Combinatorics, Probability and Computing (2017) 26, 78–98. © Cambridge University Press 2016
doi:10.1017/S0963548316000122

Saturation in the Hypercube and Bootstrap Percolation

NATASHA MORRISON, JONATHAN A. NOEL and ALEX SCOTT

Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, UK
(e-mail: {morrison,noel,scott}@maths.ox.ac.uk)

Received 1 September 2014; revised 1 June 2015; first published online 31 March 2016

Let Q_d denote the hypercube of dimension d. Given $d \geqslant m$, a spanning subgraph G of Q_d is said to be (Q_d, Q_m)-*saturated* if it does not contain Q_m as a subgraph but adding any edge of $E(Q_d) \setminus E(G)$ creates a copy of Q_m in G. Answering a question of Johnson and Pinto [27], we show that for every fixed $m \geqslant 2$ the minimum number of edges in a (Q_d, Q_m)-saturated graph is $\Theta(2^d)$.

We also study weak saturation, which is a form of bootstrap percolation. A spanning subgraph of Q_d is said to be *weakly* (Q_d, Q_m)-*saturated* if the edges of $E(Q_d) \setminus E(G)$ can be added to G one at a time so that each added edge creates a new copy of Q_m. Answering another question of Johnson and Pinto [27], we determine the minimum number of edges in a weakly (Q_d, Q_m)-saturated graph for all $d \geqslant m \geqslant 1$. More generally, we determine the minimum number of edges in a subgraph of the d-dimensional grid P_k^d which is weakly saturated with respect to 'axis aligned' copies of a smaller grid P_r^m. We also study weak saturation of cycles in the grid.

2010 *Mathematics subject classification*: Primary 05C35
Secondary 05C50, 05D99

1. Introduction

Given graphs F and H, a spanning subgraph G of F is said to be (F, H)-*saturated* if it does not contain H as a subgraph, but for every edge $e \in E(F) \setminus E(G)$, the graph $G + e$ contains a copy of H. In this language, the classical Turán problem asks for the maximum size of an (F, H)-saturated graph; this number is known as the *extremal number*, denoted $\mathrm{ex}(F, H)$. Another well-studied problem, introduced independently by Zykov [37] and Erdős, Hajnal and Moon [20], is to determine the minimum number of edges in an (F, H)-saturated graph. This is known as the *saturation number*, denoted $\mathrm{sat}(F, H)$. A spanning subgraph G of F is said to be *weakly* (F, H)-*saturated* if the edges of $E(F) \setminus E(G)$ can be added to G, one edge at a time, in such a way that every added edge creates a new copy of H. This notion was introduced by Bollobás [9]. The minimum number of edges in a weakly (F, H)-saturated graph is known as the *weak saturation number* and denoted

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

$\mathrm{wsat}(F,H)$. Note that every (F,H)-saturated graph is also weakly (F,H)-saturated and so

$$\mathrm{wsat}(F,H) \leqslant \mathrm{sat}(F,H). \tag{1.1}$$

For additional background on minimum saturation in graphs, see [21].

Given $k \geqslant 2$ and $d \geqslant 1$, the *k-grid of dimension d* is the graph P_k^d with vertex set

$$\{0,1,\ldots,k-1\}^d,$$

where two vertices are adjacent if they differ by one in some coordinate and are equal in the other $d-1$ coordinates. In particular, P_2^d is known as the *hypercube of dimension d* and denoted Q_d. In this paper, we are interested in saturation and weak saturation problems in hypercubes and grids.

With regard to $\mathrm{sat}(Q_d,Q_m)$, the special case $m=2$ was studied by Choi and Guan [14] who constructed a (Q_d,Q_2)-saturated graph with at most $(1/4+o(1))|E(Q_d)|$ edges. Santolupo (see [21]) conjectured that this construction is best possible. In their recent paper [27], Johnson and Pinto disproved this conjecture (in a strong sense) by showing that, for every fixed m, there exists a (Q_d,Q_m)-saturated graph with $o(|E(Q_d)|)$ edges. More precisely, they proved the following.

Theorem 1.1 (Johnson and Pinto [27]). *For every fixed $m \geqslant 2$, there exists $0<\varepsilon_m<1$ such that*

$$\mathrm{sat}(Q_d,Q_m) = O(d^{(1-\varepsilon_m)}2^d) = O(|E(Q_d)|/d^{\varepsilon_m}).$$

In the case $m=2$, Johnson and Pinto [27] obtained a stronger bound, namely,

$$\mathrm{sat}(Q_d,Q_2) < 10\cdot 2^d. \tag{1.2}$$

That is, for every d, there exists a (Q_d,Q_2)-saturated graph with bounded average degree. Motivated by this result, they asked the following: for which fixed values of m is

$$\mathrm{sat}(Q_d,Q_m) = O(2^d)?$$

We show that this is the case for every $m \geqslant 2$.

Theorem 1.2. *For every fixed $m \geqslant 2$, $\mathrm{sat}(Q_d,Q_m) \leqslant (1+o(1))72m^2 2^d$.*

We remark that, previously, the best known lower bound on $\mathrm{sat}(Q_d,Q_m)$ was

$$(m+1-o(1))2^{d-1},$$

due to Johnson and Pinto [27]. In particular, for fixed m, Theorem 1.2 is tight up to a (constant) factor of $O(m)$.

In the case of weak saturation, Johnson and Pinto [27] proved

$$\mathrm{wsat}(Q_d,Q_2) = 2^d - 1 \tag{1.3}$$

for all $d \geqslant 2$ by exhibiting a spanning tree of Q_d which is weakly (Q_d,Q_2)-saturated. They asked about the value of $\mathrm{wsat}(Q_d,Q_m)$ for general $d \geqslant m \geqslant 1$. We answer this question.

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

Theorem 1.3. *For $d \geqslant m \geqslant 1$,*

$$\mathrm{wsat}(Q_d, Q_m) = (m-1)2^d - \sum_{j=0}^{m-2}(m-1-j)\binom{d}{j}.$$

By combining Theorem 1.3 with (1.1), we can improve the lower bound for $\mathrm{sat}(Q_d, Q_m)$ mentioned above to $(m-1-o(1))2^d$.

Theorem 1.3 is a special case of a more general result. Given $k \geqslant r \geqslant 2$ and $d \geqslant m \geqslant 1$, say that a copy of P_r^m in P_k^d is *axis aligned* if it is induced by a set of vertices of the form $I_1 \times \cdots \times I_d$ where exactly m of the sets I_i are intervals of length r in $\{0, \ldots, k-1\}$ and the rest are singletons. Let $\mathrm{wsat}^*(P_k^d, P_r^m)$ be the minimum number of edges in a spanning subgraph G of P_k^d such that the edges of $E(P_k^d) \setminus E(G)$ can be added to G, one edge at a time, such that every added edge creates an axis aligned copy of P_r^m. We prove the following.

Theorem 1.4. *For $k \geqslant r \geqslant 2$ and $d \geqslant m \geqslant 1$,*

$$\begin{aligned}\mathrm{wsat}^*(P_k^d, P_r^m) = &\sum_{j=0}^{d}\sum_{i=0}^{d-j}(m-1+i)\binom{d}{j}\binom{d-j}{i}(k-r+1)^j(r-2)^i \\ &- \sum_{j=0}^{m-2}\sum_{i=0}^{d-j}(m-1-j)\binom{d}{j}\binom{d-j}{i}(k-r+1)^j(r-2)^i,\end{aligned}$$

where, by convention, $0^0 = 1$.

We remark that if $r = 2$ or $d = m$, then every copy of P_r^m in P_k^d is axis aligned. Thus, in these cases, $\mathrm{wsat}^*(P_k^d, P_m^r) = \mathrm{wsat}(P_k^d, P_m^r)$. In particular, Theorem 1.3 is implied by the case $k = r = 2$ of Theorem 1.4.

We also consider an extension of (1.3) to general even cycles in the grid, proving the following.

Theorem 1.5. *For $k \geqslant 2$ and $d \geqslant \ell \geqslant 2$,* $\mathrm{wsat}(P_k^d, C_{2\ell}) = k^d - 1$.

The rest of the paper is structured as follows. At the beginning of Section 2, we review some fundamental properties of hypercubes which will be used in the proof of Theorem 1.2. Then, we prove Theorem 1.2 by giving an explicit construction of a (Q_d, Q_m)-saturated graph with bounded average degree. In Section 3 we turn our attention to weak saturation and prove Theorems 1.3, 1.4 and 1.5. We conclude in Section 4 by mentioning a number of open problems.

2. Minimum saturation in the hypercube

2.1. Preliminaries

Let e_j denote the jth standard basis vector in $\mathbb{F}_2^d$, that is, the vector in which the jth coordinate is equal to one and every other coordinate is zero. Given a vertex v of Q_d, let

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

$|v|$ denote the sum of its coordinates. A basic fact about hypercubes is that, for $d \geqslant m$, every copy Q of Q_m in Q_d is induced by a set of vertices of the form

$$\left\{ v + \sum_{j \in J'} e_j : J' \subseteq J \right\},$$

where v is a fixed vertex and $J \subseteq [d]$ has cardinality m. We say that the coordinates of J and $[d] \setminus J$ are *variable* and *fixed* under Q, respectively.

We will need a standard result from Coding Theory, due to Hamming [26] (for another reference, see [30, 31]). This result was also used by Johnson and Pinto [27] (and we use some ideas from [27] in our proof).

Theorem 2.1 (Hamming [26]). *There is an independent set $C \subseteq V(Q_{2^t-1})$ such that every vertex of $V(Q_{2^t-1}) \setminus C$ has a unique neighbour in C.*

A set C as in Theorem 2.1 is often referred to as a *Hamming code*. We remark that, since every vertex of Q_{2^t-1} has exactly $2^t - 1$ neighbours,

$$|C| = \frac{2^{2^t-1}}{(2^t-1)+1} = 2^{2^t-t-1}. \tag{2.1}$$

Another ingredient of our proof of Theorem 1.2 is the following result of Conder [16].

Theorem 2.2 (Conder [16]). *For every $s \geqslant 1$, there is a 3-colouring of the edges of Q_s in which there is no monochromatic cycle of length 4 or 6.*

2.2. Definitions and proof outline

Our construction will use Theorems 2.1 and 2.2 to build a spanning subgraph G of Q_d that contains no copy of Q_m and that can be extended to a graph G' which is (Q_d, Q_m)-saturated and satisfies the bound in Theorem 1.2.

Throughout the proof, let $m \geqslant 2$ be fixed and let d be an integer which we may choose to be sufficiently large. Let t and s be the unique integers such that

$$d = 6m(2^t - 1) + s \quad \text{and} \quad 0 \leqslant s < 6m2^t. \tag{2.2}$$

We view the set $[d]$ as a union of m intervals of length $6(2^t - 1)$ indexed by the elements of $[m]$ followed by one interval of length s. Given $v \in V(Q_d)$ and $1 \leqslant i \leqslant m$, let $v(i)$ denote the vertex of $V(Q_{6(2^t-1)})$ obtained by restricting v to the coordinates of the interval corresponding to i, and we let $v(m+1)$ be the vertex of $V(Q_s)$ obtained by restricting v to its last s coordinates. Each of the first m intervals of $[d]$ is further divided into 6 subintervals of length $(2^t - 1)$ indexed by the pairs (r, γ) where $r \in \{0, 1\}$ and $\gamma \in \{0, 1, 2\}$. For $(i, r, \gamma) \in [m] \times \{0, 1\} \times \{0, 1, 2\}$, let $v(i, r, \gamma)$ be the vertex of $V(Q_{2^t-1})$ obtained by restricting $v(i)$ to the subinterval corresponding to (r, γ).

Let C be a subset of $V(Q_{2^t-1})$ as in Theorem 2.1. We will treat the vertices $v \in V(Q_d)$ differently depending on which of the triples $(i, r, \gamma) \in [m] \times \{0, 1\} \times \{0, 1, 2\}$ satisfy $v(i, r, \gamma) \in C$. For starters, let X denote the set of all vertices v for which there exists some i and $(r, \gamma) \neq (r', \gamma')$ such that $v(i, r, \gamma), v(i, r', \gamma') \in C$. The vertices of X will be isolated in

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

G and will not play a large role in the construction. Now, divide the vertices of $V(Q_d) \setminus X$ into sets $A_0, \dots, A_m$ where A_j is defined to be the set of all vertices of $V(Q_d) \setminus X$ for which there are exactly j triples (i, r, γ) such that $v(i, r, \gamma) \in C$. We remark that, by (2.1) and (2.2),

$$|V(Q_d) \setminus A_0| = O(2^d/d). \tag{2.3}$$

Our goal is to construct a graph G which does not contain Q_m as a subgraph and does not contain any edge uv of Q_d where $u, v \in A_0$ such that, for any such edge, the graph $G + uv$ contains a copy of Q_m. Given such a graph G, let G' be a graph obtained by adding a maximal set of edges to G without creating a copy of Q_m. Then G' is (Q_d, Q_m)-saturated and every edge of G' has at least one endpoint in $V(Q_d) \setminus A_0$. By (2.3), the total number of edges incident to $V(Q_d) \setminus A_0$ is at most $O(2^d)$, and so we obtain $\mathrm{sat}(Q_d, Q_m) \leqslant |E(G')| = O(2^d)$, as desired.

The difficulty of the proof is to ensure that $G + e$ contains a copy of Q_m for every edge e of Q_d joining two vertices of A_0 while simultaneously maintaining the property that G does not contain a copy of Q_m. For the latter, we will apply Theorem 2.2 and a parity condition to 'break' any potential copies of Q_m in G. In what follows, let $\phi_1 : E(Q_{6(2^t-1)}) \to \{0, 1, 2\}$ and $\phi_2 : E(Q_s) \to \{0, 1, 2\}$ be colourings as in Theorem 2.2, and let $\phi : E(Q_{6(2^t-1)}) \sqcup E(Q_s) \to \{0, 1, 2\}$ be the mapping which is equal to ϕ_1 on $E(Q_{6(2^t-1)})$ and equal to ϕ_2 on $E(Q_s)$.

2.3. The construction

We now give the details of our construction.

Step 1. Add to G every edge of Q_d which joins a vertex of A_j to a vertex of A_{j+1} for $0 \leqslant j \leqslant m - 2$.

Step 2. Suppose that uv is an edge of Q_d such that $u, v \in A_j$ for some $1 \leqslant j \leqslant m - 1$. Let k be the unique element of $[m + 1]$ such that $u(k) \neq v(k)$. We add the edge uv to G if the following three conditions hold:

(i) either $k = m + 1$ or there does not exist $(r', \gamma') \in \{0, 1\} \times \{0, 1, 2\}$ such that

$$u(k, r', \gamma') = v(k, r', \gamma') \in C,$$

and for every $(i, r, \gamma) \in [m] \times \{0, 1\} \times \{0, 1, 2\}$ such that $i \neq k$ and $v(i, r, \gamma) \in C$ we have

(ii) $\gamma = \phi(u(k)v(k))$, and

(iii) $|v| + |v(k)| + |v(i)| + j - 1 \equiv r \mod 2$.

Note that (iii) is well defined since u and v differ only on a coordinate of the interval corresponding to k, and so for every $i \neq k$ we have

$$|v| + |v(k)| + |v(i)| \equiv |u| + |u(k)| + |u(i)| \mod 2.$$

Before moving on we make a few observations, each of which can be verified by looking carefully at Steps 1 and 2.

Observation 2.3. $A_m \cup X$ is a set of isolated vertices in G.

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

Observation 2.4. A_0 is an independent set of G.

Observation 2.5. If uv is an edge of G such that $u(i,r,\gamma) = v(i,r,\gamma) \in C$, then $u(i,r',\gamma') = v(i,r',\gamma')$ for every pair $(r',\gamma') \in \{0,1\} \times \{0,1,2\}$.

Observation 2.6. If uv is an edge of G such that $u(m+1) \neq v(m+1)$, then $u, v \in A_j$ for some $1 \leqslant j \leqslant m-1$.

The next observation follows from the fact that C is an independent set in Q_{2^t-1}.

Observation 2.7. If uv is an edge of G such that $u(i,r,\gamma) \neq v(i,r,\gamma)$, then at most one of $u(i,r,\gamma)$ or $v(i,r,\gamma)$ is in C.

To complete the proof of Theorem 1.2, it suffices to establish the following two claims. We state these claims now and show that they imply Theorem 1.2 before proving the claims themselves.

Claim 2.8. *For every edge uv of Q_d such that $u, v \in A_0$, the graph $G + uv$ contains a copy of Q_m.*

Claim 2.9. *G does not contain Q_m as a subgraph.*

Proof of Theorem 1.2. If G is not (Q_d, Q_m)-saturated, then by Claim 2.9 we can extend G to a (Q_d, Q_m)-saturated graph G' by adding a maximal set of edges which do not create a copy of Q_m. By Claim 2.8 none of these additional edges are between vertices in A_0 and so, by Observation 2.4, A_0 is an independent set in G'. Thus, every edge of G' has at least one endpoint in $\left(\bigcup_{k=1}^{m} A_k\right) \cup X$. The total number of edges incident to vertices of this set is at most

$$d\left|\left(\bigcup_{k=1}^{m} A_k\right) \cup X\right| = d|A_1| + d\left|\left(\bigcup_{k=2}^{m} A_k\right) \cup X\right|$$

Note that

$$\left|\left(\bigcup_{k=2}^{m} A_k\right) \cup X\right| = O(|C|^2 2^{d-2(2^t-1)}),$$

which is $O(2^d/d^2)$ by (2.1) and (2.2). Therefore the second term of the above expression is $o(2^d)$ and it suffices to bound $d|A_1|$. We have

$$\begin{aligned} d|A_1| &= d(6m|C|2^{d-(2^t-1)}) = d(6m2^{2^t-t-1}2^{d-(2^t-1)}) = 6md2^{d-t} \\ &= 6m(6m(2^t-1)+s)2^{d-t} < 72m^2 2^d \end{aligned}$$

by (2.2). The result follows. □

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

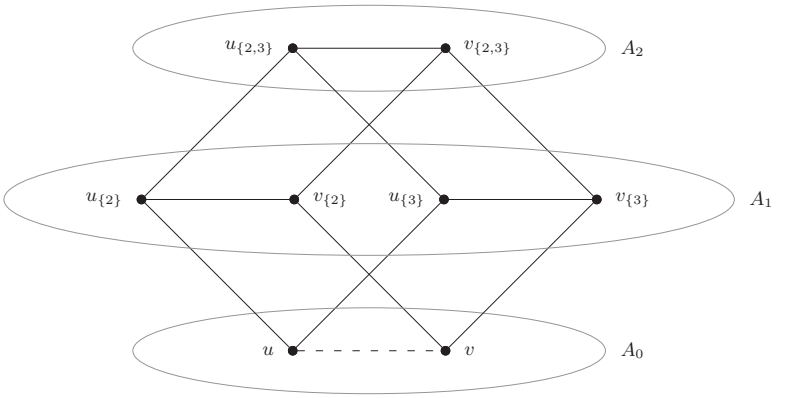


Figure 1. An illustration of the proof of Claim 2.8 in the case that $m = 3$ and $u(1) \neq v(1)$.

Remark 2.10. Note that if d is of the form $6m(2^t - 1)$ for some t, then we obtain a better bound:

$$\mathrm{sat}(Q_d, Q_m) \leqslant (1 + o(1))36m^2 2^d.$$

Thus, it suffices to prove Claims 2.8 and 2.9.

Proof of Claim 2.8. Let uv be an edge of Q_d where $u, v \in A_0$, and let k be the unique element of $[m+1]$ for which $u(k) \neq v(k)$. Define $\gamma := \phi(u(k)v(k))$ and let $I \subseteq [m] \setminus \{k\}$ be a set of size $m-1$. For $i \in I$, pick $r_i \in \{0,1\}$ so that

$$|v| + |v(k)| + |v(i)| \equiv r_i \mod 2.$$

For each $i \in I$, let c_i be the unique neighbour of $v(i, r_i, \gamma)$ in Q_{2^t-1} contained in C. Given $I' \subseteq I$, we let $v_{I'}$ and $u_{I'}$ be the vertices of Q_d such that $u_{I'}(i, r_i, \gamma) = v_{I'}(i, r_i, \gamma) = c_i$ for all $i \in I'$ and, on all other coordinates, $u_{I'}$ and $v_{I'}$ agree with u and v, respectively. In particular, $v_\emptyset = v$ and $u_\emptyset = u$. If $I' \neq I$, then for any $i' \in I \setminus I'$ we see that $v_{I'}$ is adjacent to $v_{I' \cup \{i'\}}$ and $u_{I'}$ is adjacent to $u_{I' \cup \{i'\}}$ in G via edges added in Step 1. Also, for $I' \subseteq I$, we have $v_{I'}, u_{I'} \in A_{|I'|}$ and, if $I' \neq \emptyset$, then for each $i \in I'$,

$$|v_{I'}| + |v_{I'}(k)| + |v_{I'}(i)| \equiv r_i + |I'| - 1 \mod 2.$$

Thus, for $I' \neq \emptyset$, $v_{I'}$ is adjacent to $u_{I'}$ via an edge of G added in Step 2. This implies that $\{v_{I'} : I' \subseteq I\} \cup \{u_{I'} : I' \subseteq I\}$ induces a copy of Q_m in $G + uv$, which completes the proof. □

Proof of Claim 2.9. Suppose, to the contrary, that G contains a copy Q of Q_m. Let $J \subseteq [d]$ be the set of coordinates which are variable under Q. For each $i \in [m+1]$, let $J(i)$ be the set of coordinates of J which are in the interval of $[d]$ corresponding to i. Moreover, given

$$(i, r, \gamma) \in [m] \times \{0,1\} \times \{0,1,2\},$$

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

let $J(i,r,\gamma)$ be the set of coordinates of J which are in the interval of $[d]$ corresponding to (i,r,γ).

Subclaim 2.11. *For each $i \in [m]$ there is a pair*

$$(r_i,\gamma_i) \in \{0,1\} \times \{0,1,2\}$$

such that $J(i) = J(i,r_i,\gamma_i)$.

Proof. Suppose to the contrary that there exists $(r,\gamma) \neq (r',\gamma')$ such that both $J(i,r,\gamma)$ and $J(i,r',\gamma')$ are non-empty.

First, suppose that there exists a vertex $v \in V(Q)$ with $v(i,r,\gamma) \in C$. Let u be a neighbour of v in Q obtained by changing a coordinate in $J(i,r',\gamma')$. By Observation 2.5, the edge uv is not contained in G, which is a contradiction. Thus, $v(i,r,\gamma), v(i,r',\gamma') \notin C$ for all $v \in V(Q)$.

Now, let Q' be a copy of Q_2 in Q obtained by starting with an arbitrary vertex of Q and varying one coordinate of $J(i,r,\gamma)$ and one coordinate of $J(i,r',\gamma')$. By the result of the previous paragraph, we see that $V(Q') \subseteq A_j$ for some j and so every edge of Q' was added in Step 2. By taking the edges of Q' and restricting them to the interval of $[d]$ corresponding to i, we see that all of the resulting edges must receive the same colour under ϕ. This contradicts the fact that there is no copy of Q_2 which is monochromatic under ϕ and completes the proof of the subclaim. □

Subclaim 2.12. *Suppose that $|J(i)| \geqslant 2$ for some $i \in [m]$, and let Q' be a copy of Q_2 in Q obtained by starting with an arbitrary vertex and varying two coordinates of $J(i)$. Then there is a unique vertex $v \in V(Q')$ with $v(i,r_i,\gamma_i) \in C$.*

Proof. By Subclaim 2.11 we have $J(i) = J(i,r_i,\gamma_i)$. By Observation 2.7, if uv is an edge of Q', then we cannot have $u(i,r_i,\gamma_i), v(i,r_i,\gamma_i) \in C$. Thus, if there are two vertices u,v of Q' for which $u(i,r_i,\gamma_i), v(i,r_i,\gamma_i) \in C$, then they must be non-adjacent. So, in this case, the vertices of Q' alternate between A_j and A_{j+1} for some j and every edge of Q' was added in Step 1. However, this implies that for $w \in V(Q')$ such that $w(i,r_i,\gamma_i) \notin C$, the vertex $w(i,r_i,\gamma_i)$ of Q_{2^t-1} must have two distinct neighbours in C, which is a contradiction.

Now we assume that every vertex v of Q' satisfies $v(i,r_i,\gamma_i) \notin C$. However, in this case, we see that every edge of Q' was added in Step 2. As before, we obtain a copy of Q_2 in $Q_{6(2^t-1)}$ which is monochromatic under ϕ, a contradiction. □

Subclaim 2.13. $|J(i)| \leqslant 2$ *for every* $i \in [m]$.

Proof. Suppose not. Let Q' be a copy of Q_3 in Q obtained by starting with an arbitrary vertex and varying three coordinates of $J(i)$. If we vary any pair of these coordinates, leaving the third one fixed, we obtain a copy of Q_2 in Q' which must obey Subclaim 2.12. This implies that there are precisely two vertices $x,y \in V(Q')$ such that $x(i,r_i,\gamma_i), y(i,r_i,\gamma_i) \in C$ and they are at distance 3 in Q'. However, now we get that the vertices of $V(Q') \setminus \{x,y\}$ induce a copy F of C_6 in Q, where every edge of F was added in Step 2. Taking the edges

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

of F and restricting them to the interval of $[d]$ corresponding to i, we see that all such edges must receive the same colour under ϕ. This contradicts the fact that there is no copy of C_6 which is monochromatic under ϕ, and completes the proof of the subclaim. □

Subclaim 2.14. $|J(m+1)| \leqslant 1$.

Proof. If not, then let Q' be a copy of Q_2 in Q obtained by starting at an arbitrary vertex and varying two coordinates of $J(m+1)$. Then, restricting the edges of Q' to the last s coordinates, we obtain a copy of Q_2 in Q_s which is monochromatic under ϕ, a contradiction. □

Subclaim 2.15. *Suppose that uv and vw are distinct edges of Q where $u,v,w \in A_j$ for some j. Then there is some $k \in [m]$ such that $u(k) \neq v(k)$ and $v(k) \neq w(k)$.*

Proof. We show that there cannot exist distinct integers $k,k' \in [m+1]$ such that $u(k) \neq v(k)$ and $v(k') \neq w(k')$. This is sufficient to prove the subclaim since, by Subclaim 2.14, we cannot have $u(m+1) \neq v(m+1)$ and $v(m+1) \neq w(m+1)$.

Suppose that such integers k,k' exist, and let x be the unique vertex of Q_d distinct from v which is joined to both u and w in Q_d. Note that both of the edges xu and wx are present in Q. Also, $x \in A_j$ and so all of the edges uv, vw, wx, xu were added to G in Step 2. This implies that $j \geqslant 1$ and that there is a triple (i,r,γ) with $i \notin \{k,k'\}$ such that

$$u(i,r,\gamma) = v(i,r,\gamma) = w(i,r,\gamma) = x(i,r,\gamma) \in C.$$

However, one can easily check that the parity of $|v| + |v(k)| + |v(i)|$ is different from the parity of $|x| + |x(k)| + |x(i)|$ modulo 2, and so only one can be equivalent to r. By definition of Step 2, only one of the edges uv and xw can exist in G. This contradiction completes the proof. □

Subclaim 2.16. *If $|J(m+1)| = 1$, then $|J(i)| \leqslant 1$ for all $i \in [m]$.*

Proof. If not, let Q' be a copy of Q_3 in Q obtained by starting at an arbitrary vertex and varying one coordinate of $J(m+1)$ and two of $J(i)$. Then, by Subclaim 2.12, there is an edge $uv \in E(Q')$ such that $u(i) \neq v(i)$ and $u,v \in A_j$ for some j (in fact, there are many such edges). Now, let w be the neighbour of v in Q' obtained by changing the coordinate of $J(m+1)$. By Observation 2.6, we have that $w \in A_j$ as well. However, this contradicts Subclaim 2.15. □

Subclaim 2.17. *There is at most one $i \in [m]$ for which $|J(i)| = 2$.*

Proof. If not, let $i,i' \in [m]$ such that $|J(i)| = |J(i')| = 2$ and let Q' be a copy of Q_4 obtained by starting at an arbitrary vertex and varying two coordinates of $J(i)$ and two coordinates of $J(i')$. By Subclaim 2.12, we see that there must exist edges $uv, vw \in E(Q')$ such that $u(i) \neq v(i)$, $v(i') \neq w(i')$ and $u,v,w \in A_j$ for some j. This contradicts Subclaim 2.15 and completes the proof. □

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

Now, let us complete the proof of the claim. Throughout, for each $i \in [m]$, we let (r_i, γ_i) be a pair such that $J(i) = J(i, r_i, \gamma_i)$, which exists by Subclaim 2.11. In what follows, we let j^* be the minimum integer j such that $V(Q) \cap A_j \neq \emptyset$ and let $v^* \in V(Q) \cap A_{j^*}$. Note that for every $i \in [m]$ such that $J(i) \neq \emptyset$ we must have $v^*(i, r_i, \gamma_i) \notin C$. If not, then starting with v^* and changing a coordinate of $J(i)$ yields a vertex of $V(Q) \cap A_{j^*-1}$ by Observations 2.5 and 2.7, contradicting our choice of j^*. By Subclaims 2.16 and 2.17, there is at most one $i \in [m]$ such that $J(i) = \emptyset$. This implies that

$$j^* \in \{0, 1\}, \tag{2.4}$$

where j^* must equal zero if no such i exists. We divide the proof into cases.

Case 1. $|J(i)| = 1$ for every $i \in [m]$.

In this case, we must have $j^* = 0$ and so $v^* \in A_0$. However, if we start with v and change the coordinate of $J(i)$ for any i, then we obtain a vertex u with $u(i, r_i, \gamma_i) \in C$ by Observation 2.4. Thus, changing all m of these coordinates yields a vertex w such that $w(i, r_i, \gamma_i) \in C$ for all $i \in [m]$; that is, $w \in A_m$. This is a contradiction since, by Observation 2.3, w is an isolated vertex and therefore cannot belong to Q.

Case 2. $|J(m+1)| = 1$.

Let u be the neighbour of v^* in Q obtained by changing the coordinate in $J(m+1)$. We must have $u, v^* \in A_1$ by (2.4) and Observation 2.6. This implies that there is some $i \in [m]$ such that $J(i) = \emptyset$ and $v^*(i, r_i, \gamma_i) \in C$. Also, by Subclaim 2.16 and the Pigeonhole Principle, we must have $|J(i')| = 1$ for every $i' \in [m] \setminus \{i\}$.

Now, for each $i' \in [m] \setminus \{i\}$, let $w_{i'}$ be the neighbour of v^* in Q obtained by changing the coordinate in $J(i')$. By Subclaim 2.15 and the fact that $u, v^* \in A_1$, we cannot have $w_{i'} \in A_1$. So, by our choice of j^*, we must have $w_{i'}(i', r_{i'}, \gamma_{i'}) \in C$. As in the proof of Case 1, if we start with v^* and change the coordinate of $J(i')$ for every $i' \in [m] \setminus \{i\}$, we obtain a vertex of A_m contained in Q, contradicting Observation 2.3. This completes the proof in this case.

Case 3. $|J(i)| = 2$ for some $i \in [m]$.

By Subclaim 2.12, there is a neighbour u of v^* in Q obtained by changing a coordinate in $J(i)$ such that $u \in A_{j^*}$. This immediately implies that $j^* = 1$ by (2.4) and Observation 2.4.

Now, for each $i' \in [m]$ for which $|J(i')| = 1$, let $w_{i'}$ be the neighbour of v^* in Q obtained by changing the coordinate in $J(i')$. By Subclaim 2.15 and the fact that $u, v^* \in A_1$, we must have $w_{i'}(i', r_{i'}, \gamma_{i'}) \in C$. Thus, if we let x be the vertex obtained from v^* by changing the coordinate of $J(i')$ for every such i', then $x \in A_{m-1}$.

Let Q' be the copy of Q_2 in Q obtained by starting at x and varying the two coordinates of $J(i)$. Then, by Subclaim 2.12, there must be some vertex y of $V(Q')$ such that $y(i, r_i, \gamma_i) \in C$. However, this implies that $y \in A_m$, contradicting Observation 2.3. This completes the proof of Claim 2.9 and of Theorem 1.2. □

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

3. Weak saturation

3.1. Hypercubes and grids

In this section, we discuss weak saturation in cubes and more generally in grids. We will prove Theorem 1.4, which immediately implies Theorem 1.3.

Weak saturation is part of a more general theory, known as *bootstrap percolation*. In the (F,H)-*graph bootstrap process*, we start with an initial set S_0 of 'infected' vertices in a graph F and, at the ith step of the process, a vertex

$$v \in V(F) \setminus \left(\bigcup_{j=0}^{i-1} S_j\right)$$

becomes infected and is added to S_i if there is a copy of H in the subgraph of F induced by

$$\left(\bigcup_{j=0}^{i-1} S_j\right) \cup \{v\}$$

containing v. A natural extremal problem is to determine the size of the smallest initial set S_0 such that every vertex of F is eventually infected; such a set is said to be (F,H)-*percolating*. It is easily seen that weak saturation corresponds to a bootstrap process on the edges of a graph, rather than its vertices (*i.e.*, a bootstrap process on the line graph). For an introduction to the literature on bootstrap percolation, see for instance [3, 4, 5] and the references therein.

Here we are interested in the 'edge version' of the problem for cubes and grids. The 'vertex version' of the problem was solved by Balogh, Bollobás, Morris and Riordan [6], who determined the minimum size of a subset of $V(P_k^d)$ which is percolating with respect to (the vertex sets of) axis aligned copies of P_r^m, and also the minimum size of a (K_k^d, K_r^m)-percolating set, for all $k \geqslant r \geqslant 2$ and $d \geqslant m \geqslant 1$ (here K_k^d is the graph with vertex set $\{0,\dots,k-1\}^d$ and two vertices are adjacent if they differ in exactly one coordinate). Somewhat surprisingly, the two quantities are the same.

We will use the following simple linear algebraic lemma from [6]. Given a graph F and a set $\mathcal{H}$ of subgraphs of F, let $\mathrm{wsat}(F,\mathcal{H})$ be the minimum number of edges in a graph G such that the edges of $E(F) \setminus E(G)$ can be added to G, one edge at a time, in such a way that each added edge increases the number of graphs of $\mathcal{H}$ contained in G; such a graph is said to be *weakly* $(F,\mathcal{H})$-*saturated*. In this language, $\mathrm{wsat}^*(P_k^d, P_r^m) = \mathrm{wsat}(P_k^d, \mathcal{H})$, where $\mathcal{H}$ contains all axis aligned copies of P_r^m in P_k^d.

Lemma 3.1 (Balogh, Bollobás, Morris and Riordan [6]). *Let F be a graph, let $\mathcal{H}$ be a collection of subgraphs of F, and let W be a vector space. Suppose that there exists a set $\{f_e : e \in E(F)\} \subseteq W$ such that for every $H \in \mathcal{H}$ there are non-zero scalars $\{c_e : e \in E(H)\}$ such that $\sum_{e \in E(H)} c_e f_e = 0$. Then*

$$\mathrm{wsat}(F,\mathcal{H}) \geqslant \dim(\mathrm{span}\{f_e : e \in E(F)\}).$$

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

Proof. Let G be a weakly $(F,\mathcal{H})$-saturated graph. Define $G_0 := G$ and label the edges of $E(F) \setminus E(G)$ by $e_1,\dots,e_k$ such that for $1 \leqslant i \leqslant k$ there is a subgraph $H_i \in \mathcal{H}$ of $G_i := G_{i-1} + e_i$ such that H_i contains e_i. Thus, by hypothesis, f_{e_i} can be written as a linear combination of the vectors in $\{f_e : e \in E(H_i) \setminus \{e_i\}\}$. This implies that

$$\operatorname{span}\{f_e : e \in E(G_0)\} = \operatorname{span}\{f_e : e \in E(G_1)\} = \cdots = \operatorname{span}\{f_e : e \in E(G_k)\}.$$

Since $G = G_0$ and $F = G_k$, it must be the case that

$$|E(G)| \geqslant \dim(\operatorname{span}\{f_e : e \in E(F)\}).$$

Since G was an arbitrary weakly $(F,\mathcal{H})$-saturated graph, this completes the proof. □

To handle the vertex version of the question, Balogh, Bollobás, Morris and Riordan gave a clever construction of a suitable vector space and then proved that it has the same dimension as a suitable percolating set. The result then follows by Lemma 3.1. Our proof of Theorem 1.4 uses the same approach, but requires a different construction. For other examples in which the notions of linear dependence and independence are applied to solve problems in weak saturation, see Alon [1], Kalai [29] and Pikhurko [34, 35].

Given a vertex v of P_k^d, say that a coordinate of v is *large* if it has size at least $r-1$ and *small* otherwise. We let $L(v)$ denote the number of large coordinates of v, and let $|v|$ denote the sum of the coordinates of v. Given $i \in [d]$, a *line* $\mathcal{L}$ in direction i is a path of length $k-1$ in P_k^d in which any two vertices of $\mathcal{L}$ differ on the ith coordinate.

Proof of Theorem 1.4. Let G be a spanning subgraph of P_k^d in which, for each vertex v,

- we add every edge uv of P_k^d such that $|u| = |v| - 1$ and u and v differ on a small coordinate of v, and
- we add $\min\{L(v), m-1\}$ edges uv of P_k^d such that $|u| = |v| - 1$ and u and v differ on a large coordinate of v.

Some elementary counting gives us

$$E(G) = \sum_{j=0}^{d}\sum_{i=0}^{d-j}(m-1+i)\binom{d}{j}\binom{d-j}{i}(k-r+1)^j(r-2)^i - \sum_{j=0}^{m-2}\sum_{i=0}^{d-j}(m-1-j)\binom{d}{j}\binom{d-j}{i}(k-r+1)^j(r-2)^i.$$

So, to prove the upper bound, we need only show that G is weakly (P_k^d, P_r^m)-saturated. In order of increasing $|v|$, we add all missing edges of $E(P_k^d) \setminus E(G)$ from v to vertices u with $|u| = |v| - 1$ one edge at a time. By construction, for every added edge uv with $|u| = |v| - 1$, the coordinate on which u and v differ is large in v. For any vertex x, if $L(x) \leqslant m-1$, then all edges xy with $|x| = |y| - 1$ are already present in G. Also, if v is a vertex with $L(v) \geqslant m$, then G contains $m-1$ edges wv with $|w| = |v| - 1$ such that the coordinate on which w and v differ is large in v. Putting this together, an easy inductive argument shows that, if we add edges in this order, then each edge added from v to a vertex u with $|u| = |v| - 1$ creates a new copy of P_r^m in which v is the 'top' vertex.

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

For the lower bound, we apply Lemma 3.1 where $\mathcal{H}$ consists of all axis aligned copies of P_r^m in P_k^d. So, it suffices to show that there exists a vector space W and a set

$$\{f_e : e \in E(P_k^d)\} \subseteq W$$

which satisfy the hypotheses of Lemma 3.1 such that

$$\dim(\text{span}\{f_e : e \in E(P_k^d)\}) \geqslant |E(G)|. \tag{3.1}$$

The space W that we choose is the direct sum of k^d copies of $\mathbb{R}^{m-1}$, one for each vertex of P_k^d, and dk^{d-1} copies of $\mathbb{R}^{r-2}$, one for each line in P_k^d. That is,

$$W := \left(\bigoplus_x \mathbb{R}^{m-1}\right) \oplus \left(\bigoplus_{\mathcal{L}} \mathbb{R}^{r-2}\right),$$

where x ranges over all vertices of P_k^d and $\mathcal{L}$ ranges over all lines in P_k^d. Given a vector w of W, a vertex x of P_k^d and a line $\mathcal{L}$ of P_k^d, let $\pi_x(w)$ denote the projection of w onto the copy of $\mathbb{R}^{m-1}$ corresponding to x and let $\pi_{\mathcal{L}}(w)$ denote the projection of w onto the copy of $\mathbb{R}^{r-2}$ corresponding to $\mathcal{L}$. Note that w is determined by its projections.

Let $Z := \{z_1, \ldots, z_d\}$ be a collection of d vectors of $\mathbb{R}^{m-1}$ in general position. Also, let $Y := \{y_1, \ldots, y_{k-2}\}$ be a set of $k-2$ vectors of $\mathbb{R}^{r-2}$ such that $y_1, \ldots, y_{r-2}$ are linearly independent and, for $r-1 \leqslant t \leqslant k-1$,

$$y_t := -\sum_{j=t-r+2}^{t-1} y_j. \tag{3.2}$$

Thus any consecutive $r-2$ vectors $y_i, \ldots, y_{i+r-2}$ are linearly independent and any consecutive $r-1$ vectors $y_j, \ldots, y_{j+r-1}$ sum to zero. Suppose that $e = uv$ is an edge of P_k^d such that u and v differ on coordinate $i \in [d]$ and let t be the maximum of the ith coordinates of u and v. Further, let $\mathcal{L}$ be the unique line of P_k^d containing e. We define f_e to be the vector of W such that

- $\pi_u(f_e) = \pi_v(f_e) = z_i$ and $\pi_x(f_e) = 0$ for every $x \in V(Q_d) \setminus \{u, v\}$, and
- $\pi_{\mathcal{L}} = y_t$ and $\pi_{\mathcal{L}'} = 0$ for every line $\mathcal{L}' \neq \mathcal{L}$.

In order to apply Lemma 3.1, we need to show that for every axis aligned copy P of P_r^m in P_k^d there are non-zero scalars $\{c_e : e \in E(P)\}$ such that

$$\sum_{e \in E(P)} c_e f_e = 0.$$

Let P be an axis aligned copy of P_r^m in P_k^d and let $I \subseteq [d]$ be the set of m coordinates which vary under P. Let $\{c_i : i \in I\}$ be a set of scalars such that

$$\sum_{i \in I} c_i z_i = 0.$$

Note that $c_i \neq 0$ for all $i \in I$ since the vectors of Z are in general position.

For each vertex v of P, let $M(v) \subseteq [d]$ be the set of indices j such that both $v - e_j$ and $v + e_j$ are contained in P. Define the *lines* of P to be the paths of length $r-1$ in P obtained by taking the intersection of a line of P_k^d with $V(P)$. Note that $|M(v)|$ is

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

precisely the number of lines of P in which v has degree two. For each line $\mathcal{L}$ of P, define $m(\mathcal{L}) := |M(v)|$ where v is an endpoint of $\mathcal{L}$ (clearly the endpoints give the same value). For each $i \in I$, let E_i be the set of all edges of P for which i is the variable coordinate. For each edge $e \in E_i$ contained in a line $\mathcal{L}$ of P, define $d_e := 2^{m(\mathcal{L})}c_i$. We claim that

$$\sum_{e \in E(P)} d_e f_e = 0.$$

First, let $\mathcal{L}$ be a line of P in direction i. Then, for some $t \geqslant r-1$, we have

$$\pi_{\mathcal{L}}\left(\sum_{e \in E(P)} d_e f_e\right) = \sum_{e \in E(P)} d_e \pi_{\mathcal{L}}(f_e) = \sum_{e \in E(\mathcal{L})} 2^{m(\mathcal{L})} c_i \pi_{\mathcal{L}}(f_e) = 2^{m(\mathcal{L})} c_i \sum_{j=t-r+2}^{t} y_j.$$

However, this sum is equal to zero by (3.2).

Now, fix a vertex v of P, and for each $i \in I$ let $\mathcal{L}_i$ be the line of P in direction i containing v. We have

$$\begin{aligned}\pi_v\left(\sum_{e \in E(P)} d_e f_e\right) &= \sum_{e \in E(P)} d_e \pi_v(f_e) \\ &= \sum_{i \in M(v)} \sum_{e \in E_i} 2^{m(\mathcal{L}_i)} c_i \pi_v(f_e) + \sum_{i \in I \setminus M(v)} \sum_{e \in E_i} 2^{m(\mathcal{L}_i)} c_i \pi_v(f_e),\end{aligned}$$

which is equal to

$$\sum_{i \in M(v)} 2^{m(\mathcal{L}_i)+1} c_i \pi_v(f_e) + \sum_{i \in I \setminus M(v)} 2^{m(\mathcal{L}_i)} c_i \pi_v(f_e)$$

by definition of $M(v)$. We observe that, for any $i \in M(v)$, we have $m(\mathcal{L}_i) = |M(v)| - 1$ and for any $i \in I \setminus M(v)$ we have $m(\mathcal{L}_i) = |M(v)|$. Therefore, the above sum is equal to

$$2^{|M(v)|} \sum_{i \in I} c_i z_i,$$

which is zero by our choice of $\{c_i : i \in I\}$. Combining this with the result of the previous paragraph, we get

$$\sum_{e \in E(P)} d_e f_e = 0,$$

as desired.

To complete the proof, it suffices to prove (3.1). To do so, we let G be a graph as in the proof of the upper bound and show that the vectors of $\{f_e : e \in E(G)\}$ are linearly independent. Let $\{c_e : e \in E(G)\}$ be any set of scalars, not all of which are zero, and let F be the spanning subgraph of G containing all edges $e \in E(G)$ such that $c_e \neq 0$. Let v be a vertex of non-zero degree in F such that $|v|$ is maximum.

First, consider the case that v has degree at most $m-1$ in F. Let J denote the set of $d_F(v)$ coordinates such that, for each $j \in J$, the edge from v to $v - e_j$ is present in F. Then

$$\sum_{e \in E(G)} c_e \pi_v(f_e)$$

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

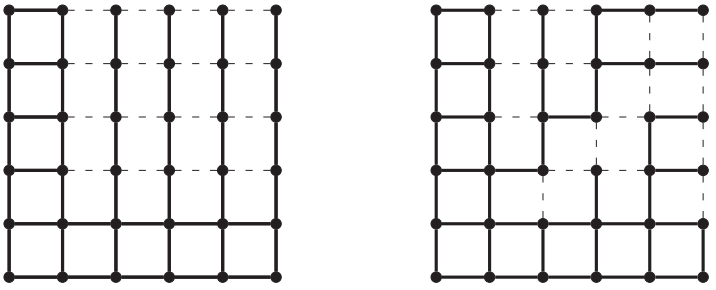


Figure 2. Two non-isomorphic weakly (P_6^2, P_3^2)-saturated graphs with $\mathrm{wsat}(P_6^2, P_3^2)$ edges.

is a linear combination of the vectors in $\{z_j : j \in J\}$ in which not all of the coefficients are zero. Since $|J| \leqslant m-1$ and the vectors of Z are in general position, this sum is non-zero and therefore

$$\sum_{e \in E(G)} c_e f_e \neq 0.$$

Now, suppose that v has degree at least m in F. By construction of G, this implies that there is a coordinate j which is small for v such that the edge e from v to $v - e_j$ is present in F. Let $\mathcal{L}$ be the unique line of P_k^d containing e. Then, by maximality of $|v|$, every edge of $\mathcal{L}$ contained in F joins two vertices for which j is a small coordinate. This implies that

$$\sum_{e \in E(F)} c_e \pi_{\mathcal{L}}(f_e)$$

is a linear combination of the vectors in $\{y_i : 1 \leqslant i \leqslant r-1\}$ in which not all of the coefficients are zero. Thus, since the vectors $y_1, \ldots, y_{r-1}$ are linearly independent, this sum is non-zero and we obtain

$$\sum_{e \in E(G)} c_e f_e \neq 0,$$

which completes the proof. □

Remark 3.2. The proof of Theorem 1.4 implies that, for most values of $k \geqslant r \geqslant 2$ and $d \geqslant m \geqslant 1$, there are many tight examples. For an example in two dimensions, see Figure 2.

3.2. Cycles in the grid

We prove Theorem 1.5 using an elementary argument.

Proof of Theorem 1.5. First, notice that

$$\mathrm{wsat}(P_k^d, C_{2\ell}) \geqslant k^d - 1$$

holds since any weakly $(P_k^d, C_{2\ell})$-saturated graph must be connected. To prove the upper bound, we proceed by induction on $d + k$.

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

Consider the base case $k = 2$ and $d = \ell$. For $0 \leqslant t \leqslant \ell$, let L_t be the set of all vertices v of Q_ℓ with $|v| = t$ and define x_t to be the vertex of L_t which consists of t ones followed by $\ell - t$ zeros. We let G be the graph such that, for $1 \leqslant t \leqslant \ell$, each vertex $v \in L_t$ is joined to a unique vertex of L_{t-1} which is chosen in the following way:

- if $t = 1$ or $v = x_t$, then v is joined to x_{t-1},
- if $t \geqslant 2$ and $v \neq x_t$, then v is joined to an arbitrary vertex $y \in L_{t-1} \setminus \{x_{t-1}\}$ such that $vy \in E(Q_\ell)$.

It is clear that G is a tree and so $|E(G)| = 2^\ell - 1$. For each vertex v, let P_v be the unique path in G from v to x_0. Note that, by construction, if $v \notin \{x_1, \ldots, x_\ell\}$ then $V(P_v) \cap \{x_1, \ldots, x_\ell\} = \emptyset$.

We show that G is weakly $(Q_\ell, C_{2\ell})$-saturated. First, add each edge $x_\ell v$ where $v \in L_{\ell-1} \setminus \{x_{\ell-1}\}$. Each of these edges creates a copy of $C_{2\ell}$ by taking the paths P_{x_ℓ} and P_v along with the edge $x_\ell v$.

Next, for each $t = \ell - 1, \ldots, 2$, in turn, we add every edge uv where $u \in L_t \setminus \{x_t\}$ and $v \in L_{t-1} \setminus \{x_{t-1}\}$. When doing so we can assume, inductively, that for each s such that $t < s \leqslant \ell$ every edge from $L_s \setminus \{x_s\}$ to $L_{s-1} \setminus \{x_{s-1}\}$ is present. In particular, this implies that there is a path R_u of length $\ell - t$ from u to x_ℓ which does not contain any vertex of $\{x_0, \ldots, x_{\ell-1}\}$. Therefore, when adding the edge uv, we obtain a copy of $C_{2\ell}$ by taking the paths P_v, P_{x_ℓ}, R_u and the edge uv.

After this, for each $t = \ell - 1, \ldots, 2$, in turn, we add every edge ux_{t-1} where $u \in L_t \setminus \{x_t\}$. In this case, let w be a vertex of $L_t \setminus \{x_t, u\}$ such that there are paths R_u and R_w from u to x_ℓ and w to x_ℓ, respectively, such that $V(R_u) \cap V(R_w) = \{x_\ell\}$. We obtain a copy of $C_{2\ell}$ by taking the paths R_u, R_w, P_w and $P_{x_{t-1}}$ along with the edge ux_{t-1}.

Finally, for each $t = \ell - 1, \ldots, 2$, in turn, we add every edge $x_t v$ where $v \in L_{t-1} \setminus \{x_{t-1}\}$. Let j be an index on which v is zero and let P be a path from x_ℓ to x_0 which contains e_j but is disjoint from $\{x_1, \ldots, x_{\ell-1}\}$. Notice that P does not contain any vertex of P_v. We obtain a copy of $C_{2\ell}$ by taking the paths $x_t x_{t+1} \cdots x_\ell$, P, and P_v along with the edge $x_t v$. This completes the proof of the base case.

Now, we assume that $d + k > \ell + 2$ and that the proposition holds for smaller values of $d + k$. We divide the proof into two cases.

Case 1. $d > \ell$.

Let G_{d-1} be a weakly $(P_k^{d-1}, C_{2\ell})$-saturated graph with $k^{d-1} - 1$ edges. We construct a weakly $(P_k^d, C_{2\ell})$-saturated graph G. For each vertex v of P_k^d, let v_{d-1} be the restriction of v to its first $d - 1$ coordinates and let v_d be the last coordinate of v. In constructing G, we add every edge uv of P_k^d such that

- $u_{d-1}v_{d-1} \in E(G_{d-1})$ and $u_d = v_d = 0$, or
- $u_{d-1} = v_{d-1}$.

It is clear that $|E(G)| = k^d - 1$. For $0 \leqslant t \leqslant k - 1$, let S_t be the set of all vertices v such that $v_d = t$. Note that S_t induces a copy of P_k^{d-1} in P_k^d.

Let us show that G is weakly $(P_k^d, C_{2\ell})$-saturated. First, add all edges uv of P_k^d where $u, v \in S_0$ in some order such that each added edge creates a new copy of $C_{2\ell}$. This is

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

possible by definition of G_{d-1}. Now, for each $t = 1, \dots, k-1$, in turn, we add all edges between vertices of S_t. By induction on t we can assume that all edges between vertices in S_{t-1} are already present. Given an edge uv where $u, v \in S_t$, let $u', v' \in S_{t-1}$ where $u'_{d-1} = u_{d-1}$ and $v'_{d-1} = v_{d-1}$. We obtain a copy of $C_{2\ell}$ by taking the edges $u'u$, uv, vv' and a path of length $2\ell - 3$ from v' to u' in S_{t-1}. This completes the proof in this case.

Case 2. $k > 2$.

Let G_{k-1} be a weakly $(P^d_{k-1}, C_{2\ell})$-saturated graph with $(k-1)^d - 1$ edges. For each set $S \subseteq [d]$, let $m(S)$ be the minimum element of S and let Y_S be the set of all vertices v of P^d_k such that S is precisely the set of coordinates on which v is equal to $k-1$. We add to G every edge uv of P^d_k such that

- $u, v \in Y_\emptyset$ and $uv \in E(G_{k-1})$, or
- $u \in Y_S$ and $v \in Y_{S \setminus \{m(S)\}}$ for some non-empty $S \subseteq [d]$.

It is clear that $|E(G)| = k^d - 1$.

Let us show that G is weakly $(P^d_k, C_{2\ell})$-saturated. First, add all edges uv where $u, v \in Y_\emptyset$ such that each added edge creates a new copy of $C_{2\ell}$. This is possible by definition of G_{k-1}.

Next, for each non-empty set $S \subseteq [d]$, in order of increasing cardinality, add all edges uv where $u, v \in Y_S$. Inductively, we can assume that all edges between vertices of $Y_{S \setminus \{m(S)\}}$ are already present. Note that $Y_{S \setminus \{m(S)\}}$ induces a copy of $P^{d-|S \setminus \{m(S)\}|}_{k-1}$ where $|S \setminus \{m(S)\}| \leqslant d-2$. Let u' and v' be the unique neighbours of u and v in $Y_{S \setminus \{m(S)\}}$, respectively. We obtain a copy of $C_{2\ell}$ by taking the edges $u'u$, uv, vv' and a path of length $2\ell - 3$ from v' to u' in $Y_{S \setminus \{m(S)\}}$.

Finally, for each set S, in order of increasing cardinality, and element j of $S \setminus \{m(S)\}$, add every edge uv where $u \in Y_S$ and v differs from u on coordinate j. Let u' be the neighbour of u in $Y_{S \setminus \{m(S)\}}$ and let v' be the neighbour of v in $Y_{S \setminus \{j, m(S)\}}$. If $\ell = 2$, then we obtain a copy of C_4 by taking the edges $u'u$, uv, vv' and $u'v'$. Otherwise, let w be the vertex of $Y_{S \setminus \{m(S)\}}$ obtained from u' by decreasing coordinate $m(S)$ by one. Further, let w' be vertex of $Y_{S \setminus \{j, m(S)\}}$ obtained from w by decreasing coordinate j by one. We obtain a copy of $C_{2\ell}$ by taking the edges $w'w$, wu', $u'u$, uv, vv' and a path of length $2\ell - 5$ from v' to w' in $Y_{S \setminus \{j, m(S)\}}$. Note that this is possible since w' is adjacent to v' and the subgraph induced by $Y_{S \setminus \{j, m(S)\}}$ is isomorphic to $P^{d-|S \setminus \{j, m(S)\}|}_{k-1}$. This completes the proof. □

4. Open problems

Many interesting open problems remain, and we mention some of them here.

4.1. Minimum saturation

Recall that we were able to prove a better upper bound on $\mathrm{sat}(Q_d, Q_m)$ when d is of the form $6m(2^t - 1)$ than for general values of d (see Remark 2.10). It is not clear whether this is simply an artifact of our proof, or part of a more general phenomenon. We ask the following.

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

Question 4.1. *For fixed $m \geqslant 2$, does*

$$\lim_{d\to\infty} \frac{\mathrm{sat}(Q_d, Q_m)}{2^d}$$

exist?

As it stands, the best known lower bound on $\mathrm{sat}(Q_d, Q_m)$ is $(m - 1 - o(1))2^d$, which follows from Theorem 1.3 and (1.1). We doubt that this bound is tight. We ask the following.

Question 4.2. *For fixed $m \geqslant 2$, let $c_m := \liminf_{d\to\infty} \mathrm{sat}(Q_d, Q_m)/2^d$. Does $c_m/m \to \infty$?*

It would also be interesting to study saturation numbers in the grid.

Question 4.3. *Fix $m \geqslant 2$, $r \geqslant 2$ and $k \geqslant \max\{r, 3\}$, and let $d \to \infty$. Is it true that*

$$\mathrm{sat}^*(P_k^d, P_r^m) = o(|E(P_k^d)|)?$$

How about

$$\mathrm{sat}^*(P_k^d, P_r^m) = O(k^d)?$$

Here, the notation sat^* indicates that we are considering axis aligned copies of P_r^m in P_k^d.

Moving beyond the finite case, one could also consider grids inside the infinite square lattice $\mathbb{Z}^d$, where vertices are adjacent if they have ℓ_1-distance 1.

Question 4.4. *For $d \geqslant m \geqslant 2$, what is the infimal density of a Q_m-saturated subgraph of the d-dimensional square lattice? More generally, what about the d-dimensional lattice that is saturated with respect to axis aligned copies of P_r^m?*

Here the density is naturally defined as the lim sup of densities inside large boxes. For both of the last two questions, it would also be interesting to know what happens if we do not require our (sub)grids to be axis aligned.

Recall that K_k^d is the graph with vertex set $\{0, \ldots, k-1\}^d$ in which two vertices are adjacent if they differ (by any amount) in exactly one coordinate. In one of the original papers on minimum saturated graphs, Erdős, Hajnal and Moon [20] determined $\mathrm{sat}(K_k, K_r)$ for all values of k and r. One could also consider the following multi-dimensional analogue.

Problem 4.5. *For $k \geqslant r \geqslant 2$ and $d \geqslant m \geqslant 1$, determine* $\mathrm{sat}(K_k^d, K_r^m)$.

The problem of determining $\mathrm{sat}(K_k, C_\ell)$ was raised by Bollobás [10] and has been extensively studied. The case $\ell = 3$ is trivial, and the cases $\ell = 4$ and $\ell = 5$ were solved by Ollmann [33] and Chen [12, 13], respectively. The first non-trivial bounds for general

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

ℓ were given by Barefoot, Clark, Entringer, Porter, Székely and Tuza [8], who proved that there are positive constants c_1 and c_2 such that, for $\ell \neq 8, 10$,

$$(1 + c_1/\ell)n \leqslant \mathrm{sat}(K_k, C_\ell) \leqslant (1 + c_2/\ell)n.$$

The value of c_2 was improved by Gould, Łuczak and Schmitt [25]. Currently, the best general upper and lower bounds are due to Füredi and Kim [22]. It is natural to ask about cycles in the hypercube.

Problem 4.6. *For $\ell \geqslant 2$ and $d \geqslant \log_2(2\ell)$, determine* $\mathrm{sat}(Q_d, C_{2\ell})$.

More generally, similar questions can be asked for cycles in P_k^d or K_k^d. We remark that the related problem of determining $\mathrm{ex}(Q_d, C_{2\ell})$ was proposed by Erdős [19] and is very well studied; see, *e.g.*, [2, 7, 11, 15, 16, 17, 23, 24].

4.2. Weak saturation

Recall that Theorem 1.4 provides the explicit value of the weak saturation number of axis aligned copies of P_r^m in P_k^d. However, for most values of m, d, r, k, the graph that we construct contains several 'bent' copies of P_r^m. We ask about the weak saturation number of general copies of P_r^m in P_k^d.

Problem 4.7. *For general $1 \leqslant m \leqslant d$ and $2 \leqslant r \leqslant k$, determine* $\mathrm{wsat}(P_k^d, P_r^m)$.

In [10], Bollobás conjectured that $\mathrm{wsat}(K_k, K_r) = \mathrm{sat}(K_k, K_r)$ for all k and r. This was proved for $r < 7$ in [9] and for general r by Alon [1] and Kalai [28, 29] using methods from exterior algebra and matroid theory. A multi-dimensional version of this problem also seems interesting.

Problem 4.8. *For general $1 \leqslant m \leqslant d$ and $2 \leqslant r \leqslant k$, determine* $\mathrm{wsat}(K_k^d, K_r^m)$.

As was mentioned in Section 3, the 'vertex version' of Problem 4.8 was solved in [6].

Theorem 1.5 leaves open the problem of determining $\mathrm{wsat}(P_k^d, C_{2\ell})$ when $d < \ell$. We ask the following.

Question 4.9. *Given $k \geqslant 2$ and $\ell > d \geqslant 2$ such that P_k^d contains a cycle of length 2ℓ, what is the value of* $\mathrm{wsat}(P_k^d, C_{2\ell})$?

4.3. Semi-saturation

Johnson and Pinto [27] also study another type of saturation problem, which is often referred to as semi-saturation [18, 22] (other terms have also been used: see [10, 28, 32, 34, 36]). Given graphs F and H, say that a spanning subgraph G of F is (F, H)-*semi-saturated* if, for every edge $e \in E(F) \setminus E(G)$, the graph $G + e$ contains more copies of H than G does. Note that G may contain H as a subgraph. Let $\mathrm{ssat}(F, H)$ denote the *semi-saturation number*, which is the minimum number of edges in a (F, H)-semi-saturated graph. Then it

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

is easy to see that

$$\mathrm{wsat}(F,H) \leqslant \mathrm{ssat}(F,H) \leqslant \mathrm{sat}(F,H).$$

Johnson and Pinto [27] proved that $\mathrm{ssat}(Q_d, Q_m) \leqslant (m^2 + m/2)2^d$. Based on this result and (1.2), they asked whether $\mathrm{ssat}(Q_d, Q_2)$ is equal to $\mathrm{sat}(Q_d, Q_2)$ for sufficiently large d and, if not, whether

$$\limsup_{d\to\infty} \frac{\mathrm{sat}(Q_d,Q_2)}{2^d} > \limsup_{d\to\infty} \frac{\mathrm{ssat}(Q_d,Q_2)}{2^d}.$$

Given Theorem 1.2, it now seems natural to pose these questions more generally.

Question 4.10. *For fixed $m \geqslant 2$, is $\mathrm{ssat}(Q_d, Q_m)$ is equal to $\mathrm{sat}(Q_d, Q_m)$ for sufficiently large d? If not, is*

$$\limsup_{d\to\infty} \frac{\mathrm{sat}(Q_d,Q_m)}{2^d} > \limsup_{d\to\infty} \frac{\mathrm{ssat}(Q_d,Q_m)}{2^d}\,?$$

Of course, the open problems on saturation mentioned also make sense for semi-saturation.

References

[1] N. Alon (1985) An extremal problem for sets with applications to graph theory. *J. Combin. Theory Ser. A* **40** 82–89.

[2] Baber, R. (2012) Turán densities of hypercubes. `arXiv:1201.3587v2`

[3] Balister, P. N., Bollobás, B., Lee, J. D. and Narayanan, B. P. Line percolation, *Random Struct. Alg.*, to appear.

[4] Balogh, J. and Bollobás, B. (2006) Bootstrap percolation on the hypercube. *Probab. Theory Rel. Fields* **134** 624–648.

[5] Balogh, J., Bollobás, B. and Morris, R. (2012) Graph bootstrap percolation. *Random Struct. Alg.* **41** 413–440.

[6] Balogh, J., Bollobás, B., Morris, R. and Riordan, O. (2012) Linear algebra and bootstrap percolation. *J. Combin. Theory Ser. A* **119** 1328–1335.

[7] Balogh, J., Hu, P., Lidický, B. and Liu, H. (2014) Upper bounds on the size of 4- and 6-cycle-free subgraphs of the hypercube. *European J. Combin.* **35** 75–85.

[8] Barefoot, C. A., Clark, L. H., Entringer, R. C., Porter, T. D., Székely, L. A. and Tuza, Z. (1996) Cycle-saturated graphs of minimum size. *Discrete Math.* **150** 31–48.

[9] Bollobás, B. (1968) Weakly k-saturated graphs. In *Beiträge zur Graphentheorie: Kolloquium, Manebach 1967*, Teubner, pp. 25–31.

[10] Bollobás, B. (1978) *Extremal graph theory*, Vol. 11 of *London Mathematical Society Monographs*, Academic.

[11] Brass, P., Harborth, H. and Nienborg, H. (1995) On the maximum number of edges in a C_4-free subgraph of Q_n. *J. Graph Theory* **19** 17–23.

[12] Chen, Y.-C. (2009) Minimum C_5-saturated graphs. *J. Graph Theory* **61** 111–126.

[13] Chen, Y.-C. (2011) All minimum C_5-saturated graphs. *J. Graph Theory* **67** 9–26.

[14] Choi, S. and Guan, P. (2008) Minimum critical squarefree subgraph of a hypercube. In *Proc. 39th Southeastern International Conference on Combinatorics, Graph Theory and Computing*, *Congr. Numer.* **189** 57–64.

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press

[15] Chung, F. R. K. (1992) Subgraphs of a hypercube containing no small even cycles. *J. Graph Theory* **16** 273–286.
[16] Conder, M. (1993) Hexagon-free subgraphs of hypercubes. *J. Graph Theory* **17** 477–479.
[17] Conlon, D. (2010) An extremal theorem in the hypercube. *Electron. J. Combin.* **17** Research Paper 111.
[18] Day, A. N. (2014) Saturated graphs of prescribed minimum degree. `arXiv:1407.6664v1`
[19] Erdős, P. (1984) On some problems in graph theory, combinatorial analysis and combinatorial number theory. In *Graph Theory and Combinatorics: Cambridge 1983*, Academic, pp. 1–17.
[20] Erdős, P., Hajnal, A. and Moon, J. W. (1964) A problem in graph theory. *Amer. Math. Monthly* **71** 1107–1110.
[21] Faudree, J. R., Faudree, R. J. and Schmitt, J. R. (2011) A survey of minimum saturated graphs. *Electron. J. Combin.* **18** Dynamic Survey 19.
[22] Füredi, Z. and Kim, Y. (2013) Cycle-saturated graphs with minimum number of edges. *J. Graph Theory* **73** 203–215.
[23] Füredi, Z. and Özkahya, L. (2009) On 14-cycle-free subgraphs of the hypercube. *Combin. Probab. Comput.* **18** 725–729.
[24] Füredi, Z. and Özkahya, L. (2011) On even-cycle-free subgraphs of the hypercube. *J. Combin. Theory Ser. A* **118** 1816–1819.
[25] Gould, R., Łuczak, T. and Schmitt, J. (2006) Constructive upper bounds for cycle-saturated graphs of minimum size. *Electron. J. Combin.* **13** Research Paper 29.
[26] Hamming, R. W. (1950) Error detecting and error correcting codes. *Bell System Tech. J.* **29** 147–160.
[27] Johnson, J. R. and Pinto, T. (2014) Saturated subgraphs of the hypercube. `arXiv:1406.1766v1`
[28] Kalai, G. (1984) Weakly saturated graphs are rigid. In *Convexity and Graph Theory: Jerusalem 1981*, Vol. 87 of *North-Holland Mathematics Studies*, North-Holland, pp. 189–190.
[29] Kalai, G. (1985) Hyperconnectivity of graphs. *Graphs Combin.* **1** 65–79.
[30] MacWilliams, F. J. and Sloane, N. J. A. (1977) *The Theory of Error-Correcting Codes I*, Vol. 16 of *North-Holland Mathematical Library*, North-Holland.
[31] MacWilliams, F. J. and Sloane, N. J. A. (1977) *The Theory of Error-Correcting Codes II*, Vol. 16 of *North-Holland Mathematical Library*, North-Holland.
[32] Morrison, N., Noel, J. A. and Scott, A. (2014) On saturated k-Sperner systems. *Electron. J. Combin.* **21** Paper 3.22.
[33] Ollmann, L. T. (1972) $K_{2,2}$ saturated graphs with a minimal number of edges. In *Proc. 3rd Southeastern Conference on Combinatorics, Graph Theory, and Computing: Boca Raton 1972*, pp. 367–392.
[34] Pikhurko, O. (1999) Extremal hypergraphs. PhD thesis, University of Cambridge.
[35] Pikhurko, O. (2001) Weakly saturated hypergraphs and exterior algebra. *Combin. Probab. Comput.* **10** 435–451.
[36] Tuza, Z. (1992) Asymptotic growth of sparse saturated structures is locally determined. *Discrete Math.* **108** 397–402.
[37] Zykov, A. A. (1949) On some properties of linear complexes. *Mat. Sbornik* (N.S.) **24 (66)** 163–188.

https://doi.org/10.1017/S0963548316000122 Published online by Cambridge University Press