $h(n)$ that is bounded above everywhere by $\frac{n}{6}$,*

$$T_n(h(n)) \geq \Omega_n\left(\frac{n^3}{h(n)^2}\right).$$

PROOF. First we will need a (weakened) generalization of Lemma 4.

Definition 5. *Let $G=(V,E)$ be a graph on n vertices. For integers a,b, we say that G is* (a,b)-crossing *if for all disjoint $X,Y \subseteq V$ such that $|X|=a$ and $|Y|=b$, there is an edge from X to Y. (That is, $\exists x \in X, y \in Y : (x,y) \in E$.)*

Lemma 6. *For positive $i \le n/2$, every $(i,n/2)$-crossing graph on n vertices has at least $\frac{n^2}{2(n-i)}\left(\frac{n}{2i}-1\right)$ edges.*

PROOF. Note that the bound is vacuous for $i=n/2$, so assume $i \le n/2-1$. First observe that every subset of size i must have at least $n/2-i+1$ edges leaving it. This can be seen by contradiction: assume that some set X of size i has at most $n/2-i$ edges leaving it, and hence at most $n/2-i$ vertices in its neighborhood. Then take as Y the set of vertices in $V \setminus X$ with no edge to X. This set is of size at least $n-|X|-(n/2-i)=n/2$. Since there is no edge from X to Y, the graph cannot be $(i,n/2)$-crossing.

We now show that the lemma in fact holds for any graph having the above boundary property. There are $\binom{n}{i}$ vertex sets of size i. We count at least $n/2-i+1$ edges out of each set. Each edge can only be counted for the sets that it leaves; there are $2\binom{n-2}{i-1}$ of these, because we must choose one of the two vertices, not choose the other one, and choose $i-1$ other vertices. Hence, the total number of edges is at least

$$\begin{aligned}\frac{\binom{n}{i}}{2\binom{n-2}{i-1}}(n/2-i+1) &= \frac{n!(i-1)!(n-i-1)!}{2i!(n-i)!(n-2)!}(n/2-i+1)\\ &= \frac{n(n-1)}{2i(n-i)}(n/2-i+1)\\ &= \frac{n}{2(n-i)}\left((n-1)\left(\frac{n}{2i}-1\right)+\frac{n-1}{i}\right)\\ &> \frac{n}{2(n-i)}\left((n-1)\left(\frac{n}{2i}-1\right)+\frac{n}{2i}-1\right)\\ &= \frac{n^2}{2(n-i)}\left(\frac{n}{2i}-1\right).\end{aligned}$$

□

Now consider an $h(n)$-disjoint hypergraph $H=(V,E)$ on n vertices. For convenience, let $h=h(n)$. Since H is h-disjoint, there must exist a hyperedge for every A,B,C satisfying the conditions of Definition 2. In particular, consider a $C \subseteq V$ of size $|C|=h$. As in Section 3.1, we will show that the graph projection of H by C has many edges.

For $C \subseteq V$ having size $|C|=h$, let $G_C=(V_C,E_C)$ be the projection of H by C. We bound the size of each $|E_C|$. H is h-disjoint, so for any disjoint $A,B \subseteq V \setminus C$ such that $|A|,|B| \ge h$ and $|A|+|B| \ge \frac{n+3h}{2}-h=\frac{n+h}{2}$, there exists an edge $(x,y,z) \in E$ with $x \in A, y \in B, z \in C$. In other words, graph G_C has the following property: for $D \subseteq V_C$ with $|D| \ge \frac{n+h}{2}$, for all $A \subseteq D, B=D \setminus A$ with $|A|,|B| \ge h$, there is an edge in E_C from A to B.

When $h=1$, the above says that each subset of V_C of size $\frac{n+1}{2}$ is connected. In generality, we would like to lower bound the number of edges in $|E_C|$, which we can do with Lemma 6. First let $n'=|V_C|$, so $n'=n-h$. Observe that G_C is $(h,n'/2)$-crossing, by choosing A as any set of size h, B as any disjoint set of size $\frac{n'}{2}$, and $D=A \cup B$, so that $|D|=|A|+|B|=h+\frac{n'}{2}=h+\frac{n-h}{2}=\frac{n+h}{2}$. Hence $|E_C| \ge \frac{n'^2}{2(n'-h)}\left(\frac{n'}{2h}-1\right)$.

Now we must bound the extent to which each hyperedge is overcounted. A hyperedge can only be counted towards a given E_C if exactly one of its vertices is contained in C. There are then $\binom{n-3}{h-1}$ ways to choose the rest of the vertices. So each hyperedge contributes to $|E_C|$ for at most $3\binom{n-3}{h-1}$ values of C.

There are exactly $\binom{n}{h}$ sets C. Hence the total number of edges in H must be at least

$$|E| \ge \frac{\binom{n}{h}}{3\binom{n-3}{h-1}}\left(\frac{n'^2}{2(n'-h)}\right)\left(\frac{n'}{2h}-1\right).$$

We have assumed that $h \le n/6$, so $\frac{3}{2} \le \frac{n}{4h}$ and hence

$$\frac{n'}{2h}-1=\frac{n-h}{2h}-1=\frac{n}{2h}-\frac{3}{2} \ge \frac{n}{4h}.$$

Similarly,

$$\frac{n'^2}{2(n'-h)}=\frac{(n-h)^2}{2(n-2h)} \ge \frac{(n-(n/6))^2}{2(n-2)}=\frac{25n^2}{72(n-2)}.$$

Then

$$\begin{aligned}|E| &\ge \frac{\binom{n}{h}}{3\binom{n-3}{h-1}}\left(\frac{25n^3}{288h(n-2)}\right)\\ &= \frac{n!(h-1)!(n-2-h)!}{3h!(n-3)!(n-h)!}\left(\frac{25n^3}{288h(n-2)}\right)\\ &\ge \frac{25n^4}{864h^2(n-2)}=\Omega\left(\frac{n^3}{h^2}\right).\end{aligned}$$

□

4. UPPER BOUNDS

In this section, we give an asymptotically tight upper bound on $T_n(h(n))$ for almost all n, and all $1 \le h(n) \le n/3$. We do this by constructing near-Ramanujan expander graphs and converting them to "lifted" hypergraphs with expander-like properties. Our construction hence depends on the existence of sufficiently good expanders. These are probabilistically guaranteed to exist for all n; recall, however, that we are primarily interested in explicit constructions. Our result is fully constructive, with the exception that it relies on expander graphs that can be explicitly constructed for an infinite but incomplete set of values of n. As such, our result is only fully constructive for these n, which we do not consider a substantial weakness. To 'fill in the missing values' would require advances in explicit expander construction, which would immediately imply corresponding extensions of our algorithm.

Much of the difficulty of our analysis comes in explicitly bounding the degree. This is necessary to achieve an eigenvalue gap that can guarantee edges are well-distributed enough to induce a hyperedge across all "reasonable" colorings.

In what follows, an *algebraic* (n,d,λ)*-expander* will refer to an n-vertex, d-regular graph whose adjacency matrix has $\max(|\lambda_2|, |\lambda_n|)=\lambda$, where $\lambda_1 \ge \ldots \ge \lambda_n$ are the matrix eigenvalues.

Definition 6. *For graph $G=(V,E)$, we define a* lifted *3-uniform hypergraph $L(G)=(V,E')$ as follows. The edge set E' contains (x,y,z) if and only if at least two of the edges (x,y), (y,z), and (x,z) are present in G.*

In other words, for a given vertex x in G, we make hyperedges out of x and every pair of its neighbors. We claim that for an (n,d,λ)-expander G with the right parameters, hypergraph $H=L(G)$ is $h(n)$-disjoint and has a number of edges given by:

Theorem 7. $T_n(h(n)) \le O\left(\frac{n^3}{h(n)^2}\right)$.

PROOF. We construct a lifted 3-uniform hypergraph $H = L(G)$, where G is an (n,d,λ)-expander. Our goal is to determine the minimum λ such that H is $h(n)$-disjoint, as a function of d. Then using an expander G for which an upper bound on λ is known, we can derive a sufficient lower bound on d and hence the number of hyperedges in H.

To demonstrate $h(n)$-disjointness, we consider each partial 3-coloring A,B,C satisfying the conditions of Definition 2, and show that H contains a trichromatic edge for each such coloring. By the construction of H, it suffices to show that some set of vertices in C has edges in G to both A and B. Our main tool will be the Expander Mixing Lemma, which states that if $G = (V,E)$ is an (n,d,λ)-expander, then for any $S,T \subseteq V$, $\left||E(S,T)| - \frac{d|S|\cdot|T|}{n}\right| \le \lambda\sqrt{|S|\cdot|T|}$. Additionally, we will need the following variant of the expander mixing lemma. We have not found this precise lemma in the literature, so we give a proof of it below.

Lemma 8 (Expander Vertex-Boundary Lemma)**.** *Let $G = (V,E)$ be an (n,d,λ)-expander. Then for any sets $S,T \subseteq V$,*

$$|S \cap N(T)| \ge |S| - \frac{2\lambda n}{d}\sqrt{\frac{|S|}{|T|}},$$

where $N(T)$ is the set of vertices having a neighbor in T.

PROOF. The intuition is that if S is large, then the subset of S with edges to T cannot be small, because by the expander mixing lemma, a small set would have a small number of edges to T, but S must have a large number of edges to T. Let $S_T = S \cap N(T)$ and $s = |S|, t = |T|, s_T = |S_T|$. Then by definition $E(S,T) = E(S_T,T)$. By the expander mixing lemma, we have:

$$|E(S,T)| \ge \frac{d}{n}st - \lambda\sqrt{st} \quad \text{and} \quad |E(S_T,T)| \le \frac{d}{n}s_T t + \lambda\sqrt{s_T t}.$$

Combining the inequalities and solving for s_T gives:

$$\begin{aligned} s_T &\ge s - \lambda\frac{n}{d\sqrt{t}}(\sqrt{s}+\sqrt{s_T}) \\ &\ge s - \frac{2\lambda n}{d}\sqrt{s/t}, \end{aligned}$$

where the last inequality follows because $s_T \le s$. □

Using these tools, we prove the following main lemma:

Lemma 9. *Let $A,B,C \subseteq V$ be colors of sizes $a \le b \le c$, respectively, of the vertices of $H = L(G)$. Define $\mathscr{F}(a,b,c) = \sqrt{\frac{cb}{a}}\left(\sqrt{a+b}-\sqrt{b}\right)$. If*

$$\lambda < \frac{d}{n}\mathscr{F}(a,b,c),$$

then H contains a trichromatic edge.

Before proving the lemma, we show how it implies the theorem. By picking a G with $\lambda < \frac{d}{n}\mathscr{F}(a,b,c)$, we ensure that H contains a trichromatic edge for all colorings A,B,C. But the conditions of h-disjointness do not require all colorings to have this property, and in particular we can show:

Claim 10. *For $a \le b \le c$ satisfying the conditions of Definition 2, $\mathscr{F}(a,b,c) \ge k\sqrt{h(n-h)})$ for a fixed constant $k > 0$.*

PROOF. We wish to show that:

$$\sqrt{\frac{cb}{a}}\left(\sqrt{a+b}-\sqrt{b}\right) \ge k\sqrt{n(n-h)}$$

for some constant $k > 0$. To do this, we will show that $\sqrt{c} \ge k_1\sqrt{n-h}$ and $\sqrt{\frac{b}{a}}\left(\sqrt{a+b}-\sqrt{b}\right) \ge k_2\sqrt{h}$ for constants $k_1,k_2 > 0$, from which the theorem follows (with $k = k_1k_2$) because the LHS and RHS are non-negative in both inequalities. The conditions $a \le b \le c$ and $a+b+c = \frac{n+3h}{2}$ imply that $c \ge \frac{1}{3}\left(\frac{n+3h}{2}\right)$. Since both sides are positive, we have:

$$\sqrt{c} \ge \sqrt{\frac{1}{6}(n+3h)} > \sqrt{\frac{1}{6}(n-h)} = k_1\sqrt{n-h}$$

for $k_1 = \sqrt{\frac{1}{6}}$.

To lower bound the $\sqrt{\frac{b}{a}}\left(\sqrt{a+b}-\sqrt{b}\right)$ expression, we relax the $c \ge b$ and $a+b+c = \frac{n+3h}{2}$ constraints, and leave only the constraints $b \ge a$ and $a \ge h$. Now, since $a,b > 0$, we have for fixed a:

$$\begin{aligned} \frac{d}{db}\left(\sqrt{\frac{b}{a}}\left(\sqrt{a+b}-\sqrt{b}\right)\right) &= \frac{d}{db}\left(\sqrt{\frac{b^2}{a}+b}-\frac{b}{\sqrt{a}}\right) \\ &= \frac{\frac{2b}{a}+1}{2\sqrt{\frac{b^2}{a}+b}} - \frac{1}{\sqrt{a}} \\ &= \frac{2+\frac{a}{b}}{2\sqrt{a}\sqrt{1+\frac{a}{b}}} - \frac{1}{\sqrt{a}} \\ &= \frac{2+\frac{a}{b}-2\sqrt{1+\frac{a}{b}}}{2\sqrt{a}\sqrt{1+\frac{a}{b}}} \\ &= \frac{\left(1-\sqrt{1+\frac{a}{b}}\right)^2}{2\sqrt{a}\sqrt{1+\frac{a}{b}}} \end{aligned}$$

which is always positive, indicating that the function is monotonically increasing in b. Thus in order to minimize the function for fixed a, we choose b as small as possible, namely $b = a$. This gives:

$$\sqrt{\frac{b}{a}}\left(\sqrt{a+b}-\sqrt{b}\right) \ge \left(\sqrt{2}-1\right)\sqrt{a} \ge k_2\sqrt{h(n)}$$

for $k_2 = \sqrt{2}-1$. Thus the claim holds for $k = k_1k_2 = \frac{\sqrt{2}-1}{\sqrt{6}}$. □

Thus it suffices to construct a G with $\lambda < \frac{dk}{n}\sqrt{h(n-h)}$. A Ramanujan graph has $\lambda \le 2\sqrt{d-1} < 2\sqrt{d}$ and hence can be used if $2\sqrt{d} < \frac{dk}{n}\sqrt{h(n-h)}$. Rearranging gives $d > \frac{4n^2}{k^2h(n-h)}$, which is satisfied if $d > \frac{6n}{k^2h}$, since $h \le n/3$. That is, if G is Ramanujan with $d = \Theta\left(\frac{n}{h}\right)$, then $H = L(G)$ is h-disjoint. Since H has a hyperedge for every pair of edges from a given vertex in G, H has maximum degree $O\left(\frac{n^2}{h^2}\right)$ and thus at most $O\left(\frac{n^3}{h^2}\right)$ hyperedges.

To prove the bound, it suffices to assert the existence of Ramanujan graphs for every n and d. In fact, a much stronger theorem holds: for every $\varepsilon > 0$ and even $d \ge 4$, a random d-regular graph on n vertices satisfies $\lambda \le 2\sqrt{d-1}+\varepsilon$ with high probability [24]. Both ε and the requirement that d be even have an insubstantial effect on the final number of edges. □

We now give a proof of the main lemma. The idea is to first apply the expander vertex-boundary lemma to A and C, then the expander mixing lemma to B and the subset of C with neighbors in A. In so doing, we certify that A contains a vertex with edges to both B and C, in G. By the definition of a lifted hypergraph, this ensures that H contains a hyperedge crossing all three colors. Since we show this for arbitrary A,B,C satisfying the size bounds of Definition 2, this verifies the h-disjointness of H.

PROOF OF LEMMA 9. Let $C_A = C \cap N(A)$ and $a = |A|, b = |B|$, $c = |C|, c_A = |C_A|$. By Lemma 8,

$$c_A \geq c - \frac{2\lambda n}{d}\sqrt{\frac{c}{a}}. \tag{2}$$

To prove the existence of a trichromatic edge in H, it suffices to show that $|E_G(C_A, B)| > 0$. By the expander mixing lemma,

$$|E_G(C_A,B)| \geq \frac{dbc_A}{n} - \lambda\sqrt{bc_A}.$$

Hence there exists an edge from C_A to B when $\frac{dbc_A}{n} > \lambda\sqrt{bc_A}$, which is equivalent to $c_A > \frac{\lambda^2 n^2}{bd^2}$, because all variables are non-negative. Substituting c_A with (2) and solving gives $\frac{n^2}{bd^2}\lambda^2 + \frac{2n}{d}\sqrt{\frac{c}{a}}\lambda - c < 0$. By the quadratic equation, this is equivalent to

$$\left(\lambda - \frac{\sqrt{c}bd}{n}\left(\sqrt{\frac{1}{b}+\frac{1}{a}} - \sqrt{\frac{1}{a}}\right)\right)\cdot$$
$$\left(\lambda + \frac{\sqrt{c}bd}{n}\left(\sqrt{\frac{1}{b}+\frac{1}{a}} + \sqrt{\frac{1}{a}}\right)\right) < 0.$$

Because λ is positive, the LHS is negative when the first term is negative. Thus we need:

$$\begin{aligned}\lambda &< \frac{\sqrt{c}bd}{n}\left(\sqrt{\frac{1}{b}+\frac{1}{a}} - \sqrt{\frac{1}{a}}\right)\\ &= \frac{\sqrt{c}bd}{n}\frac{\sqrt{a+b}-\sqrt{b}}{\sqrt{ab}}\\ &= \frac{d}{n}\sqrt{\frac{cb}{a}}\left(\sqrt{a+b}-\sqrt{b}\right).\end{aligned}$$

This concludes the proof of the lemma, and hence of Theorem 7. □

We also give a (slightly less general) explicit construction of such hypergraphs.

Theorem 11. *There is an algorithm that, for an infinite number of integers n, and any $h(n)$ bounded above by $n/3$, efficiently constructs an $h(n)$-disjoint hypergraph with $O\left(\frac{n^3}{h(n)^2}\right)$ hyperedges.*

In other words, by applying an explicit Ramanujan construction, we constructively achieve the result of Theorem 7, for an infinite number of values of n.

PROOF. Extending the classic works of Lubotzky-Phillips-Sarnak [33], Margulis [35], and Morgenstern [36] on explicit constructions of Ramanujan graphs, Cioabă and Murty [13] give a construction that comes very close to the Ramanujan bound for nearly any graph size and degree.

Theorem 12 (Cioabă and Murty [13])**.** *Let $d \in \mathbb{Z}^+$ be such that $d-1$ is composite. For any positive ε, there exists an infinite sequence of graphs $\{G_i\}_{i=0}^{\infty}$ such that G_i is an $(n_i, d, (2+\varepsilon)\sqrt{d-1})$-expander, and $n_i > n_{i-1}$ $\forall$ $i > 0$.*

Recall that Lemma 9 and Claim 10 together imply it suffices to construct an (n,d,λ)-expander G, where $\lambda < \frac{dk}{n}\sqrt{h(n)(n-h(n))}$ for a fixed k.

Pick $d = \frac{2(2+\varepsilon)^2 n}{k^2 h(n)}$. Then [13] gives an algorithm to construct (n,d,λ)-expanders with $\lambda = \frac{(2+\varepsilon)^2}{k}\sqrt{\frac{2n}{h(n)}}$. Then observe:

$$\begin{aligned}\frac{dk}{n}\sqrt{h(n)(n-h(n))} &= \frac{2(2+\varepsilon)^2 nk}{k^2 h(n) n}\sqrt{h(n)(n-h(n))}\\ &\geq \frac{2(2+\varepsilon)^2}{k}\sqrt{\frac{n(2/3)}{h(n)}}\\ &> \frac{(2+\varepsilon)^2}{k}\sqrt{\frac{2n}{h(n)}} = \lambda,\end{aligned}$$

proving the theorem. □

5. A SUFFICIENCY CONDITION FOR h-DISJOINTNESS

In this section we consider a complementary question to that of the previous sections. Namely: how many hyperedges are necessary to ensure that *every* 3-uniform hypergraph of that size is $h(n)$-disjoint? Equivalently, what is the densest 3-uniform hypergraph that is *not* $h(n)$-disjoint? This question is relevant in practice, as it may be impossible in some systems to implement the set of 3-hyperedges exactly. The theorem below gives guarantees on reliability in such an oblivious setting.

Definition 7. *For integer $h \leq n/3$, the* sufficiency number *$U_n(h)$ is the minimum integer such that, for a 3-uniform hypergraph $H = (V,E)$ on n vertices, $|E| \geq U_n(h)$ implies that H is h-disjoint.*

Theorem 13. *For $h \leq n/3$,*

$$U_n(h) = \binom{n}{3} - \frac{n-h}{2}\cdot h^2 + 1.$$

In other words, $\frac{n-h}{2}\cdot h^2$ is the minimum number of edges one can remove from the complete 3-uniform n-vertex hypergraph, in order to ensure it is not h-disjoint.

PROOF. Let $H = (V,E)$ be an n-vertex 3-uniform hypergraph that is not h-disjoint. By definition, there must be some partial coloring A,B,C of the vertices with $a = |A|, b = |B|, c = |C|$, such that there is no edge crossing A,B,C, and moreover $a,b,c \geq h$ and $a+b+c \geq \frac{n+3h}{2}$. For any 3-coloring, the complete 3-uniform hypergraph contains exactly abc crossing hyperedges. Hence for some a,b,c having the properties above, abc is the smallest number of edges that can be removed from the complete graph to make it not h-disjoint.

Claim 14. *For integers $a,b,c \geq h$ such that $a+b+c \geq \frac{n+3h}{2}$, abc is minimized by taking $a = h, b = h, c = \frac{n-h}{2}$.*

PROOF. We will assume for the proof that $n \equiv h \pmod 2$. The second case is proved similarly.

First note that since $a+b+c \geq \frac{n+3h}{2}$, abc is minimized by taking $a+b+c = \frac{n+3h}{2}$. Decreasing the sum can always decrease the product. Hence, we may assume w.l.o.g. that $c = \frac{n+3h}{2} - a - b$, and minimize $g(a,b) = ab\left(\frac{n+3h}{2} - a - b\right)$. This gives the following optimization problem, for arbitrary n and h:

$$\begin{aligned} &\text{minimize} \quad g(a,b) \\ &\text{subject to} \quad a \geq h, \qquad b \geq h, \qquad a+b \leq \frac{n+h}{2} \end{aligned}$$

The constraints are linear and hence define halfspaces in the (a,b)-plane. These halfspaces define P, a polytope (triangle) in which the solution must lie. In particular, the optimal solution must either be a global minimum of $g(a,b)$ (and hence a root of the gradient); a minimum along one of the faces of the polytope; or a vertex of the polytope. We check each case in turn.

Proposition 15. *The following three facts about the constrained optima of $g(a,b)$ hold.*

- *The gradient of g has a single root inside P, and it is a global maximum.*
- *The minimum value of g along the faces of P is $\frac{(n+h)^2}{16}h$.*
- *The minimum value of g at a vertex of P is $\frac{h^2(n-h)}{2}$.*

The proof of the proposition appears in the appendix. We now observe how it implies the theorem.

Comparison. There are only two possible minima in the polytope: $\frac{h(n+h)^2}{16}$ and $\frac{h^2(n-h)}{2}$. Observe that

$$\begin{aligned} \frac{(n+h)^2}{16} &= \frac{n^2+h^2+2nh}{16} \\ &= \frac{(n^2+9h^2-6nh)-8h^2+8nh}{16} \\ &= \frac{(n-3h)^2}{16} + \frac{h(n-h)}{2} \\ &\geq \frac{h(n-h)}{2}. \end{aligned}$$

Therefore $\frac{h(n+h)^2}{16} \geq \frac{h^2(n-h)}{2}$, so $\frac{h^2(n-h)}{2}$ is the minimum of the constrained $g(a,b)$. Recall that this was obtained by setting two faces to tight. In other words, set any two of a,b,c to h, and the other to $\frac{n-h}{2}$. □

By the claim, removing $\frac{h^2(n-h)}{2}$ edges from the complete 3-uniform hypergraph ensures that a given valid coloring has no trichromatic edge. As a result, the hypergraph cannot be h-disjoint. Conversely, removing fewer edges cannot remove all the edges crossing any valid coloring. Hence $U_n(h) = \binom{n}{3} - \frac{n-h}{2} \cdot h^2 + 1$.

This completes the proof of Theorem 13. □

6. HARDNESS OF DECIDING h-DISJOINTNESS

In this section, we take a first step towards addressing algorithmic questions related to h-disjointness. In particular, given a 3-uniform hypergraph H, we would like to determine the minimum value $h_{opt}(n)$ such that H is $h_{opt}(n)$-disjoint, or barring this, an approximate value h that is as close to $h_{opt}(n)$ as possible. This question has practical value because it allows us to evaluate an existing hypergraph, or perhaps one constructed via the methods described in Section 4, for h-disjointness. Here we show that deciding whether H is h-disjoint is *co-NP*-complete.

Theorem 1 (restated). *Given a 3-uniform hypergraph $H = (V,E)$ with $|V| = n$, it is* co-NP*-complete to decide, for integers $h \leq n/3$, whether H is h-disjoint.*

PROOF. The complement problem to h-disjointness is that of finding a disjoint $A,B,C \subseteq V$ satisfying the conditions of Definition 2 such that $\forall x \in A, y \in B, z \in C$, it holds that $(x,y,z) \notin E$. A certificate for this problem is the sets A,B,C, and it can be verified in $O(|A||B||C|)$ time by checking that all hyperedges (x,y,z) are not in E. Since complement h-disjointness is in *NP*, it follows that h-disjointness is in *co-NP*.

We show that complement h-disjointness is *NP*-hard by a reduction from balanced bipartite independent set (BBIS), which is *NP*-complete [3]. Given a balanced bipartite graph $G(X \cup Y; E)$ with $|X| = |Y| = n/2$ and a positive integer t, the decision BBIS problem is to find sets $A \subseteq X$, $B \subseteq Y$ with $|A| = |B| = t$ with no edges between A and B. Given an instance of the BBIS problem, we construct an instance of complement h-disjointness as follows. Create an empty 3-uniform hypergraph H with $n' = n + (n-t)$ vertices, where the first n vertices represent the vertices of G. On the $n-t$ vertices, create a complete 3-uniform hypergraph Z. Add a hyperedge (u,v,w) for each pair of vertices $u,v \in Z$ and every $w \in G$. Add a hyperedge (u,v,w) for each pair of vertices $u,v \in X$ and every $w \in Z$; do the same for every pair of vertices $u,v \in Y$. Finally, add a hyperedge (u,v,w) for each *edge* $(u,v) \in G$ and every $w \in Z$. The input to the complement h-disjointness problem is the hypergraph H and the positive integer $h = t$.

Given a solution (A,B) with $|A| = |B| = t$ to BBIS, we claim that A,B,Z is a solution to complement h-disjointness. Since $t \leq |X| = |Y| = n/2$, it follows that $|Z| = n-t \geq t$, so all $|A|,|B|,|Z| \geq t$. Also, $|A|+|B|+|Z| = t+t+(n-t) = \frac{n+(n-t)+3t}{2} = \frac{n'+3t}{2}$. Now, for a hyperedge to cross the sets A,B,Z, there must be some $u \in A, v \in B, w \in Z$ such that $(u,v,w) \in H$. By our construction, all but one type of hyperedge in H involve vertices in at most two of the sets $A \subseteq X$, $B \subseteq Y$, and Z. The exception are hyperedges (u,v,w) where $(u,v) \in G$ and $w \in Z$. But since there are no edges crossing A,B, there is no hyperedge that crosses A,B,Z. Thus A,B,Z is a solution to complement h-disjointness.

In the reverse direction, suppose we have a solution A,B,C to complement h-disjointness. Since $|A|+|B|+|C| \geq \frac{n'+3t}{2} = n+t$, some vertices in Z must appear in the sets A,B,C. We claim that these vertices must appear in exactly one set. This is because if Z appears in all three sets, then there would exist a hyperedge crossing all three sets, since Z is a complete 3-uniform hypergraph. Similarly, Z cannot appear in exactly two sets, because then there would exist a hyperedge connecting two vertices of Z (one in each set) and a vertex in the third set (a vertex in G), also by our construction. Thus Z participates in exactly one set; assume w.l.o.g. that this set is C. We now claim that the vertices in X appear in at most two sets, and if they appear in two sets, one of those sets must be C. If X appears in both A and B, then there would exist a hyperedge connecting two vertices of X (one in each set) to a vertex of C, because C contains at least some vertices of Z by our argument above. Therefore, X appears in either A or B, but not both. The same argument shows that this is also true of Y. Combining these arguments with the fact that A,B,C are non-empty, it follows that the vertices of X and Y are split across A,B (though they may appear together in C). Since Z appears in C, A and B consist entirely of vertices in G. Finally, the same argument used in the forward direction above shows that there cannot exist an edge $(u,v) \in G$ between A and B, since then there would exist a hyperedge (u,v,w) to a vertex $w \in Z$ in C. Thus, since $|A|,|B| \geq t$, we can remove excess vertices so that $|A| = |B| = t$ and the resulting sets are a solution to BBIS. □

7. CONCLUSIONS & OPEN PROBLEMS

This paper studies the price of equivocation in distributed systems. Our tight bounds on the number of 3-processor partial broadcast channels required for Byzantine agreement describe the amount of equivocation a system can tolerate for a given level of redundancy. Our results thus capture the *equivocation vs. redundancy trade-off*, an important metric in the cost-benefit analysis of a fault-tolerant system.

Several interesting theoretical questions remain. For example, given the hardness of deciding a system's resilience (h-disjointness) based on its partial broadcast channels, we are interested in approximation algorithms for this value. We would also like to understand the combinatorial properties of h-disjointness in greater depth. Can it be shown that h-disjoint hypergraphs *must* fundamentally be built on an underlying expander? How do our definitions and results scale to k-uniform hypergraphs, for $k > 3$?

On the practical side, it would be very interesting to find a realistic network that achieves h-disjointness naturally based on its broadcast (*e.g.*, network hub) and point-to-point (*e.g.*, network switch) connections, instead of constructing partial broadcast channels explicitly.

References

[1] I. Abraham, M. K. Aguilera, and D. Malkhi. Fast asynchronous consensus with optimal resilience. In *24th International Symposium on Distributed Computing (DISC)*, pages 4–19, 2010.

[2] A. S. Aiyer, L. Alvisi, A. Clement, M. Dahlin, J.-P. Martin, and C. Porth. BAR fault tolerance for cooperative services. In *20th Symposium on Operating Systems Principles (SOSP)*, pages 45–58, 2005.

[3] Alon, Duke, Lefmann, Rodl, and Yuster. The algorithmic aspects of the regularity lemma. *J. Algorithms*, 16, 1994.

[4] L. Alvisi, A. Clement, M. Dahlin, M. Marchetti, and E. Wong. Making Byzantine fault tolerant systems tolerate Byzantine faults. In *6th Symposium on Networked Systems Design and Implementation (NSDI)*, pages 153–168, 2009.

[5] S. Amitanand, I. Sanketh, K. Srinathant, V. Vinod, and C. P. Rangan. Distributed consensus in the presence of sectional faults. In *22nd Symposium on Principles of Distributed Computing (PODC)*, pages 202–210, 2003.

[6] J. L. Arocha and J. Tey. The size of minimum 3-trees. *J. Graph Theory*, 54(2):103–114, 2007.

[7] J. L. Arocha, J. Bracho, and V. Neumann-Lara. On the minimum size of tight hypergraphs. *J. Graph Theory*, 16(4): 319–326, 1992.

[8] G. Bracha. An asynchronous $[(n-1)/3]$-resilient consensus protocol. In *3rd Symposium on Principles of Distributed Computing (PODC)*, pages 154–162, 1984.

[9] G. Bracha and S. Toueg. Asynchronous consensus and broadcast protocols. *J. ACM*, 32(4):824–840, 1985.

[10] C. Bujtas and Z. Tuza. Smallest set-transversals of k-partitions. *Graph. Comb.*, 25:807–816, 2009.

[11] M. Castro. *Practical Byzantine Fault-Tolerance*. PhD thesis, Massachusetts Institute of Technology, 2000.

[12] B.-G. Chun, P. Maniatis, S. Shenker, and J. Kubiatowicz. Attested append-only memory: Making adversaries stick to their word. In *21st Symposium on Operating Systems Principles (SOSP)*, pages 189–204, 2007.

[13] S. M. Cioaba. *Eigenvalues, expanders and gaps between primes*. PhD thesis, Queen's University, 2005.

[14] J. Considine, M. Fitzi, M. K. Franklin, L. A. Levin, U. M. Maurer, and D. Metcalf. Byzantine agreement given partial broadcast. *J. Cryptology*, 18(3):191–217, 2005.

[15] J. Cowling, D. Myers, B. Liskov, R. Rodrigues, and L. Shrira. HQ replication: A hybrid quorum rotocol for Byzantine fault tolerance. In *7th Symposium on Operating Systems Design and Implementation (OSDI)*, pages 177–190, 2006.

[16] K. Diao, P. Zhao, and H. Zhou. About the upper chromatic number of a co-hypergraph. *Discrete Math.*, 220:67–73, 2000.

[17] K. Diao, G. Liu, D. Rautenbach, and P. Zhao. A note on the least number of edges of 3-uniform hypergraphs with upper chromatic number 2. *Discrete Math.*, 306(7):670–672, 2006.

[18] C. Dwork, N. Lynch, and L. Stockmeyer. Consensus in the presence of partial synchrony. *J. ACM*, 35(2):288–323, 1988.

[19] M. J. Fischer, N. A. Lynch, and M. S. Paterson. Impossibility of distributed consensus with one faulty process. *J. ACM*, 32 (2):374–382, 1985.

[20] M. Fitzi and U. Maurer. From partial consistency to global broadcast. In *32nd Symposium on Theory of Computing (STOC)*, pages 494–503, 2000.

[21] M. Fitzi and U. M. Maurer. Efficient byzantine agreement secure against general adversaries. In *12th International Symposium on Distributed Computing (DISC)*, pages 134–148, 1998.

[22] M. Franklin and M. Yung. Secure hypergraphs: Privacy from partial broadcast (extended abstract). In *27th Symposium on the Theory of Computing (STOC)*, pages 36–44, 1995.

[23] M. K. Franklin and R. N. Wright. Secure communications in minimal connectivity models. In *Advances in Cryptology (EUROCRYPT)*, pages 346–360, 1998.

[24] J. Friedman. A proof of Alon's second eigenvalue conjecture. In *35th Symposium on Theory of Computing (STOC)*, pages 720–724, 2003.

[25] J. A. Garay and K. J. Perry. A continuum of failure models for distributed computing. In *6th International Workshop on Distributed Algorithms (WDAG)*, pages 153–165, 1992.

[26] A. Haeberlen, P. Kouznetsov, and P. Druschel. Peerreview: practical accountability for distributed systems. In *21st Symposium on Operating Systems Principles (SOSP)*, pages 175–188, 2007.

[27] R. Kapitza, J. Behl, C. Cachin, T. Distler, S. Kuhnle, S. V. Mohammadi, W. Schröder-Preikschat, and K. Stengel. CheapBFT: resource-efficient byzantine fault tolerance. In *European Conference on Computer Systems (EuroSys)*, pages 295–308, 2012.

[28] A. Karlin and A. C. Yao. Probabilistic lower bounds for byzantine agreement and clock synchronization. Unpublished manuscript, 1984.

[29] R. Kotla, L. Alvisi, M. Dahlin, A. Clement, and E. Wong. Zyzzyva: Speculative Byzantine fault tolerance. In *21st Symposium on Operating Systems Principles (SOSP)*, pages 45–58, 2007.

[30] L. Lamport. Lower bounds for asynchronous consensus. In *Future Directions in Distributed Computing*, pages 22–23,

2003.

[31] L. Lamport, R. Shostak, and M. Pease. The Byzantine generals problem. *ACM Trans. Program. Lang. Syst.*, 4(3): 382–401, 1982.

[32] D. Levin, J. R. Douceur, J. R. Lorch, and T. Moscibroda. TrInc: Small trusted hardware for large distributed systems. In *6th Symposium on Networked Systems Design and Implementation (NSDI)*, pages 1–14, 2009.

[33] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[34] N. A. Lynch. *Distributed Algorithms*. Morgan Kaufmann, 1996.

[35] G. A. Margulis. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Probl. Inf. Transm.*, 24(1): 39–46, 1988.

[36] M. Morgenstern. Existence and explicit constructions of $q+1$ regular ramanujan graphs for every prime power q. *J. Comb. Theory Ser. B*, 62(1):44–62, 1994.

[37] T. Rabin and M. Ben-Or. Verifiable secret sharing and multiparty protocols with honest majority (extended abstract). In *21st Symposium on Theory of Computing (STOC)*, pages 73–85, 1989.

[38] D. V. S. Ravikant, M. Venkitasubramaniam, V. Srikanth, K. Srinathan, and C. P. Rangan. On byzantine agreement over (2, 3)-uniform hypergraphs. In *18th International Symposium on Distributed Computing (DISC)*, pages 450–464, 2004.

[39] M. Serafini, P. Bokor, D. Dobre, M. Majuntke, and N. Suri. Scrooge: Reducing the costs of fast byzantine replication in presence of unresponsive replicas. In *41st International Conference on Dependable Systems and Networks (DSN)*, pages 353–362, 2010.

[40] F. Sterboul. An extremal problem in hypergraph coloring. *J. Comb. Theory Ser. B*, 22(2):159 – 164, 1977.

[41] V. I. Voloshin. On the upper chromatic number of a hypergraph. *Australas. J. Combin.*, 11:25–45, 1995.

[42] V. I. Voloshin. *Coloring Mixed Hypergraphs: Theory, Algorithms and Applications*, volume 17 of *Fields Institute Monographs*. AMS, 2002.

[43] T. Wood, R. Singh, A. Venkataramani, P. J. Shenoy, and E. Cecchet. ZZ and the art of practical BFT execution. In *6th European Conference on Computer Systems (EuroSys)*, pages 123–138, 2011.

[44] J. Yin, J.-P. Martin, A. Venkataramani, L. Alvisi, and M. Dahlin. Separating agreement from execution for Byzantine fault tolerant services. In *19th Symposium on Operating Systems Principles (SOSP)*, pages 253–267, 2003.

APPENDIX

Proof for Sufficiency Condition

PROOF OF PROPOSITION 15. We explore the three claims below.

Global minimum. The gradient of g has only a single root in the polytope.

$$\nabla g(a,b) = \left[\frac{b(n+3h)}{2} - 2ab - b^2, \frac{a(n+3h)}{2} - 2ab - a^2\right].$$

We may assume that $a \neq 0, b \neq 0$, since otherwise $[a,b]$ is not in the polytope. Hence $\frac{b(n+3h)}{2} - 2ab - b^2 = 0$ is equivalent to $a = \frac{n+3h}{4} - \frac{b}{2}$. Symmetrically, $b = \frac{n+3h}{4} - \frac{a}{2}$. So $a = \frac{n+3h}{4} - \left(\frac{n+3h}{8} - \frac{a}{4}\right)$, *i.e.* $\frac{3}{4}a = \frac{n+3h}{8}$, *i.e.* $a = \frac{n+3h}{6}$. Symmetrically, $b = \frac{n+3h}{6}$. Therefore the only root of the gradient that may be in the polytope is $\left[\frac{n+3h}{6}, \frac{n+3h}{6}\right]$.

We now show that $\left[\frac{n+3h}{6}, \frac{n+3h}{6}\right]$ must be a maximum by examining the second derivatives of $g(a,b)$. The second derivatives are

- $g_{aa}(a,b) = -2b = -\frac{n+3h}{3}$
- $g_{bb}(a,b) = -2a = -\frac{n+3h}{3}$
- $g_{ab}(a,b) = \frac{n+3h}{2} - 2a - 2b = -\frac{n+3h}{6}$.

The second derivative test says that if $g_{aa}(a,b)$ is negative, and $g_{aa}(a,b) \cdot g_{bb}(a,b) - g_{ab}(a,b)^2$ is positive, then $[a,b]$ is a local maximum. Indeed, $-\frac{n+3h}{3}$ is negative, and $\left(-\frac{n+3h}{3}\right)^2 - \left(-\frac{n+3h}{6}\right)^2$ is positive. Therefore the only root that may be within the polytope is a global maximum, so it cannot possibly be the minimum point of the polytope.

Faces. We consider the points along the faces of each constraint. In other words, we set the constraints to tight, then globally optimize the resulting function.

First make $a \geq h$ tight. $a = h$ means the new objective function is $hb\left(\frac{n+3h}{2} - h - b\right) = hb\left(\frac{n+h}{2} - b\right)$. We differentiate to find the optimum value of b for this function.

$$\frac{d}{db} hb\left(\frac{n+h}{2} - b\right) = \frac{h(n+h)}{2} - 2hb,$$

which is 0 only at $b = \frac{n+h}{4}$. Thus $g\left(h, \frac{n+h}{4}\right) = h\left(\frac{n+h}{4}\right) \cdot \left(\frac{n+h}{2} - \frac{n+h}{4}\right) = \frac{h(n+h)^2}{16}$ is a potential global minimum for g.

By symmetry, setting $b \geq h$ tight gives the same potential minimum.

Finally, set $a + b \leq \frac{n+h}{2}$ tight. Then the new objective function is $\left(\frac{n+h}{2} - b\right) b \left(\frac{n+3h}{2} - \left(\frac{n+h}{2} - b\right) - b\right) = \left(\frac{n+h}{2} - b\right) bh$. We differentiate to find the optimum value of b of this new function.

$$\frac{d}{db}\left(\frac{n+h}{2} - b\right) bh = \frac{h(n+h)}{2} - 2bh,$$

which is 0 only when $b = \frac{n+h}{4}$. Plugging this and $a = \frac{n+h}{4}$ (by symmetry) into g gives $g(a,b) = (\frac{n+h}{4})^2(\frac{n+3h}{2} - \frac{n+h}{2}) = \frac{(n+h)^2}{16} h$, the same value found for the other two constraints.

Vertices. We now consider the value of g at the three intersections of the three halfspaces. (Note that the faces only intersect in at most one point because they are unique and on two variables.)

Setting $a = h$ and $a + b = \frac{n+h}{2}$ gives $b = \frac{n-h}{2}$, hence $g(a,b) = h \cdot \frac{n-h}{2}\left(\frac{n+3h}{2} - h - \frac{n-h}{2}\right) = h \cdot \frac{n-h}{2} \cdot h$.

By symmetry, setting $b = h$ and $a + b = \frac{n+h}{2}$ also gives $g(a,b) = \frac{h^2(n-h)}{2}$.

Finally, setting $a = h$ and $b = h$ gives $g(a,b) = h^2\left(\frac{n+3h}{2} - 2h\right) = \frac{h^2(n-h)}{2}$, again. □