On the Query Complexity of Testing Orientations for being Eulerian*

Eldar Fischer [†] Oded Lachish [‡] Arie Matsliah [§]

Ilan Newman [¶] Orly Yahalom [‖]

Abstract

We consider testing directed graphs for being Eulerian in the orientation model introduced in [15]. Despite the local nature of the property of being Eulerian, it turns out to be significantly harder for testing than other properties studied in the orientation model. We show a non-constant lower bound on the query complexity of 2-sided tests and a linear lower bound on the query complexity of 1-sided tests for this property. On the positive side, we give several 1-sided and 2-sided tests, including a sub-linear query complexity 2-sided test for general graphs. For special classes of graphs, including bounded-degree graphs and expander graphs, we provide improved results. In particular, we give a 2-sided test with constant query complexity for dense graphs, as well as for expander graphs with a constant expansion parameter.

1 Introduction

Property Testing deals with the following relaxation of decision problems: Given a property $\mathcal{P}$, an input structure S and $\epsilon > 0$, distinguish between the case where S satisfies $\mathcal{P}$ and the case where S is ϵ-far from satisfying $\mathcal{P}$. Roughly speaking, an input S is said to be *ϵ-far* from satisfying a property $\mathcal{P}$ if more than an ϵ-fraction of its values must be modified in order to make it satisfy the property. Algorithms which distinguish with high probability between the two cases are called *property testers* or simply *testers* for $\mathcal{P}$. Furthermore, a tester for $\mathcal{P}$ is said to be *1-sided* if it never rejects an input that satisfies $\mathcal{P}$. Otherwise, the tester is called *2-sided*. We say that a tester is *adaptive* if some of the choices of the locations for which the input is queried may depend on the returned values (answers) of previous queries. Otherwise, the tester is called *non-adaptive*. Property Testing normally deals with problems involving a very large input or a costly retrieval procedure. Thus, the number of queries of input values, rather than the computation time, is considered to be the most expensive resource.

*A preliminary version appeared in the proceedings of the 12^{th} RANDOM and 11^{th} APPROX (2008).

†Computer Science Department, Technion – Israel Institute of Technology, Haifa 32000, Israel. Email: eldar@cs.technion.ac.il. Research supported in part by an ERC-2007-StG grant number 202405-2 and by an ISF grant number 1101/06.

‡Centre for Discrete Mathematics and its Applications (DIMAP), Warwick, UK. Email: oded@dcs.warwick.ac.uk.

§Centrum voor Wiskunde en Informatica (CWI), Amsterdam, Holland, E-mail: ariem@cwi.nl

¶Computer Science Department, Haifa University, Haifa 31905, Israel. Email: ilan@cs.haifa.ac.il. Research supported in part by an ISF grant number 1101/06.

‖Department of Software Engineering, ORT Braude College, P.O. Box 78, Karmiel, 21982, Israel. Email: oyahalom@braude.ac.il.

Property Testing has been a very active field of research since it was initiated by Blum, Luby and Rubinfeld [5]. The general definition of Property Testing was formulated by Rubinfeld and Sudan [25], who were interested mainly in testing algebraic properties. The study of Property Testing for combinatorial objects, and mainly for labelled graphs, began in the seminal paper of Goldreich, Goldwasser and Ron [12]. They introduced the *dense graph model*, where the graph is assumed to be represented by an adjacency matrix, and the distance function is computed accordingly. For comprehensive surveys on Property Testing see [24, 8].

The dense graph model is in a sense too lenient, since for n-vertex graphs, the distance function allows adding and removing $o(n^2)$ edges, regardless of the number of actual edges in the graph. Thus, many interesting properties, such as connectivity in undirected or directed graphs, are trivially testable in this model, as all the graphs are close to satisfying the property. In recent years, researchers have studied several alternative models for graph testing, including the *bounded-degree* graph model of [13], in which a sparse representation of sparse graphs is considered, and the *general density* model (also called the *mixed* model) of [21] and [17]. In these models, the distance function allows edge insertions and deletions whose number is at most a fraction of the number of the edges in the original graph.

Property Testing of directed graphs has also been studied in the context of the above models [1, 3]. Here we continue the study of testing properties of directed graphs in the orientation model, which started in [15] and followed in [14] and [7]. In this model, an underlying undirected graph $G = (V, E)$ is given in advance, and the actual input is an orientation $\overrightarrow{G}$ of G, in which every edge in E has a direction. Our testers may access the input using edge queries. That is, every query concerns an edge $e \in E$, and the answer to the query is the direction of e in $\overrightarrow{G}$. An orientation $\overrightarrow{G}$ of G is called *ϵ-close* to a property $\mathcal{P}$ if it can be made to satisfy $\mathcal{P}$ by inverting at most an ϵ-fraction of the edges of G, and otherwise $\overrightarrow{G}$ is said to be *ϵ-far* from $\mathcal{P}$.

Note that the distance function in the orientation model naturally depends on the size of the underlying graph and is independent of representation details. Moreover, the testing algorithm may strongly depend on the structure of the underlying graph. The model is strict in that the distance function allows only edge inversions, but no edge insertions or deletions. On the other hand, we assume that our algorithms have a full knowledge of the underlying graph, whose size is roughly the same as the input size. Viewing the underlying graph as a parameter that the testing algorithm receives in advance, we say that the orientation model is an example of a *massively parameterized* model. Other examples that can be thought of as massively parameterized models appear in [20], where the property is represented by a known bounded-width branching program, in [9], where the input is a vertex-coloring of a known graph, and in other works.

In this paper we consider the property of being Eulerian, which was presented in [14] as one of the natural orientation properties whose query complexity was still unknown. A directed graph $\overrightarrow{G}$ is called *Eulerian* if for every vertex v in the graph, the in-degree of v is equal to its out-degree (in addition, it is common to define Eulerian graphs as connected, but as we explain later, our algorithms and proofs work equally well whether we require connectivity or not). An undirected graph G has an Eulerian orientation $\overrightarrow{G}$ if and only if all the degrees of G are even. Such an undirected graph is called Eulerian also. Throughout the paper we assume that our underlying undirected graph G is Eulerian.

Eulerian graphs and Eulerian orientations have attracted researchers since the dawn of graph theory in 1736, when Leonard Euler published his solution for the famous "Königsberg bridge problem". Throughout the years, Eulerian graphs have been the subject of extensive research (e.g. [23, 18, 26, 19, 6, 2]; see [10, 11] for an extensive survey). Aside from their appealing theoretic characteristics, Eulerian graphs have been studied in the context of networking [16] and genetics [22].

Testing for being Eulerian in the orientation model is equivalent to the following problem. We have a known network (a communication network, a transportation system or a piping system) where every edge can transport a unit of "flow" in both directions. Our goal is to know whether the network is "balanced", or far from being balanced, where being balanced means that the number of flows entering every node in the network is equal to the number of flows exiting it (so there is no "accumulation" in the nodes). To examine the network, we detect the flow direction in selected individual edges, which is deemed to be the expensive operation.

The main difficulty in testing orientations for being Eulerian arises from the fact that an orientation might have a small number of unbalanced vertices, and each of them with a small imbalance, and yet be far from being Eulerian. This is since trying to balance an unbalanced vertex by inverting some of its incident edges may violate the balance of its balanced neighbors. Thus, we must continue to invert edges along a directed path between a vertex with a positive imbalance and a vertex with a negative imbalance. We call such a path a *correction path*. A main component of our work is giving upper bounds for the length of the correction paths. In this context we note the work of Babai [2], who studied the ratio between the diameter of Eulerian digraphs and the diameter of their underlying undirected graphs. While he gave an upper bound for this ratio for vertex-transitive graphs, he showed an infinite family of undirected graphs with diameter 2 which have an Eulerian orientation with diameter $\Omega(n^{1/3})$.

Our upper bounds are based on three "generic" tests, one 1-sided test and two 2-sided tests. Instead of receiving ϵ as a parameter, the generic tests receive a parameter p, which stands for the number of required correction paths in an orientation that is far from being Eulerian. We hence call these tests *p-tests*. We later derive ϵ-tests from the p-tests by proving two lower bounds for p. The first one gives an efficient test for dense graphs and the second one gives an efficient test for expander graphs. Finally, we show how to use variations of the expander tests for obtaining a 1-sided test and a 2-sided test for general graphs, using a decomposition ("chopping") procedure into subgraphs that are roughly expanders. The 2-sided test that we obtain this way has a sublinear query complexity for every graph. Unfortunately, our chopping procedure is adaptive and has an exponential computational time in $|E|$. All of our other algorithms are non-adaptive and their computational complexity is of the same order as their query complexity.

On the negative side, we provide several lower bounds. We show that any 1-sided test for being Eulerian must use $\Omega(m)$ queries for some graphs. For bounded-degree graphs, we use the toroidal grid to prove non-constant 1-sided and 2-sided lower bounds. These bounds are noteworthy, as bounded-degree graphs have a constant size witness for not being Eulerian, namely the edges incident with one unbalanced vertex. In contrast, the *st*-connectivity property, whose witness must include a cut in the graph, is testable with a constant number of queries in the orientation model [7]. In other testing models there are known super-constant lower bounds also for properties which have constant-size witness, e.g., [4] prove a linear lower bound for testing whether a truth assignment satisfies a known 3CNF formula. However, most of these bounds are for properties that have stronger expressive power than that of being Eulerian.

Tables 1 and 2 summarize our upper and lower bounds, respectively. Here and throughout the paper, we set $n = |V|$ and $m = |E|$, let Δ be the maximum vertex-degree in G, and set $d \stackrel{\text{def}}{=} m/n$. The tilde notation hides polylogarithmic factors.

Table 1: Upper bounds

Result	1-sided tests	2-sided tests
Graphs with large d	$O\left(\frac{\Delta m}{\epsilon^2 d^2}\right)$	$\min\left\{\widetilde{O}\left(\frac{m^3}{\epsilon^6 d^6}\right), \widetilde{O}\left(\frac{\sqrt{\Delta} m}{\epsilon^2 d^2}\right)\right\}$
α-expanders	$O\left(\frac{\Delta \log(1/\epsilon)}{\alpha\epsilon}\right)$	$\min\left\{\widetilde{O}\left(\left(\frac{\log(1/\epsilon)}{\alpha\epsilon}\right)^3\right), \widetilde{O}\left(\frac{\sqrt{\Delta}\log(1/\epsilon)}{\alpha\epsilon}\right)\right\}$
General graphs (adaptive)	$O\left(\frac{(\Delta m \log m)^{2/3}}{\epsilon^{4/3}}\right)$	$\min\left\{\widetilde{O}\left(\frac{\Delta^{1/3} m^{2/3}}{\epsilon^{4/3}}\right), \widetilde{O}\left(\frac{\Delta^{3/16} m^{3/4}}{\epsilon^{5/4}}\right)\right\}$

Table 2: Lower bounds

Result	1-sided tests	2-sided tests
General graphs	$\Omega(m)$	—
Bounded-degree graphs, non-adaptive tests	$\Omega(m^{1/4})$	$\Omega\left(\sqrt{\frac{\log m}{\log\log m}}\right)$
Bounded-degree graphs, adaptive tests	$\Omega(\log m)$	$\Omega(\log\log m)$

The rest of the paper is organized as follows. Section 2 includes general definitions and lemmas to be used in the sequel. In Section 3 we provide a simple 1-sided lower bound for general graphs. Section 4 is dedicated to our three p-tests, which distinguish between Eulerian orientations and orientations with many correction paths. In Section 5 we give a lower bound on the number of correction paths as a function of the average degree in the graph, and derive our tests for graphs with high average degree. In Section 6 we give such a bound and tests for expander graphs. Section 7 considers testing subgraphs that we call "lame" expanders, providing results for them that are similar to those obtained for expanders, and are used in the sequel. Section 8 presents our most general tests, which use the results of Section 7. Section 9 gives our lower bound for bounded-degree graphs. Finally, Section 10 summarizes the still open questions.

2 Preliminaries

Throughout the paper, we assume a fixed and known underlying graph $G = (V, E)$ which is Eulerian, that is, for every $v \in V$, the degree $\deg(v)$ of v is even.

Given an orientation $\overrightarrow{G} = (V, \overrightarrow{E})$ and a vertex $v \in V$, let $\mathrm{indeg}_{\overrightarrow{G}}(v)$ denote the in-degree of v with respect to $\overrightarrow{G}$ and let $\mathrm{outdeg}_{\overrightarrow{G}}(v)$ denote the out-degree of v with respect to $\overrightarrow{G}$. We define the *imbalance* of v in $\overrightarrow{G}$ as $\mathrm{ib}_{\overrightarrow{G}}(v) \stackrel{\mathrm{def}}{=} \mathrm{outdeg}_{\overrightarrow{G}}(v) - \mathrm{indeg}_{\overrightarrow{G}}(v)$. In the following, we sometimes omit the subscript $\overrightarrow{G}$ whenever it is obvious from the context. We say that a vertex $v \in V$ is a *spring* in $\overrightarrow{G}$ if $\mathrm{ib}_{\overrightarrow{G}}(v) > 0$. We say that v is a *drain* in $\overrightarrow{G}$ if $\mathrm{ib}_{\overrightarrow{G}}(v) < 0$. If $\mathrm{ib}_{\overrightarrow{G}}(v) = 0$ then we say that v

is *balanced* in $\overrightarrow{G}$. We say that $\overrightarrow{G}$ is *Eulerian* if all its vertices are balanced. Since all the vertices of G are of even degree, there always exists some Eulerian orientation $\overrightarrow{G}$ of G.

We note that it is common to require connectivity from an Eulerian undirected graph and strong connectivity from a directed Eulerian graph. However, all our algorithms and proofs work also for the case where our underlying graph G is not connected. Furthermore, it is easy to see that if G is connected, then every Eulerian orientation $\overrightarrow{G}$ of G is strongly connected. Thus, if we are interested in testing whether $\overrightarrow{G}$ is strongly connected, in addition to having balanced vertices, we simply add to our tests a phase which checks the underlying graph G and rejects if it is not connected. We henceforth ignore the connectivity criterion.

We next observe that testing whether $\overrightarrow{G}$ is Eulerian is trivial for $\epsilon \geq \frac{1}{2}$.

Observation 2.1 *Every orientation $\overrightarrow{G}$ of G is $\frac{1}{2}$-close to being Eulerian.*

Proof. Let $\overrightarrow{G_1}$ be an arbitrary Eulerian orientation of G. Let $\overrightarrow{G_2} = \overleftarrow{G_1}$, namely, the orientation derived from G_1 by inverting all the edges. Clearly, $\overrightarrow{G_2}$ is Eulerian as well, since inverting all the edges maintains the absolute value of the imbalance of all vertices. Now, for every edge $e \in E$, the direction of e in $\overrightarrow{G}$ is the same as in $\overrightarrow{G_1}$ if and only if it is opposite to the direction of e in $\overrightarrow{G_2}$. Hence, $\mathrm{dist}(\overrightarrow{G}, \overrightarrow{G_1}) = 1 - \mathrm{dist}(\overrightarrow{G}, \overrightarrow{G_2})$, and therefore, $\overrightarrow{G}$ is $\frac{1}{2}$-close to either $\overrightarrow{G_1}$ or $\overrightarrow{G_2}$. ∎

We conclude with some notation that is useful in the following. Given a set $U \subseteq V$, we let

$$E(U) \stackrel{\text{def}}{=} \{\{u,v\} \in E \mid u,v \in U\},$$

$$\overrightarrow{E}(U) \stackrel{\text{def}}{=} \{(u,v) \in \overrightarrow{E} \mid u,v \in U\},$$

$$\partial U \stackrel{\text{def}}{=} \{\{u,v\} \in E \mid u \in U, v \notin U\},$$

and

$$\overrightarrow{\partial} U \stackrel{\text{def}}{=} \{(u,v) \in \overrightarrow{E} \mid u \in U, v \notin U\}.$$

Given two disjoint sets $U, W \subseteq V$, let

$$E(U,W) \stackrel{\text{def}}{=} \{\{u,w\} \in E \mid u \in U, w \in W\}$$

and

$$\overrightarrow{E}(U,W) \stackrel{\text{def}}{=} \{(u,w) \in \overrightarrow{E} \mid u \in U, w \in W\}.$$

2.1 Correction subgraphs and p-tests

Let $\overrightarrow{G}$ be an orientation of G. Given a subgraph $\overrightarrow{H} = (V_H, \overrightarrow{E_H})$ of $\overrightarrow{G}$ (that is, a directed graph where $V_H \subseteq V$ and $\overrightarrow{E_H} \subseteq \overrightarrow{E}$) we define $\overrightarrow{G}_{\overleftarrow{H}} \stackrel{\text{def}}{=} (V, \overrightarrow{E}_{\overleftarrow{H}})$ to be the orientation of G derived from $\overrightarrow{G}$ by inverting all the edges of $\overrightarrow{H}$. Namely,

$$\overrightarrow{E}_{\overleftarrow{H}} \;=\; \overrightarrow{E} \setminus \overrightarrow{E_H} \;\cup\; \{(v,u) \in (V_H)^2 \mid (u,v) \in \overrightarrow{E_H}\}.$$

We say that $\overrightarrow{H}$ is a *correction subgraph* of $\overrightarrow{G}$ if $\overrightarrow{G}_{\overleftarrow{H}}$ is Eulerian. Note that in such a case, $\overrightarrow{G}$ is $\frac{|\overrightarrow{E_H}|}{m}$-close to being Eulerian.

Lemma 2.2 *A subgraph $\overrightarrow{H}$ of $\overrightarrow{G}$ is a correction subgraph if and only if the following conditions hold for every $v \in V$:*

1. *If $v \notin V_H$, then v is balanced in $\overrightarrow{G}$.*
2. *If $v \in V_H$, then $\mathrm{ib}_{\overrightarrow{H}}(v) = \frac{1}{2} \cdot \mathrm{ib}_{\overrightarrow{G}}(v)$.*

In particular, a vertex v is a spring in $\overrightarrow{G}$ if and only if it is a spring in $\overrightarrow{H}$, and v is drain in $\overrightarrow{G}$ if and only if it is a drain in $\overrightarrow{H}$.

Proof. Remember that $\overrightarrow{H}$ is a correction subgraph of $\overrightarrow{G}$ if and only if $\mathrm{ib}_{\overrightarrow{G}_{\overleftarrow{H}}}(v) = 0$ for every $v \in V$. The proof follows from the following facts, which can be easily verified:

1. If $v \notin V_H$, then $\mathrm{ib}_{\overrightarrow{G}_{\overleftarrow{H}}}(v) = \mathrm{ib}_{\overrightarrow{G}}(v)$.
2. If $v \in V_H$, then $\mathrm{ib}_{\overrightarrow{G}_{\overleftarrow{H}}}(v) = \mathrm{ib}_{\overrightarrow{G}}(v) - 2 \cdot \mathrm{ib}_{\overrightarrow{H}}(v)$.

■

Since we assume that G is Eulerian, there always exists some correction subgraph of $\overrightarrow{G}$. Furthermore, without loss of generality, we may focus on acyclic correction graphs.

Observation 2.3 *For any orientation $\overrightarrow{G}$ of G and for any correction subgraph $\overrightarrow{H_1}$ of $\overrightarrow{G}$, there exists an acyclic correction subgraph $\overrightarrow{H}$ of $\overrightarrow{G}$ that is a subgraph of $\overrightarrow{H_1}$.*

Proof. We obtain $\overrightarrow{H}$ from $\overrightarrow{H_1}$ as follows. While $\overrightarrow{H_1}$ is not acyclic, we arbitrarily choose a directed cycle and remove its edges from the subgraph. Note that this operation maintains the balance factors of all the vertices of G with respect to $\overrightarrow{H_1}$. The proof thus follows from Lemma 2.2. ■

Let S be the set of springs in $\overrightarrow{G}$ and let T be the set of drains in $\overrightarrow{G}$. We say that a directed path $\overrightarrow{P} = \langle u_0, \ldots, u_k \rangle$ in $\overrightarrow{G}$ is a *spring-drain path* if $u_0 \in S$ and $u_k \in T$. Note that in this case, for any correction subgraph $\overrightarrow{H}$ of $\overrightarrow{G}$, we have by Lemma 2.2 that u_0 is a spring in $\overrightarrow{H}$ and u_k is a drain in $\overrightarrow{H}$. The following observations and lemmas follow easily.

Observation 2.4 *If $\overrightarrow{G}$ is not Eulerian then any correction subgraph $\overrightarrow{H}$ of $\overrightarrow{G}$ contains a spring-drain path.* ■

Observation 2.5 *Let $\overrightarrow{H}$ be a correction subgraph of $\overrightarrow{G}$ and let $\overrightarrow{P}$ be a spring-drain path in $\overrightarrow{H}$. Define $\overrightarrow{H} \setminus \overrightarrow{P}$ to be the graph obtained from $\overrightarrow{H}$ by removing all the edges of $\overrightarrow{P}$, that is, $\overrightarrow{E} \setminus \overrightarrow{P} \stackrel{\text{def}}{=} (V_H, \overrightarrow{E} \setminus \overrightarrow{P})$. Then $\overrightarrow{H} \setminus \overrightarrow{P}$ is a correction subgraph of $\overrightarrow{G}_{\overleftarrow{P}}$, the graph obtained from $\overrightarrow{H}$ by inverting all the edges of $\overrightarrow{P}$. Moreover, if $\overrightarrow{H}$ is acyclic then $\overrightarrow{H} \setminus \overrightarrow{P}$ is edgeless if and only if $\overrightarrow{G}_{\overleftarrow{P}}$ is Eulerian.* ■

Lemma 2.6 *If $\overrightarrow{G}$ is not Eulerian then any acyclic correction subgraph $\overrightarrow{H}$ of $\overrightarrow{G}$ is a union of $p = \frac{1}{4} \sum_{u \in V} |\mathrm{ib}(u)|$ edge-disjoint spring-drain paths. Moreover, every decomposition of $\overrightarrow{H}$ into edge-disjoint spring-drain paths has exactly p paths. In particular, removing any spring-drain path from $\overrightarrow{H}$ results in a graph that is a collection of $p - 1$ edge-disjoint spring-drain paths.*

Proof. From Observations 2.4 and 2.5, it is clear that $\overrightarrow{H}$ may be decomposed into edge-disjoint spring-drain paths. Note that when removing a spring-drain path, we reduce the sum $\sum_{u \in V} |\mathrm{ib}(u)|$ by exactly four (as we reduce the absolute value of the imbalance of both the spring and the drain by two). Therefore, every decomposition of $\overrightarrow{H}$ into disjoint spring-drain paths contains $p = \frac{1}{4} \sum_{u \in V} |\mathrm{ib}(u)|$ paths. ∎

Lemma 2.7 *Let $\overrightarrow{G}$ be an orientation of G and let $\overrightarrow{H}$ be an acyclic correction subgraph of $\overrightarrow{G}$ which is a union of p edge-disjoint spring-drain paths. Let S be the set of springs in $\overrightarrow{H}$. Then $\sum_{u \in S} \mathrm{ib}(u) = 2p$.*

Proof. By Lemma 2.2, for any spring $u \in S$, the number of spring-drain paths starting at u is $\mathrm{ib}(u)/2$. Thus $\sum_{u \in S} \mathrm{ib}(u)/2 = p$. ∎

If every correction subgraph of an orientation $\overrightarrow{G}$ is a union of at least p disjoint spring-drain paths, we will say that $\overrightarrow{G}$ is *p-far* from being Eulerian. An algorithm will be called a *p-test* for being Eulerian for some positive number p if it can distinguish between Eulerian orientations and p-far orientations. Namely, the algorithm should accept an Eulerian orientation with probability at least $2/3$ and reject a p-far orientation with probability at least $2/3$. Similarly to ϵ-tests, if a p-test accepts every Eulerian orientation with probability 1 then it is called *1-sided*, and otherwise it is called *2-sided*.

2.2 β-correction subgraphs and (p, β)-tests

Given $\beta > 0$, we say that a vertex v is *β-small* if $\deg(v) \leq \beta$ and *β-big* if $\deg(v) > \beta$. An orientation $\overrightarrow{G}$ is called *β-Eulerian* if all the β-small vertices in V are balanced in $\overrightarrow{G}$. Note that for $\beta \geq \Delta$, $\overrightarrow{G}$ is β-Eulerian if and only if $\overrightarrow{G}$ is Eulerian. A directed subgraph $\overrightarrow{H} = (V_H, \overrightarrow{E_H})$ of $\overrightarrow{G}$ is a *β-correction subgraph of $\overrightarrow{G}$* if:

1. $\overrightarrow{G}_{\overleftarrow{H}}$ is β-Eulerian.
2. For every $v \in V$ we have: $|\mathrm{ib}_{\overrightarrow{G}_{\overleftarrow{H}}}(v)| \leq |\mathrm{ib}_{\overrightarrow{G}}(v)|$, and if v is a spring (resp. drain) in $\overrightarrow{H}$ then it is either balanced or a spring (resp. drain) in $\overrightarrow{G}_{\overleftarrow{H}}$.

That is, $\overrightarrow{H}$ fixes the balance of all the β-small vertices without increasing or changing the sign of the imbalance of β-big vertices.

A directed path $\overrightarrow{P} = \langle u_0, \ldots, u_k \rangle$ in an orientation $\overrightarrow{G}$ is called a *β-spring-drain path* if $\overrightarrow{P}$ is a spring-drain path and at least one of u_0 and u_k is β-small.

Observation 2.8 *Every orientation $\overrightarrow{G}$ which is not β-Eulerian has a β-correction subgraph which is a union of edge-disjoint β-spring-drain paths.*

For $0 \leq \epsilon \leq 1$ and $\beta > 0$, we say that an orientation $\overrightarrow{G}$ is *(ϵ, β)-amendable* if there exists a β-correction subgraph $\overrightarrow{H} = (V_H, \overrightarrow{E_H})$ of $\overrightarrow{G}$ with $|\overrightarrow{E_H}| \leq \epsilon m$. Otherwise we say that $\overrightarrow{G}$ is *(ϵ, β)-unamendable*.

For $p, \beta > 0$, we say that an orientation $\overrightarrow{G}$ is *p-close to being β-Eulerian* if there exists a β-correction subgraph $\overrightarrow{H}$ of $\overrightarrow{G}$ that is a union of at most p edge-disjoint β-spring-drain paths.

Otherwise, we say that $\overrightarrow{G}$ is *p-far from being β-Eulerian.* Note that in fact the requirement from the correction subgraph to be composed of β-spring-drain paths only is not restricting in the definition of being p-far.

One can show that all the lemmas and observations that we have proved in Subsection 2.1 for correction subgraphs and spring-drain paths can be adapted to β-correction subgraphs and β-spring-drain paths. In particular:

Lemma 2.9 *If $\overrightarrow{G}$ is not β-Eulerian then any acyclic β-correction subgraph $\overrightarrow{H}$ of $\overrightarrow{G}$ is a union of at least $\frac{1}{4}\sum_{u\in V,\deg(u)\leq\beta}|\mathrm{ib}(u)|$ and at most $\frac{1}{2}\sum_{u\in V,\deg(u)\leq\beta}|\mathrm{ib}(u)|$ edge-disjoint β-spring-drain paths. Hence, if $\overrightarrow{G}$ is p-far from being β-Eulerian then*

$$\sum_{u\in V,\deg(u)\leq\beta}|\mathrm{ib}(u)| > 2p.$$

Proof. One can show that any acyclic β-correction subgraph is a union of edge-disjoint β-spring-drain paths, using similar techniques to those that we used in Subsection 2.1, as well as the fact that a β-correction subgraph may not increase or change the sign of the imbalance of β-big vertices. To bound the number of β-spring-drain paths in any β-correction subgraph $\overrightarrow{H}$ of an orientation $\overrightarrow{G}$, note that when we remove a β-spring-drain path from $\overrightarrow{H}$, we reduce the sum $\sum_{u\in V,\deg(u)\leq\beta}|\mathrm{ib}(u)|$ by either two or four (as we reduce the absolute value of the imbalance of both the spring and the drain by two, and at least one of them is β-small). Therefore, every decomposition of $\overrightarrow{H}$ into disjoint β-spring-drain paths contains p paths for some p between $\frac{1}{4}\sum_{u\in V,\deg(u)\leq\beta}|\mathrm{ib}(u)|$ and $\frac{1}{2}\sum_{u\in V,\deg(u)\leq\beta}|\mathrm{ib}(u)|$. ■

Given $p,\beta>0$, an algorithm is called a (p,β)*-test for being Eulerian* for some positive number p if it can distinguish between β-Eulerian orientations and orientations that are p-far from being β-Eulerian. Namely, the algorithm should accept a β-Eulerian orientation with probability at least $2/3$ and reject an orientation that is p-far from being β-Eulerian with probability at least $2/3$. As usual, a (p,β)-test is said to be *1-sided* if it accepts every β-Eulerian orientation with probability 1. Otherwise, the test is said to be *2-sided*.

3 A linear lower bound for 1-sided tests

In this section we prove that there exists no sub-linear 1-sided test for Eulerian orientations of general graphs.

Consider an algorithm that tests an orientation $\overrightarrow{G}$ of G. At a given moment, we represent the edges that the algorithm has queried so far by a directed *knowledge graph* $\overrightarrow{H}=(V,\overrightarrow{E_H})$, where $\overrightarrow{E_H}\subseteq\overrightarrow{E}$. We say that a cut $M=(U,V\setminus U)$ of G is *valid* with respect to a knowledge graph $\overrightarrow{H}$ if

$$|\overrightarrow{E_H}(U,V\setminus U)| \;\leq\; \frac{1}{2}|E(U,V\setminus U)| \quad\text{and}\quad |\overrightarrow{E_H}(V\setminus U,U)| \;\leq\; \frac{1}{2}|E(U,V\setminus U)|.$$

Otherwise, M is called *invalid*. Clearly, if $\overrightarrow{G}$ is Eulerian, then every knowledge graph $\overrightarrow{H}$ of $\overrightarrow{G}$ contains only valid cuts. We show that any 1-sided test for being Eulerian must obtain a knowledge graph that contains some invalid cut in order to reject $\overrightarrow{G}$.

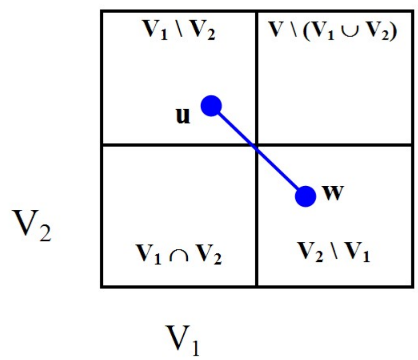
We say that a (valid) cut $M=(U,V\setminus U)$ of G is *restricting* with respect to $\overrightarrow{H}$ if

$$|\overrightarrow{E_H}(U, V \setminus U)| \;=\; \frac{1}{2}|E(U, V \setminus U)| \quad \text{or} \quad |\overrightarrow{E_H}(V \setminus U, U)| \;=\; \frac{1}{2}|E(U, V \setminus U)|.$$

Note that, given that $\overrightarrow{G}$ is Eulerian, a restricting cut with respect to $\overrightarrow{H}$ forces the orientations of all the unqueried edges in the cut. We say that two restricting cuts M_1 and M_2 are *conflicting* (with respect to a knowledge graph $\overrightarrow{H}$) if they force contrasting orientations of at least one unqueried edge.

Lemma 3.1 *Let $\overrightarrow{H}$ be a knowledge graph of $\overrightarrow{G}$ and suppose that all the cuts in G are valid with respect to $\overrightarrow{H}$. Then there are no conflicting cuts with respect to $\overrightarrow{H}$.*

Figure 1: *A schematic illustration of the conflicting cuts $M_1 = (V_1, V \setminus V_1)$ and $M_2 = (V_2, V \setminus V_2)$ and the edge $\{u, w\}$.*



Proof. Assume, on the contrary, that $M_1 = (V_1, V \setminus V_1)$ and $M_2 = (V_2, V \setminus V_2)$ are conflicting with respect to $\overrightarrow{H}$. That is, there exists an edge $\{u, w\} \in E$, that was not queried and hence is not oriented in $\overrightarrow{H}$, which is forced to have contrasting orientations by M_1 and M_2 (see figure 1). Without loss of generality, assume that $u \in V_1 \setminus V_2$, $w \in V_2 \setminus V_1$, and

$$|\overrightarrow{E_H}(V_1,\ V \setminus V_1)| \;=\; \frac{1}{2} \cdot |E(V_1,\ V \setminus V_1)|, \tag{1}$$

$$|\overrightarrow{E_H}(V_2,\ V \setminus V_2)| \;=\; \frac{1}{2} \cdot |E(V_2,\ V \setminus V_2)|. \tag{2}$$

Thus, M_1 forces e to be oriented from w to u, whereas M_2 forces e to be oriented from u to w. Summing Equation (1) with Equation (2) yields

$$|\overrightarrow{E_H}(V_1,\ V \setminus V_1)| + |\overrightarrow{E_H}(V_2,\ V \setminus V_2)| = \frac{1}{2}\left(|E(V_1,\ V \setminus V_1)| \;+\; |E(V_2,\ V \setminus V_2)|\right). \tag{3}$$

Recall now that all the cuts in G are valid with respect to $\overrightarrow{H}$. Consider the cuts $(V_1 \cap V_2,\ V \setminus (V_1 \cap V_2))$ and $(V_1 \cup V_2,\ V \setminus (V_1 \cup V_2))$. We have

$$|\overrightarrow{E_H}(V_1 \cap V_2,\ V \setminus (V_1 \cap V_2))| \;\le\; \frac{1}{2} \cdot |E(V_1 \cap V_2,\ V \setminus (V_1 \cap V_2))| \tag{4}$$

and

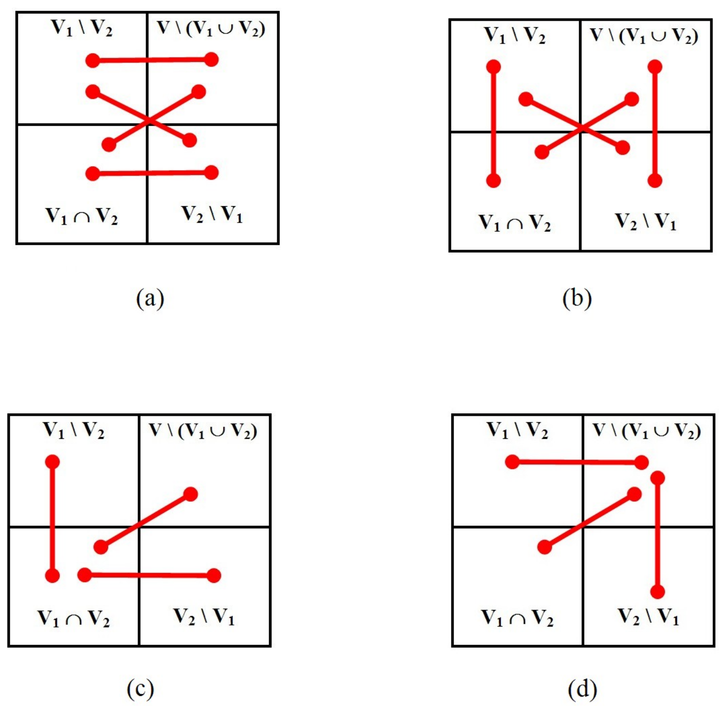
$$|\overrightarrow{E_H}(V_1 \cup V_2,\ V \setminus (V_1 \cup V_2))| \ \leq \ \frac{1}{2} \cdot |E(V_1 \cup V_2,\ V \setminus (V_1 \cup V_2))| \tag{5}$$

since these cuts are valid. Summing Inequality (4) with Inequality (5) yields

$$\begin{aligned} &\overrightarrow{E_H}(V_1 \cap V_2,\ V \setminus (V_1 \cap V_2))| \ + \ |\overrightarrow{E_H}(V_1 \cup V_2,\ V \setminus (V_1 \cup V_2))| \leq \\ &\frac{1}{2}\left(|E(V_1 \cap V_2,\ V \setminus (V_1 \cap V_2))| \ + \ |E(V_1 \cup V_2,\ V \setminus (V_1 \cup V_2))|\right). \end{aligned} \tag{6}$$

Now, note that

Figure 2: *The cuts of Equation (7) and the types of edges forming them. (a)* $(V_1,\ V \setminus V_1)$*, (b)* $(V_2,\ V \setminus V_2)$*, (c)* $(V_1 \cap V_2,\ V \setminus (V_1 \cap V_2))$*, (d)* $(V_1 \cup V_2,\ V \setminus (V_1 \cup V_2))$*.*



$$\begin{aligned} &|E(V_1 \cap V_2,\ V \setminus (V_1 \cap V_2))| \ + \ |E(V_1 \cup V_2,\ V \setminus (V_1 \cup V_2))| \\ &\quad = |E(V_1,\ V \setminus V_1)| + |E(V_2,\ V \setminus V_2)| - 2 \cdot |E(V_1 \setminus V_2,\ V_2 \setminus V_1)| \end{aligned} \tag{7}$$

(see Figure 2) and, similarly,

$$\begin{aligned} &|\overrightarrow{E_H}(V_1 \cap V_2,\ V \setminus (V_1 \cap V_2))| \ + \ |\overrightarrow{E_H}(V_1 \cup V_2,\ V \setminus (V_1 \cup V_2))| \\ &\quad = |\overrightarrow{E_H}(V_1,\ V \setminus V_1)| + |\overrightarrow{E_H}(V_2,\ V \setminus V_2)| - |\overrightarrow{E_H}(V_1 \setminus V_2,\ V_2 \setminus V_1)| - |\overrightarrow{E_H}(V_2 \setminus V_1,\ V_1 \setminus V_2)|. \end{aligned} \tag{8}$$

Substituting Equations (7) and (8) in Inequality (6) we obtain:

$$|\overrightarrow{E_H}(V_1,\ V\setminus V_1)| + |\overrightarrow{E_H}(V_2,\ V\setminus V_2)| - |\overrightarrow{E_H}(V_1\setminus V_2,\ V_2\setminus V_1)| - |\overrightarrow{E_H}(V_2\setminus V_1,\ V_1\setminus V_2)| \tag{9}$$
$$\le \frac{1}{2}\left(|E(V_1,\ V\setminus V_1)|\ +\ |E(V_2,\ V\setminus V_2)|\right) - |E(V_1\setminus V_2,\ V_2\setminus V_1)|.$$

Now, from Equation (3) we have:

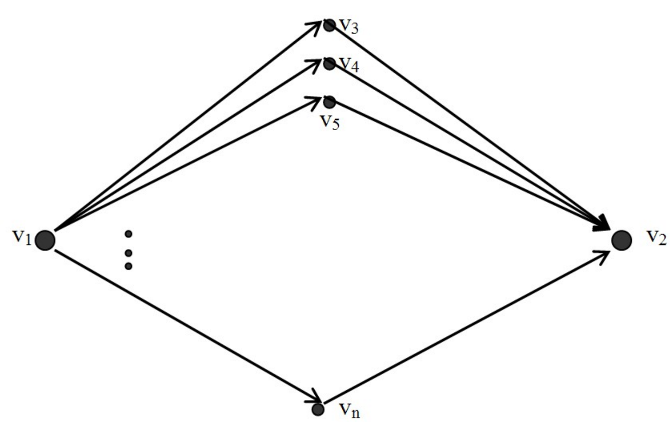
$$|\overrightarrow{E_H}(V_1\setminus V_2,\ V_2\setminus V_1)|\ +\ |\overrightarrow{E_H}(V_2\setminus V_1,\ V_1\setminus V_2)|\ \ge\ |E(V_1\setminus V_2,\ V_2\setminus V_1)|.$$

That is, all the edges in $E(V_1\setminus V_2,\ V_2\setminus V_1)$ are oriented in $\overrightarrow{H}$. This is a contradiction to our assumption that $\{u,w\}\in E(V_1\setminus V_2,\ V_2\setminus V_1)$ was not yet oriented. ■

Lemma 3.2 *Suppose that $\overrightarrow{H}$ is a knowledge graph that does not contain invalid cuts. Then $\overrightarrow{H}$ is extensible to an Eulerian orientation $\overrightarrow{G}=(V,\overrightarrow{E_G})$ of G. That is, $\overrightarrow{E_H}\subseteq\overrightarrow{E_G}$.*

Proof. We orient unoriented edges in the following manner. If there exists a restricting cut with unoriented edges, we orient one of them as obliged by the cut. According to Lemma 3.1, this will not invalidate any of the other cuts in the graph, and so we may continue. If there are no restricting cuts in the graph, we arbitrarily orient one unoriented edge and repeat (and this cannot violate any cut in the graph since there were no restricting cuts). Eventually, after orienting all the edges, we receive a complete orientation of G whose cuts are all valid, and thus it is Eulerian. ■

Figure 3: *The graph G_n. Every cut which separates v_1 and v_2 is of size $n-2$.*



Theorem 3.3 *There exists an infinite family of graphs for which every 1-sided test for being Eulerian must use $\Omega(m)$ queries.*

Proof. For every even n, let $G_n \stackrel{\text{def}}{=} K_{2,n-2}$, namely, the graph with a set of vertices $V = \{v_1, \ldots, v_n\}$ and a set of edges $E = \{\{v_i, v_j\} \mid i \in \{1,2\},\ j \in \{3, \ldots, n\}\}$ (see Figure 3). Clearly, G_n is Eulerian and $n = \Omega(m)$.

Consider the orientation $\overrightarrow{G}_n$ of G_n in which all the edges incident with v_1 are outgoing and all the edges incident with v_2 are incoming. Clearly, $\overrightarrow{G}_n$ is $\frac{1}{2}$-far from being Eulerian. According to Lemma 3.2, every 1-sided test must query at least half of the edges in some unbalanced cut (because otherwise it would clearly not obtain an invalid cut in the knowledge graph). However, one can easily see that every cut which does not separate v_1 and v_2 is balanced, while every cut which separates v_1 and v_2 is of size $n - 2 = \Omega(m)$. ∎

4 Generic tests

In this section we present one 1-sided p-test and two 2-sided p-tests for being Eulerian. Namely, our tests distinguish with high probability between the case where $\overrightarrow{G}$ is Eulerian and the case where $\overrightarrow{G}$ is p-far from being Eulerian (see Section 2.1). In later sections we devise several lower bounds on p for every orientation $\overrightarrow{G}$ that is ϵ-far from being Eulerian, thus obtaining corresponding upper bounds on the tests below. In fact, the 1-sided and 2-sided tests that we give in Subsection 4.2 are (p, β)-tests (see Section 2.2), which are, in particular, p-tests when $\beta = \Delta$. We will use these tests also for $\beta < \Delta$ in Section 7.

4.1 A 2-sided p-test

We give a simple 2-sided p-test that is independent of the maximum degree Δ. This p-test will yield efficient ϵ-tests for dense graphs (Section 5) and expanders (Section 6). To simplify notation, we denote $\delta \stackrel{\text{def}}{=} \frac{p}{4m}$.

Algorithm 4.1 *SIMPLE-2($\overrightarrow{G}, p$):*

1. *Repeat $\frac{4}{\delta}$ times independently:*
   - *Select an edge $e \in E$ uniformly and query it. Denote the start vertex of e in $\overrightarrow{E}$ by u and the end vertex of e in $\overrightarrow{E}$ by v.*
   - *Query $\frac{16 \ln(12/\delta)}{\delta^2}$ edges incident with u uniformly and independently and reject if the sample contains at least $(1+\delta)\frac{8\ln(12/\delta)}{\delta^2}$ outgoing edges.*
2. *Accept if the input was not rejected earlier.*

Lemma 4.2 *SIMPLE-2 is a 2-sided p-test for being Eulerian with query complexity $\widetilde{O}\left(\frac{1}{\delta^3}\right) = \widetilde{O}\left(\frac{m^3}{p^3}\right)$.*

Proof. The query complexity is clearly as stated. To prove the correctness of our algorithm, suppose first that $\overrightarrow{G}$ is Eulerian. For every vertex $u \in V$, the expected number of outgoing edges in the sample of u's incident edges is $\frac{8\ln(12/\delta)}{\delta^2}$. Applying Chernoff's upper tail bound, the probability of having at least $(1+\delta)\frac{8\ln(12/\delta)}{\delta^2}$ outgoing edges in a sample is at most $\exp\left(-\frac{\delta^2}{2}\frac{8\ln(12/\delta)}{\delta^2}\right) < \frac{\delta}{12}$. Since we sample $\frac{4}{\delta}$ vertices, the probability of rejecting an Eulerian orientation is at most $\frac{1}{3}$.

Assume now that $\overrightarrow{G}$ is p-far from being Eulerian. A vertex $u \in V$ will be called a δ-*spring* in $\overrightarrow{G}$ if $\mathrm{ib}(u) > 3\delta \cdot \deg(u)$. Let S_δ be the set of all δ-springs in $\overrightarrow{G}$. Note that $S_\delta \subseteq S$, where S is the set of all springs in $\overrightarrow{G}$. We have

$$\sum_{u \in S} \mathrm{ib}(u) \le \sum_{u \in S_\delta} \deg(u) + 3\delta \cdot \sum_{u \in S \setminus S_\delta} \deg(u).$$

Let $y = \sum_{u \in S_\delta} \deg(u)$. Since $\sum_{u \in S} \deg(u) < 2m$, we obtain

$$\sum_{u \in S} \mathrm{ib}(u) < y + 3\delta \cdot (2m - y) = (1 - 3\delta)y + 6\delta m.$$

However, from Lemma 2.7, we have $\sum_{u \in S} \mathrm{ib}(u) = 2p = 8\delta m$, and therefore,

$$y > \frac{2\delta m}{1 - 3\delta} > 2\delta m.$$

Thus, $\sum_{u \in S_\delta} \deg(u) > 2\delta m$, and since for every $u \in S_\delta$ we have $\deg_{out}(u) > \frac{1}{2}\deg(u)$, there exist at least δm edges $(u, v) \in \overrightarrow{E}$ such that u is a δ-spring. As SIMPLE-2 samples $\frac{4}{\delta}$ edges $(u, v) \in \overrightarrow{E}$ uniformly and independently, the probability of not sampling any edge (u, v) such that u is a δ-spring is at most $(1 - \delta)^{\frac{4}{\delta}} < e^{-4} < 0.02$.

Suppose now that the algorithm has sampled an edge (u, v) such that u is a δ-spring. Then $\overrightarrow{G}$ will be rejected unless fewer than $(1 + \delta)\frac{8\ln(12/\delta)}{\delta^2}$ of the queried edges (u, w) are outgoing. Since $\mathrm{ib}(u) > 3\delta \cdot \deg(u)$, the expected number of outgoing edges is at least $(1 + 3\delta/2)\frac{8\ln(12/\delta)}{\delta^2}$. Note that $(1 + \delta)\frac{8\ln(12/\delta)}{\delta^2} < (1 - \frac{1}{4}\delta)(1 + \frac{3}{2}\delta)\frac{8\ln(12/\delta)}{\delta^2}$. Therefore, by the Chernoff bound, the probability of sampling fewer than $(1 + \delta)\frac{8\ln(12/\delta)}{\delta^2}$ outgoing edges is at most

$$\exp\left(-\frac{\delta^2}{32}\left(1 + \frac{3\delta}{2}\right)\frac{8\ln(12/\delta)}{\delta^2}\right) < \exp\left(-\frac{\ln(12/\delta)}{4}\right).$$

Note that $\frac{\ln(12/\delta)}{4} > \ln(10/3)$, and therefore, the probability of not detecting that u is a δ-spring is at most 0.3. We thus conclude that if $\overrightarrow{G}$ is p-far from being Eulerian, then SIMPLE-2 accepts $\overrightarrow{G}$ with probability smaller than $\frac{1}{3}$, which completes the proof of the theorem. ■

4.2 (p, β)-tests

We now give a simple 1-sided (p, β)-test, which has a better query complexity than SIMPLE-2 for $\Delta \ll \frac{m^2}{p^2}\ln(\frac{m}{p})$.

Algorithm 4.3 *GENERIC-1*$(\overrightarrow{G}, p, \beta)$*:*

1. *Repeat* $\frac{\ln 3\ m}{p}$ *times independently:*

   - *Select an edge* $e \in E$ *uniformly and query it. Denote the start vertex of* e *in* $\overrightarrow{E}$ *by* u *and the end vertex of* e *in* $\overrightarrow{E}$ *by* v*.*

- *If* $\deg(u) \leq \beta$ *then: Query all the edges* $\{u, w\} \in E$ *and reject if* u *is unbalanced, namely, if* $\mathrm{ib}(u) \neq 0$*.*

2. *Repeat* $\frac{\ln 3 \ m}{p}$ *times independently:*

   - *Select an edge* $e \in E$ *uniformly and query it. Denote the start vertex of* e *in* $\overrightarrow{E}$ *by* u *and the end vertex of* e *in* $\overrightarrow{E}$ *by* v*.*
   - *If* $\deg(v) \leq \beta$ *then: Query all the edges* $\{w, v\} \in E$ *and reject if* v *is unbalanced, namely, if* $\mathrm{ib}(v) \neq 0$*.*

3. *Accept if the input was not rejected by the above.*

Lemma 4.4 *GENERIC-1 is a 1-sided* (p, β)*-test for being Eulerian with query complexity* $O\left(\frac{\beta m}{p}\right)$*. In particular, for* $\beta = \Delta$*, GENERIC-1 is a 1-sided* p*-test with query complexity* $O\left(\frac{\Delta m}{p}\right)$*.*

Proof. The query complexity is clearly as stated. Since Algorithm 4.3 rejects only when an unbalanced β-small vertex is discovered, it always accepts a β-Eulerian $\overrightarrow{G}$. Suppose now that $\overrightarrow{G}$ is p-far from being β-Eulerian. Clearly, to reject $\overrightarrow{G}$, it suffices to sample one edge $e = (u, v) \in \overrightarrow{E}$ such that u is a spring in Step 1 or an edge $e = (u, v) \in \overrightarrow{E}$ such that v is a drain in Step 2. By Lemma 2.9, E contains at least p edges of at least one of these two kinds of edges, and so their fraction is at least p/m. We thus conclude that the probability of accepting $\overrightarrow{G}$ is at most

$$\left(1 - \frac{p}{m}\right)^{\frac{\ln 3 \ m}{p}} < \frac{1}{3}.$$

■

We conclude this section with a 2-sided (p, β)-test, which gives better query complexity than GENERIC-1 for $\beta \gg \log^2 m$ and better query complexity than SIMPLE-2 for $p \ll \frac{m}{\sqrt{\beta}}$. The main idea of the algorithm is to perform roughly $O((\log \beta)^2)$ testing stages, each designed to detect unbalanced vertices whose degree and imbalance lie in a certain interval. We show that with high probability, a β-Eulerian orientation $\overrightarrow{G}$ is accepted by all the testing stages, while an orientation that is p-far from being β-Eulerian is rejected by at least one of them.

Algorithm 4.5 *MULTISTAGE-2*$(\overrightarrow{G}, p, \beta)$*:*

For $i = 1, \ldots, \lceil \log \beta \rceil - 1$*, do:*

1. *Let* $V_i \stackrel{\text{def}}{=} \{u \in V \mid \deg(u) \in [2^i, 2^{i+1})\ \}$ *and* $n_i \stackrel{\text{def}}{=} |V_i|$*.*
2. *Let* $j = \lceil i/2 \rceil$*. If* $2^j \cdot n_i > \frac{2p}{(\log \beta)^2}$ *then:*
   - *Sample* $x_{ij} = \frac{\ln 12 \ (\log \beta)^2 \ 2^{j+1} \ n_i}{2p}$ *vertices in* V_i *uniformly and independently.*
   - *For every sampled vertex* u*, query all the edges incident with* u*, and reject if* u *is unbalanced.*
3. *For every* $j \in \{\lceil i/2 \rceil + 1, \ldots, i - 1\}$ *such that* $2^j \cdot n_i > \frac{2p}{(\log \beta)^2}$ *do:*
   - *Sample* $x_{ij} = \frac{\ln 12 \ (\log \beta)^2 \ 2^{j+1} \ n_i}{2p}$ *vertices in* V_i *uniformly and independently.*

- *For every sampled vertex u, query $q_{ij} = 256 \cdot \ln(\ 6(\log\beta)^2\ x_{ij}) \cdot 2^{2(i-j)}$ edges adjacent to u, uniformly and independently, and reject if the absolute difference between the number of incoming and outgoing edges in the sample is at least $\frac{q_{ij}}{4\cdot 2^{i-j}}$.*

Accept if the input was not rejected earlier.

Lemma 4.6 *MULTISTAGE-2 is a 2-sided (p,β)-test for being Eulerian with query complexity $\widetilde{O}\left(\frac{\sqrt{\beta}\ m}{p}\right)$. In particular, for $\beta = \Delta$, MULTISTAGE-2 is a 2-sided p-test for being Eulerian with query complexity $\widetilde{O}\left(\frac{\sqrt{\Delta}\ m}{p}\right)$.*

Proof. First, we compute the asymptotic query complexity of Algorithm 4.5. Since

$$x_{ij} = O\left(\frac{(\log\beta)^2 \cdot 2^j \cdot n_i}{p}\right)$$

and

$$q_{ij} = O\left(\log\left(\frac{\log\beta \cdot 2^j \cdot n_i}{p}\right) \cdot 2^{2(i-j)}\right) = O\left(\log\left(\frac{\log\beta \cdot m}{p}\right) \cdot 2^{2(i-j)}\right),$$

the total query complexity is at most

$$\begin{aligned}
&\sum_{i=1}^{\log\beta}\left(x_{i\lceil i/2\rceil}2^{\lceil i/2\rceil+1} + \sum_{j=\lceil i/2\rceil+1}^{i-1} x_{ij}\cdot q_{ij}\right) \\
&= O\left(\log\left(\frac{\log\beta\cdot m}{p}\right)\frac{(\log\beta)^2}{p}\sum_{i=1}^{\log\beta} 2^{2i}\cdot n_i \sum_{j=\lceil i/2\rceil}^{i-1}\frac{1}{2^j}\right) \\
&= O\left(\log\left(\frac{\log\beta\cdot m}{p}\right)\frac{(\log\beta)^2}{p}\sum_{i=1}^{\log\beta} 2^{3i/2}\cdot n_i\right) \\
&= O\left(\log\left(\frac{\log\beta\cdot m}{p}\right)\frac{(\log\beta)^2}{p}\,2^{\log\beta/2}\sum_{i=1}^{\log\beta} 2^{i}\cdot n_i\right) \\
&= O\left(\log\left(\frac{\log\beta\cdot m}{p}\right)\frac{(\log\beta)^2}{p}\sqrt{\beta}m\right).
\end{aligned}$$

Now suppose that $\overrightarrow{G}$ is β-Eulerian. Then $\overrightarrow{G}$ can only be rejected in Step 3, where we randomly sample q_{ij} edges incident with a (β-small) vertex $u \in V_i$. Since u is balanced, the expected number of incoming edges in the sample is $\frac{q_{ij}}{2}$. By the Chernoff bound, the probability of sampling fewer than $(1-\frac{1}{4\cdot 2^{i-j}})\frac{q_{ij}}{2}$ incoming edges is at most $\exp\left(-\frac{1}{2^{2(i-j)}}\cdot\frac{q_{ij}}{64}\right) < \frac{1}{6(\log\beta)^2 x_{ij}}$. Similarly, the probability of sampling fewer than $(1-\frac{1}{4\cdot 2^{i-j}})\frac{q_{ij}}{2}$ outgoing edges is at most $\frac{1}{6(\log\beta)^2 x_{ij}}$. Since for every pair (i,j) we sample x_{ij} vertices, and since there are less than $(\log\beta)^2$ pairs, the total probability of rejecting $\overrightarrow{G}$ is at most $\frac{1}{3}$.

Suppose now that $\overrightarrow{G}$ is p-far from being β-Eulerian. From Lemma 2.9, we have

$$\sum_{u \in V, \deg(u) \le \beta} |\mathrm{ib}(u)| > 2p.$$

We define a partition of the unbalanced β-small vertices in $\overrightarrow{G}$ as follows. For every $1 \le i \le \lceil \log \beta \rceil - 1$ and for every $1 \le j \le i$, let

$$V_{ij} \stackrel{\text{def}}{=} \{ u \in V_i \mid |\mathrm{ib}(u)| \in [2^j, 2^{j+1}) \}.$$

Then there exist i_0 and j_0 such that

$$\sum_{u \in V_{i_0 j_0}} |\mathrm{ib}(u)| > \frac{2p}{(\log \beta)^2}. \tag{10}$$

As $|\mathrm{ib}(u)| < 2^{j_0+1}$ for every $u \in V_{i_0 j_0}$, we have

$$|V_{i_0 j_0}| > \frac{2p}{(\log \beta)^2 \cdot 2^{j_0+1}}.$$

Recall that $|V_{i_0}| = n_{i_0}$. Hence, when we sample $x_{i_0 j_0}$ vertices in Step 2 or in Step 3, the probability of not sampling any vertex $u \in V_{i_0 j_0}$ is smaller than

$$\left(1 - \frac{2p}{(\log \beta)^2 \cdot 2^{j_0+1} \cdot n_{i_0}}\right)^{x_{i_0 j_0}} < \exp\left(-\frac{2p}{(\log \beta)^2 \cdot 2^{j_0+1} \cdot n_{i_0}} \cdot x_{i_0 j_0}\right) = \frac{1}{12}.$$

Assume from now on that the algorithm has sampled a vertex $u \in V_{i_0 j_0}$. If $j_0 \le \lceil i_0/2 \rceil$ then all the edges incident with u are queried (Step 2), and since u is unbalanced, $\overrightarrow{G}$ is now rejected with probability 1. Otherwise, $q_{i_0 j_0}$ edges incident with u are queried independently in Step 3. Since $\deg(u) < 2^{i_0+1}$ and $|\mathrm{ib}(u)| \ge 2^{j_0}$, either the expected number of incoming edges in the sample is at least $\left(1 + \frac{1}{2 \cdot 2^{i_0 - j_0}}\right) \frac{q_{i_0 j_0}}{2}$ or the expected number of outgoing edges in the sample is at least $\left(1 + \frac{1}{2 \cdot 2^{i_0 - j_0}}\right) \frac{q_{i_0 j_0}}{2}$. To accept the input, we must sample fewer than $\left(1 + \frac{1}{4 \cdot 2^{i_0 - j_0}}\right) \frac{q_{i_0 j_0}}{2} < \left(1 - \frac{1}{8 \cdot 2^{i_0 - j_0}}\right) \cdot \left(1 + \frac{1}{2 \cdot 2^{i_0 - j_0}}\right) \frac{q_{i_0 j_0}}{2}$ edges in the majority direction. By the Chernoff bound, the probability of doing so is at most

$$\exp\left(-\frac{q_{i_0 j_0}}{256 \cdot 2^{2(i_0 - j_0)}}\right) = \exp\left(-\ln(6(\log \beta)^2 \; x_{i_0 j_0})\right) = \frac{1}{6(\log \beta)^2 \; x_{i_0 j_0}}.$$

Note that from Inequality (10) we must have $2^{j_0+1} \cdot n_{i_0} > \frac{4p}{(\log \beta)^2}$ and hence $x_{i_0 j_0} > \ln 12$. In addition, we may assume that $\beta \ge 2$, since otherwise $\overrightarrow{G}$ is trivially β-Eulerian. It thus follows that the probability of not rejecting the input when sampling u's edges is smaller than $1/4$. We conclude that the probability of accepting an orientation that is p-far from being β-Eulerian is at most $1/3$. ■

We note that the query complexity bound of MULTISTAGE-2 can be made tighter for some graphs, as the algorithm skips pairs (i, j) where $2^j \cdot n_i \le \frac{2p}{(\log \beta)^2}$ (and thus i and j cannot be the i_0 and j_0 which satisfy Equation (10)). In particular, for regular graphs, only one value of i is relevant for testing, which eliminates the square over $\log \beta$ in the query complexity. However, this does not change the power of β or the dependency on p and m.

5 Testing graphs with high average degree

In this section we obtain a general lower bound for the required number p of correction paths as a function of $d = m/n$. This, together with the analysis of the previous section, will provide us with an efficient test for dense graphs.

Lemma 5.1 *Suppose that $\overrightarrow{G}$ is not Eulerian and that $\overrightarrow{H}$ is an acyclic correction subgraph of $\overrightarrow{G}$ which is a union of p edge-disjoint spring-drain paths. Then $\overrightarrow{H}$ contains a spring-drain path of length smaller than $\frac{n}{\sqrt{p}}$.*

Proof. Consider a Breadth-First Search (BFS) traversal of $\overrightarrow{H}$ starting at the set of springs, S. We define a partition of V_H into levels $L_0, \ldots, L_t$ for some $t > 0$ as follows. Let $L_0 = S$. Now, for any $i > 0$, while $\bigcup_{j=0}^{i-1} L_j \neq V_H$, define

$$L_i \stackrel{\text{def}}{=} \{v \in V_H \setminus \bigcup_{j=0}^{i-1} L_j \mid \text{there exists } u \in L_{i-1} \text{ such that } (u, v) \in \overrightarrow{E_H}\}.$$

Note that if $v \in L_i$ then $\overrightarrow{H}$ contains a path of length i from some spring to v. Let ℓ be the minimum index of a level L_i which contains a drain. We prove the claim by showing that $\ell < \frac{n}{\sqrt{p}}$.

For every $i = 0, \ldots, t$ we let $n_i = |L_i|$. Recall that there are p edge-disjoint spring-drain paths in $\overrightarrow{H}$. We first show that for every $0 \leq i < \ell$ we have $n_{i+1} \geq p/n_i$. Consider the level L_i for some $0 \leq i < \ell$. Since there are no drains in the levels $L_0, \ldots, L_i$, there exist at least p edges from L_i to L_{i+1}. Therefore, there exists a vertex $v \in L_i$ which has at least p/n_i neighbors in L_{i+1}. Hence, $n_{i+1} \geq p/n_i$.

Summing over all $i = 0, \ldots, \ell - 1$ we obtain

$$\sum_{i=0}^{\ell-1} n_{i+1} \geq p \cdot \sum_{i=0}^{\ell-1} \frac{1}{n_i},$$

and so

$$p \leq \frac{\sum_{i=0}^{\ell-1} n_{i+1}}{\sum_{i=0}^{\ell-1} \frac{1}{n_i}}.$$

Now, for a given $\sum_{i=0}^{\ell-1} n_i$, the minimum value of $\sum_{i=0}^{\ell-1} \frac{1}{n_i}$ is reached when $n_i = \sum_{i=0}^{\ell-1} n_i/\ell$ for every $i = 0, \ldots, \ell - 1$. Thus,

$$p \leq \frac{\sum_{i=0}^{\ell-1} n_{i+1}}{\ell^2 / \sum_{i=0}^{\ell-1} n_i} < n^2/\ell^2,$$

which proves the lemma. ■

Lemma 5.2 *If $\overrightarrow{G}$ is ϵ-far from being Eulerian then it is p-far from being Eulerian for $p > \epsilon^2 d^2/4$.*

Proof. Let $\overrightarrow{H_0}$ be an acyclic correction subgraph of $\overrightarrow{G}$. In each step $j \geq 1$, while $\overrightarrow{H_{j-1}}$ is not empty, we choose a shortest spring-drain path $\overrightarrow{P_{j-1}}$ in $\overrightarrow{H_{j-1}}$ and set $\overrightarrow{H_j} = \overrightarrow{H_{j-1}} \setminus \overrightarrow{P_{j-1}}$. By Lemma

2.6, $\overrightarrow{H_0}$ is a union of $p = \frac{1}{4}\sum_{u\in V}|\text{ib}(u)|$ edge-disjoint spring-drain paths, and moreover, every $\overrightarrow{H_j}$ is a union of $p-j$ disjoint spring-drain paths. Hence, the graphs $\overrightarrow{H_0}, \ldots, \overrightarrow{H_{p-1}}$ are non-empty. Furthermore, every subgraph $\overrightarrow{H_j}$ is clearly acyclic, and hence by Observation 2.5, for $j = 0, \ldots, p-1$ $\overrightarrow{H_j}$ is a correction subgraph for some non-Eulerian orientation of G. Let ℓ_j be the length of $\overrightarrow{P_j}$ for $j = 0, \ldots, p-1$. Then by Lemma 5.1, we have $\ell_j < \frac{n}{\sqrt{p-j}}$. Summing over j, we obtain

$$\sum_{j=0}^{p-1} \ell_j < n \cdot \sum_{j=0}^{p-1} \frac{1}{\sqrt{p-j}} = n \sum_{j=1}^{p} \frac{1}{\sqrt{j}} = n \left(1 + \sum_{j=2}^{p} \frac{1}{\sqrt{j}}\right). \tag{11}$$

Since $f(x) = \frac{1}{\sqrt{x}}$ is monotone decreasing for every $x > 0$, we have

$$\frac{1}{\sqrt{j}} < \int_{x=j-1}^{j} \frac{dx}{\sqrt{x}}$$

for every $j \geq 1$, and therefore

$$\sum_{j=2}^{p} \frac{1}{\sqrt{j}} < \int_{x=1}^{p} \frac{dx}{\sqrt{x}} = 2\sqrt{p} - 2.$$

Substituting this in (11) we obtain

$$\sum_{j=0}^{p-1} \ell_j < 2\sqrt{p}n.$$

Note that $\sum_{j=0}^{p-1} \ell_j$ is the total number of edges in $\overrightarrow{H}$. As $\overrightarrow{G}$ is ϵ-far from being Eulerian, we have $\epsilon m < \sum_{j=0}^{p-1} \ell_j$, and thus, $p > \epsilon^2 m^2/4n^2 = \epsilon^2 d^2/4$. ■

Substituting the lower bound for p of Lemma 5.2 in Lemmas 4.4, 4.2 and 4.6, we obtain the following theorem.

Theorem 5.3

1. *SIMPLE-2($\overrightarrow{G}, \epsilon^2 d^2/4$) is a 2-sided ϵ-test for being Eulerian with query complexity* $\widetilde{O}\left(\frac{m^3}{\epsilon^6 d^6}\right) = \widetilde{O}\left(\frac{n^3}{\epsilon^6 d^3}\right)$.

2. *GENERIC-1($\overrightarrow{G}, \epsilon^2 d^2/4, \Delta$) is a 1-sided ϵ-test for being Eulerian with query complexity* $O\left(\frac{\Delta m}{\epsilon^2 d^2}\right) = O\left(\frac{\Delta n}{\epsilon^2 d}\right)$.

3. *MULTISTAGE-2($\overrightarrow{G}, \epsilon^2 d^2/4, \Delta$) is a 2-sided ϵ-test for being Eulerian with query complexity* $\widetilde{O}\left(\frac{\sqrt{\Delta} m}{\epsilon^2 d^2}\right) = \widetilde{O}\left(\frac{\sqrt{\Delta} n}{\epsilon^2 d}\right)$.

SIMPLE-2 gives a sub-linear complexity for $d = \omega\left(\frac{\sqrt{n}\cdot(\log n)^{1/4}}{\epsilon^{3/2}}\right)$, GENERIC-1 gives a sub-linear query complexity for $d = \omega\left(\frac{\sqrt{\Delta}}{\epsilon}\right)$, and MULTISTAGE-2 gives a sub-linear complexity for $d = \omega\left(\frac{\Delta^{1/4}}{\epsilon}\right)$. All of the tests yield their lowest query complexity relative to m when $m = \Theta(n^2)$ (i.e., $d = \Theta(n)$): $\widetilde{O}(1/\epsilon^6)$ for SIMPLE-2, $O(n/\epsilon^2)$ for GENERIC-1, and $\widetilde{O}(\sqrt{n}/\epsilon^2)$ for MULTISTAGE-2.

6 Testing orientations of an expander graph

In this section we obtain a lower bound for the required number p of correction paths in an expander graph. This bound, together with the analysis of Section 4 will provide us with ϵ-tests for expanders.

A graph $G = (V, E)$ is called an *α-expander* for some $\alpha > 0$ if it is connected and for every $U \subseteq V$ such that $0 < |E(U)| \leq m/2$ we have

$$\frac{|\partial U|}{|E(U)|} \geq \alpha.$$

Note that while the diameter of G is $O(\log_{(1+\alpha)} m)$, the "oriented-diameter" of $\overrightarrow{G}$ is not necessarily low, even if we assume that the orientation is Eulerian, as was shown by [2].

In the following, $\log_b^{(k)}(x)$ denotes the k-nested logarithm with base b of x, that is, $\log_b^{(1)}(x) \stackrel{\text{def}}{=} \log_b(x)$ and $\log_b^{(k+1)}(x) \stackrel{\text{def}}{=} \log_b(\log_b^{(k)}(x))$ for any natural $k \geq 1$.

Lemma 6.1 *Let G be an Eulerian α-expander and let $k \geq 1$ be a natural number such that $\log^{(k-1)}_{(1+\alpha/2)} m \geq \log_{(1+\alpha/2)} \left(\frac{4}{\epsilon}\right)$. Then:*

1. *Every non-Eulerian orientation $\overrightarrow{G}$ of G contains a spring-drain path of length at most*

$$\ell_k \stackrel{\text{def}}{=} 2 \cdot \log^{(k)}_{(1+\alpha/2)} m + 2 \cdot \log_{(1+\alpha/2)} \left(\frac{4}{\epsilon}\right). \tag{12}$$

2. *Every orientation $\overrightarrow{G}$ of G that is ϵ-far from being Eulerian is p_k-far from being Eulerian for*

$$p_k \stackrel{\text{def}}{=} \frac{\epsilon m}{\ell_k} = \frac{\epsilon m}{2 \cdot \log^{(k)}_{(1+\alpha/2)} m + 2 \cdot \log_{(1+\alpha/2)} \left(\frac{4}{\epsilon}\right)}$$

Proof. We prove the lemma by induction on k. In every inductive step, we use the known bounds of ℓ_k and p_k to devise ℓ_{k+1} and p_{k+1} in an iterative manner. We start by proving the lemma for the base case, $k = 1$.

To prove Item 1 of the lemma for $k = 1$, let $\overrightarrow{G}$ be a non-Eulerian orientation of $\overrightarrow{G}$. Consider a BFS traversal of $\overrightarrow{G}$ starting from the set S of springs. For every $i \geq 0$, let L_i be the ith level of the traversal, where $L_0 = S$, and let $U_{<i} \stackrel{\text{def}}{=} \bigcup_{0 \leq j < i} L_j$ and $U_{\geq i} \stackrel{\text{def}}{=} \bigcup_{j \geq i} L_j$. For every $i > 0$, let f_i be the number of directed edges going from L_{i-1} to L_i. Let L_ℓ be the first level that contains a drain. By the expander property of G, for every $i > 0$ while $|E(U_{<i})| \leq m/2$ we have $|\partial(U_{<i})| \geq \alpha |E(U_{<i})|$. Note that for every $i \leq \ell$, the set $U_{<i}$ contains no drains, and all the directed edges that exit it are from L_{i-1} to L_i. Hence $f_i > \frac{1}{2}|\partial(U_{<i})|$. We thus obtain that

$$f_i > \frac{\alpha}{2} |E(U_{<i})|$$

for every $0 < i \leq \ell$ while $|E(U_{<i})| \leq m/2$, and therefore

$$|E(U_{<i+1})| > \left(1 + \frac{\alpha}{2}\right) |E(U_{<i})|.$$

By induction, we obtain that

$$|E(U_{<i})| > \left(1+\frac{\alpha}{2}\right)^{i-1} f_1 \geq \left(1+\frac{\alpha}{2}\right)^{i-1} \tag{13}$$

for every $0 < i \leq \ell$ for which $|E(U_{<i})| \leq m/2$. Now, if for every $0 < i \leq \ell$ we have $|E(U_{<i})| \leq m/2$, then clearly $|E(U_{<\ell})| > \left(1+\frac{\alpha}{2}\right)^{\ell-1}$, and hence $\ell - 1 < \log_{(1+\alpha/2)} |E(U_{<\ell})| \leq \log_{(1+\alpha/2)} m$ and $\ell < \log_{(1+\alpha/2)} m$.

Otherwise, let $r > 0$ be the minimal index for which $|E(U_{<r})| > m/2$. Then, for every $r \leq i \leq \ell$ we have $|E(U_{\geq i})| < m/2$, and therefore $|\partial(U_{\geq i})| \geq \alpha |E(U_{\geq i})|$. Note that for every $i \geq r$, the set $U_{\geq i}$ contains no springs, and all the directed edges that enter it go from L_{i-1} to L_i. Therefore, $f_i > \frac{1}{2}|\partial(U_{\geq i})|$. We obtain that for every $r \leq i \leq \ell$

$$f_i > \frac{\alpha}{2}|E(U_{\geq i})|,$$

and thus

$$|E(U_{\geq i-1})| > \left(1+\frac{\alpha}{2}\right)|E(U_{\geq i})|.$$

By induction, we have

$$|E(U_{\geq i-1})| > \left(1+\frac{\alpha}{2}\right)^{\ell-i+1} |E(U_{\geq \ell})| \geq \left(1+\frac{\alpha}{2}\right)^{\ell-i+1} \tag{14}$$

for every $r \leq i \leq \ell$.

From (13) and (14) we obtain that both $r - 1 < \log_{(1+\alpha/2)} m$ and $\ell - r + 1 < \log_{(1+\alpha/2)} m$, and therefore $\ell < 2 \cdot \log_{(1+\alpha/2)} m$. Hence, every non-Eulerian orientation of G contains a spring-drain path of length at most $\ell_1 = 2 \cdot \log_{(1+\alpha/2)} m + 2 \cdot \log_{(1+\alpha/2)} \left(\frac{4}{\epsilon}\right)$.

To prove Item 2 of the lemma for $k = 1$, let $\overrightarrow{G}$ be an orientation of G that is ϵ-far from being Eulerian. While $\overrightarrow{G}$ is not Eulerian, choose a shortest spring-drain path in $\overrightarrow{G}$ and invert all its edges. By Item 1, every chosen spring-drain path is of length at most ℓ_1. Let $\overrightarrow{H}$ be the union of the spring-drain paths inverted. Clearly, $\overrightarrow{H}$ is a correction subgraph of $\overrightarrow{G}$. As $\overrightarrow{G}$ is ϵ-far from being Eulerian, $\overrightarrow{H}$ contains at least ϵm edges, and thus it is necessarily a union of at least $p_1 = \frac{\epsilon m}{\ell_1}$ disjoint spring-drain paths. By Lemma 2.6, every correction subgraph of $\overrightarrow{G}$ contains the same number of disjoint spring-drain paths, which completes the proof of the base case.

Suppose now that the lemma holds for some natural $k \geq 1$ and assume that $\log^{(k)}_{(1+\alpha/2)} m \geq \log_{(1+\alpha/2)} \left(\frac{4}{\epsilon}\right)$. The proof of both items of the lemma for $k+1$ is very similar to that of the base case. However, in Inequality (13) we know that $f_1 \geq p_k$, and in Inequality (14) we know that $|E(U_{\geq \ell})| \geq p_k$. Hence, every non-Eulerian orientation $\overrightarrow{G}$ of G contains a spring-drain path of length at most

$$\begin{aligned}\ell_{k+1} &\leq 2 \cdot \log_{(1+\alpha/2)} \left(\frac{m}{p_k}\right) = 2 \cdot \log_{(1+\alpha/2)} \left(\frac{\ell_k}{\epsilon}\right) \\ &\leq 2 \cdot \log_{(1+\alpha/2)} \left(2 \cdot \log^{(k)}_{(1+\alpha/2)} m + 2 \cdot \log_{(1+\alpha/2)} \left(\frac{4}{\epsilon}\right)\right) + 2 \cdot \log_{(1+\alpha/2)} \left(\frac{1}{\epsilon}\right). \end{aligned} \tag{15}$$

Since $\log^{(k)}_{(1+\alpha/2)} m \geq \log_{(1+\alpha/2)} \left(\frac{4}{\epsilon}\right)$ we have

$$\ell_{k+1} \leq 2 \cdot \log_{(1+\alpha/2)} \left(4 \cdot \log^{(k)}_{(1+\alpha/2)} m\right) + 2 \cdot \log_{(1+\alpha/2)} \left(\frac{1}{\epsilon}\right) = 2 \cdot \log^{(k+1)}_{(1+\alpha/2)} m + 2 \cdot \log_{(1+\alpha/2)} \left(\frac{4}{\epsilon}\right),$$

which proves Item 1. The proof of Item 2 is the same as for the base case. ■

Lemma 6.2 *Let G be an Eulerian α-expander. Let $\overrightarrow{G}$ be an orientation of G that is ϵ-far from being Eulerian. Then $\overrightarrow{G}$ is p-far from being Eulerian for*

$$p = \Omega\left(\frac{\alpha \epsilon m}{\log(\frac{1}{\epsilon})}\right).$$

Proof. Let k be the minimum natural number such that $\log^{(k)}_{(1+\alpha/2)} m < \log_{(1+\alpha/2)}\left(\frac{4}{\epsilon}\right)$. Then, using the same arguments as we did in the proof of Lemma 6.1 for smaller k's, we obtain that every non-Eulerian orientation of G contains a spring-drain path of length at most ℓ_{k+1}, where ℓ_{k+1} satisfies Inequality (15). However, since $\log^{(k)}_{(1+\alpha/2)} m < \log_{(1+\alpha/2)}\left(\frac{4}{\epsilon}\right)$, we now have

$$\ell_{k+1} < 2 \cdot \log_{(1+\alpha/2)}\left(4 \cdot \log_{(1+\alpha/2)}\left(\frac{4}{\epsilon}\right)\right) + 2 \cdot \log_{(1+\alpha/2)}\left(\frac{1}{\epsilon}\right) = O\left(\log_{(1+\alpha/2)}\left(\frac{1}{\epsilon}\right)\right).$$

Similarly to our proof of Item 2 in Lemma 6.1, we obtain that every orientation of G that is ϵ-far from being Eulerian is p-far from being Eulerian for

$$p = \frac{\epsilon m}{\ell_{k+1}} = \Omega\left(\frac{\epsilon m}{\log_{(1+\alpha/2)}\left(\frac{1}{\epsilon}\right)}\right) = \Omega\left(\frac{\alpha \epsilon m}{\log(\frac{1}{\epsilon})}\right).$$

■

Substituting the lower bound for p of Lemma 6.2 in Lemmas 4.2, 4.4 and 4.6, we obtain the following theorem.

Theorem 6.3 *Let G be an α-expander (for some $\alpha > 0$) with m edges and maximum degree Δ.*

1. *SIMPLE-2$\left(\overrightarrow{G}, \Omega\left(\frac{\alpha\epsilon m}{\log(1/\epsilon)}\right)\right)$ is a 2-sided ϵ-test for being Eulerian with query complexity $\widetilde{O}\left(\left(\frac{\log(1/\epsilon)}{\alpha\epsilon}\right)^3\right)$.*

2. *GENERIC-1$\left(\overrightarrow{G}, \Omega\left(\frac{\alpha\epsilon m}{\log(1/\epsilon)}\right), \Delta\right)$ is a 1-sided ϵ-test for being Eulerian with query complexity $O\left(\frac{\Delta\log(1/\epsilon)}{\alpha\epsilon}\right)$.*

3. *MULTISTAGE-2$\left(\overrightarrow{G}, \Omega\left(\frac{\alpha\epsilon m}{\log(1/\epsilon)}\right), \Delta\right)$ is a 2-sided ϵ-test for being Eulerian with query complexity $\widetilde{O}\left(\frac{\sqrt{\Delta}\log(1/\epsilon)}{\alpha\epsilon}\right)$.*

Note that for a constant α, the query complexity of SIMPLE-2 depends only on ϵ (while the other tests depend also on Δ).

7 Testing orientations of "lame" directed expanders

In this section we discuss a variation of the expander test, which will serve us in Section 8 for devising tests for general graphs. Given an orientation $\overrightarrow{G}$ of G, we now test a subgraph $\overrightarrow{G}[U]$ of $\overrightarrow{G}$, induced by a subset $U \subseteq V$. We refer to the edges in $E(U)$ as the *internal edges* of $\overrightarrow{G}[U]$, and denote $m_U \stackrel{\text{def}}{=} |E(U)|$. We say that $\overrightarrow{G}[U]$ is *Eulerian* if and only if all the vertices in U are balanced in $\overrightarrow{G}$. We say that $\overrightarrow{G}[U]$ is *β-Eulerian* if and only if all the β-small vertices in U are balanced in $\overrightarrow{G}$. Note that these definitions rely also on the edges in ∂U, which we will henceforth call *external edges*. We assume that the orientations of all the external edges are known, and furthermore, we use a distance function that does not allow inverting external edges. Namely, we will say that $\overrightarrow{G}[U]$ is *ϵ-close* to being Eulerian if and only if it has a correction subgraph of size at most ϵm_U which includes only internal edges. Otherwise, we say that $\overrightarrow{G}[U]$ is *ϵ-far* from being Eulerian. Similarly, we will say that $\overrightarrow{G}[U]$ is *(ϵ, β)-amendable* if and only if it has a β-correction subgraph of size at most ϵm_U which includes only internal edges. Otherwise, we say that $\overrightarrow{G}[U]$ is *(ϵ, β)-unamendable*. Note that we can view the external edges as comprising a knowledge graph (see Section 3). To ensure that $\overrightarrow{G}[U]$ can be made Eulerian (or β-Eulerian) by inverting internal edges only, we always assume that all the cuts in $\overrightarrow{G}$ are valid with respect to the orientation $\overrightarrow{\partial} U$ of the external edges. This implies in particular that

$$\overrightarrow{E}(U, V \setminus U) = \overrightarrow{E}(V \setminus U, U). \tag{16}$$

The next lemma shows that this assumption allows us to apply the same techniques as we did for the general testing problem.

Lemma 7.1 *If all the cuts in $\overrightarrow{G}$ are valid with respect to $\overrightarrow{\partial} U$, then:*

1. *$\overrightarrow{G}[U]$ can be made Eulerian by inverting internal edges along spring-drain paths.*
2. *$\overrightarrow{G}[U]$ can be made β-Eulerian by inverting internal edges along β-spring-drain paths.*
3. *If $\overrightarrow{G}[U]$ is β-Eulerian then it can be made Eulerian by inverting internal edges along spring-drain paths, where in each such path, both the spring and the drain are β-big.*

Proof. We first give a proof of Item 1, and later explain how to modify it so as to prove Item 2 and Item 3. Assume that there is a spring $s \in \overrightarrow{G}[U]$ with no path to any drain that is contained entirely in $\overrightarrow{G}[U]$. Let

$$X = \{u \in U \mid \text{ there is a directed path of internal edges from } s \text{ to } u\}.$$

As X contains no drains but at least one spring, more edges exit X than enter it. Furthermore, by the definition of X, all the edges that exit X are in $\overrightarrow{\partial} U$. Hence, the cut $(X, V \setminus X)$ is invalid with respect to $\overrightarrow{\partial} U$, a contradiction. Since inverting an internal spring-drain path does not change the orientation of the edges in ∂U, we may continue to invert such paths until $\overrightarrow{G}[U]$ becomes balanced.

To prove Item 2, note that any correction subgraph of internal edges, which exists by Item 1, contains a β-correction subgraph as a subgraph, and thus it is also internal in $\overrightarrow{G}[U]$. To obtain this

β-correction subgraph, we modify the correction subgraph by removing paths from β-big springs to β-big drains one by one as long as they exist.

To prove Item 3 we use the same proof as for Item 1. However, here we know that all our springs and drains are β-big, since $\overrightarrow{G}[U]$ is β-Eulerian. ■

We will be interested in induced subgraphs $\overrightarrow{G}[U]$ that are "lame directed expanders".

Definition 7.2 ((α, β)-expander) *Given a subset $U \subseteq V$ and a parameter $\beta > 0$, we say that a cut (A, B) of U is a β-cut of U if*

$$|E(B)| \geq |E(A)| \geq \beta.$$

Given parameters $\alpha, \beta > 0$, we say that the subgraph $\overrightarrow{G}[U]$ of G is an (α, β)-expander if for every β-cut (A, B) of U we have

$$|E(A,B)| - \left| |\overrightarrow{E}(V \setminus U, A)| - |\overrightarrow{E}(A, V \setminus U)| \right| \geq 2\alpha|E(A)|. \tag{17}$$

Note that to decide whether $\overrightarrow{G}[U]$ is an (α, β)-expander we do not need to know the orientation of its internal edges, but only that of its external edges. In particular, if the entire domain graph G is an α-expander, then $\overrightarrow{G}$ itself is always an (α, β)-expander for every $\beta > 0$ (because $G[V]$ has no external edges).

In the next two lemmas we give lower bounds for the numbers of internal spring-drain paths and β'-spring-drain paths (for some β') in an (α, β)-expander. Using our (p, β)-tests from Section 4 with these bounds, we will later obtain ϵ-tests for (α, β)-expanders. In the following, $m_U = |E(U)|$ and Δ_U is the maximum degree of a vertex in U (where the degree of a vertex $u \in U$ is the total number of edges incident with u, including external edges).

Lemma 7.3 *Let $\overrightarrow{G}[U]$ be an (α, β)-expander for some $0 < \alpha < 1$ and $0 \leq \beta$ and let $m_U \stackrel{\text{def}}{=} |E(U)|$. Suppose that all the cuts in G are valid with respect to $\overrightarrow{\partial} U$, and that $\overrightarrow{G}[U]$ is (ϵ', β')-unamendable for some $0 < \epsilon' < 1$ and $0 < \beta' \leq \Delta_U$. Then $\overrightarrow{G}[U]$ is p-far from being β'-Eulerian for $p(U, \alpha, \beta, \epsilon', \beta') = \Omega\left(\frac{\epsilon' m_U}{\log_{(1+\alpha)} m_U + \beta}\right) = \Omega\left(\frac{\epsilon' m_U}{\log m_U / \alpha + \beta}\right)$.*

Proof. We show that in any orientation $\overrightarrow{G}$ such that $\overrightarrow{G}[U]$ is not β'-Eulerian, there exists a spring-drain path inside $\overrightarrow{G}[U]$ of length $O(\log_{(1+\alpha)} m_U + \beta)$. Note that an (α, β)-expander $\overrightarrow{G}[U]$ remains an (α, β)-expander after we invert some of its internal edges. Thus, if $\overrightarrow{G}[U]$ is (ϵ', β')-unamendable, then we have to invert internal edges along $\Omega\left(\frac{\epsilon' m_U}{\log_{(1+\alpha)} m + \beta}\right)$ β'-spring-drain paths (as inverting spring-drains paths in which both the spring and the drain are β'-big is irrelevant for being β'-Eulerian). Thereof the lemma will follow.

Assume that $\overrightarrow{G}[U]$ is not β'-Eulerian, and let ℓ be the minimum length of a spring-drain path of internal edges in $\overrightarrow{G}[U]$. Such a path exists by Lemma 7.1. If $\ell \leq \beta$ then the claim is obviously true. We thus assume that $\ell > \beta$. Let S_U be the set of springs in U. Consider a BFS traversal of $\overrightarrow{G}[U]$ starting from $L_0 \stackrel{\text{def}}{=} S_U$. For every $i > 0$, set by induction

$$L_i \stackrel{\text{def}}{=} \{v \in U \mid \text{there exists } u \in L_{i-1} \text{ s.t. } (u, v) \in \overrightarrow{E}\}.$$

In addition, let $A_i \stackrel{\text{def}}{=} \bigcup_{0 \leq j < i} L_j$ and $B_i \stackrel{\text{def}}{=} \bigcup_{j \geq i} L_j$. Consider a level i such that $\beta < i \leq \ell$ and $|E(A_i)| \leq |E(B_i)|$. Clearly, (A_i, B_i) is a β-cut, and thus, since U is an (α, β)-expander, we have

$$|E(A_i, B_i)| + |\overrightarrow{E}(V \setminus U, A_i)| - |\overrightarrow{E}(A_i, V \setminus U)| \geq 2\alpha |E(A_i)|. \tag{18}$$

Note that for every $i \leq \ell$, the set A_i contains springs but no drains. Hence, more edges exit A_i then enter it:

$$|\overrightarrow{E}(A_i, B_i)| - |\overrightarrow{E}(B_i, A_i)| - |\overrightarrow{E}(V \setminus U, A_i)| + |\overrightarrow{E}(A_i, V \setminus U)| > 0,$$

and thus

$$|\overrightarrow{E}(A_i, B_i)| > \frac{1}{2}\left(|\overrightarrow{E}(A_i, B_i)| + |\overrightarrow{E}(B_i, A_i)| + |\overrightarrow{E}(V \setminus U, A_i)| - |\overrightarrow{E}(A_i, V \setminus U)|\right).$$

Substituting the above in Inequality (18) we have

$$|\overrightarrow{E}(A_i, B_i)| > \alpha |E(A_i)|$$

for every i such that $\beta < i \leq \ell$ and $|E(A_i)| \leq |E(B_i)|$. Recall that all the directed edges in $G[U]$ that exit A_i enter A_{i+1}. We thus have

$$|E(A_{i+1})| > (1+\alpha)|E(A_i)|,$$

and by induction,

$$|E(A_i)| > (1+\alpha)^{i-\beta}|E(A_\beta)| \geq (1+\alpha)^{i-\beta}\beta.$$

Therefore,

$$i - \beta \leq \log_{(1+\alpha)}\left(\frac{|E(A_i)|}{\beta}\right) < \log_{(1+\alpha)} m_U \tag{19}$$

for every i such that $\beta < i \leq \ell$ and $|E(A_i)| \leq |E(B_i)|$. Now, if for every $\beta < i \leq \ell$ we have $|E(A_i)| \leq |E(B_i)|$, then, from Inequality (19) we have $\ell - \beta < \log_{(1+\alpha)} m_U$ and thus $\ell = O(\log_{(1+\alpha)} m_U + \beta)$.

Otherwise, let $k > \beta$ be the minimal index i for which $|E(A_i)| > |E(B_i)|$. From Equation (19) we have

$$k - 1 - \beta < \log_{(1+\alpha)} m_U. \tag{20}$$

For every $k \leq i \leq \ell$ while $|E(B_i)| \geq \beta$, (B_i, A_i) is a β-cut, and thus from Inequality (17) we have

$$|E(A_i, B_i)| - |\overrightarrow{E}(V \setminus U, B_i)| + |\overrightarrow{E}(B_i, V \setminus U)| \geq 2\alpha |E(B_i)|. \tag{21}$$

Note that the set B_i contains drains but no springs. Hence, more edges enter B_i then exit it:

$$|\overrightarrow{E}(A_i, B_i)| - |\overrightarrow{E}(B_i, A_i)| + |\overrightarrow{E}(V \setminus U, B_i)| - |\overrightarrow{E}(B_i, V \setminus U)| > 0,$$

and thus

$$|\overrightarrow{E}(A_i, B_i)| > \frac{1}{2}\left(|\overrightarrow{E}(A_i, B_i)| + |\overrightarrow{E}(B_i, A_i)| - |\overrightarrow{E}(V \setminus U, B_i)| + |\overrightarrow{E}(B_i, V \setminus U)|\right).$$

Substituting the above in Inequality (21) we have

$$|\overrightarrow{E}(A_i, B_i)| > \alpha |E(B_i)|$$

for every i such that $k \le i \le \ell$ and $|E(A_i)| > |E(B_i)| \ge \beta$. Since all the edges in $\overrightarrow{E}(A_i, B_i)$ enter L_i, we have

$$|E(B_i)| > (1+\alpha)|E(B_{i+1})|,$$

and by induction

$$|E(B_i)| > (1+\alpha)^{j-i}|E(B_j)|$$

for every i, j such that $k \le i \le j \le \ell$ and $|E(A_j)| > |E(B_j)| \ge \beta$. Hence,

$$j - i \le \log_{(1+\alpha)} \left(\frac{|E(B_i)|}{|E(B_j)|} \right) < \log_{(1+\alpha)} m_U \tag{22}$$

for every i, j such that $k \le i \le j \le \ell$ and $|E(A_j)| > |E(B_j)| \ge \beta$. Now, if $|E(A_\ell)| > |E(B_\ell)| \ge \beta$ then, taking $i = k$ and $j = \ell$ we have

$$\ell - k < \log_{(1+\alpha)} m_U.$$

Combined with Inequality (20) we obtain

$$\ell < 2\log_{(1+\alpha)} m_U + \beta + 1 = O(\log_{(1+\alpha)} m_U + \beta).$$

Otherwise, let r be the minimum index i such that $|E(B_i)| < \beta$. Then, for $i = k$ and $j = r - 1$, Inequality (22) yields

$$r - 1 - k < \log_{(1+\alpha)} m_U. \tag{23}$$

Since $|E(B_r)| < \beta$, there are less then β levels between r and ℓ and thus $\ell < r + \beta$. Hence, with Inequalities (20) and (23) we achieve

$$\ell < 2\log_{(1+\alpha)} m_U + 2\beta + 2 = O(\log_{(1+\alpha)} m_U + \beta). \qquad \blacksquare$$

Lemma 7.4 *Let $\overrightarrow{G}[U]$ be an (α, β)-expander for some $0 < \alpha < 1$ and $0 \le \beta \le \frac{\Delta_U}{2}$ and let $m_U \stackrel{\text{def}}{=} |E(U)|$. Suppose that all the cuts in G are valid with respect to $\overrightarrow{\partial} U$. Suppose further that for some $\epsilon > 0$, $\overrightarrow{G}[U]$ is ϵ-far from being Eulerian, but still $(\frac{\epsilon}{2}, 2\beta)$-amendable. Then $\overrightarrow{G}[U]$ is p'-far from being Eulerian for $p'(U, \alpha, \beta, \epsilon) = \Omega\left(\frac{\epsilon m_U}{\log_{(1+\alpha)} m_U}\right) = \Omega\left(\frac{\alpha \epsilon m_U}{\log m_U}\right)$.*

Proof. As $\overrightarrow{G}[U]$ is $(\frac{\epsilon}{2}, 2\beta)$-amendable, there exists a 2β-correction subgraph $\overrightarrow{H}$ of $\overrightarrow{G}[U]$ of size at most $\epsilon m_U/2$. Consider $\overrightarrow{G}_1[U] \stackrel{\text{def}}{=} \overrightarrow{G}[U]_{\overleftarrow{H}}$, that is, the digraph obtained from $\overrightarrow{G}[U]$ by inverting all the edges in the 2β-correction subgraph $\overrightarrow{H}$. Then $\overrightarrow{G}_1[U]$ is 2β-balanced. However, since $\overrightarrow{G}[U]$ is ϵ-far from being balanced, at least $\epsilon m_U/2$ more edges must be inverted in order to make it Eulerian. By Item 3 of Lemma 7.1, there exists a correction subgraph for $\overrightarrow{G}[U]$ which is a union of internal paths from 2β-big springs to 2β-big drains.

To complete the proof, we show that while $\overrightarrow{G}_1[U]$ is not Eulerian, it contains an internal spring-drain path of length $\ell = O(\log_{(1+\alpha)} m_U) = O(\log m_U/\alpha)$. This is done similarly to the proof in Lemma 7.3 which shows the existence of a short β'-spring-drain path. However, note that now the set $E(A_2)$ includes all the edges outgoing from at least one 2β-big spring, and thus $E(A_i, B_i)$ is a β-cut for every $i \geq 2$ such that $|E(A_i)| \leq |A(B_i)|$. Hence, instead of Inequality (19) we have

$$i - 2 \leq \log_{(1+\alpha)} \left(\frac{|E(A_i)|}{|E(A_2)|} \right) < \log_{(1+\alpha)} m_U \tag{24}$$

for every i such that $1 < i \leq \ell$ and $|E(A_i)| \leq |E(B_i)|$. Furthermore, since all the drains are 2β-big, the set $E(B_{\ell-1})$ includes all the edges incoming to at least one 2β-big drain, and thus $E(B_i, A_i)$ is a β-cut for every $i \leq \ell - 1$ such that $|E(A_i)| > |A(B_i)|$. Hence, putting $j = \ell - 1$ in Inequality (22), we obtain

$$\ell - i - 1 \leq \log_{(1+\alpha)} \left(\frac{|E(B_i)|}{|E(B_{\ell-1})|} \right) < \log_{(1+\alpha)} m_U \tag{25}$$

for every $i \leq \ell - 1$ such that $|E(A_i)| > |A(B_i)|$. Let k be the minimum value of i for which $|E(A_i)| > |A(B_i)|$. Then, from Inequality (24) we have $k < \log_{(1+\alpha)} m_U + 3$ and from Inequality (25) we have $\ell - k < \log_{(1+\alpha)} m_U + 1$, and hence $\ell = O(\log_{(1+\alpha)} m_U)$. ■

Suppose that $\overrightarrow{G}$ is ϵ-far from being Eulerian and that the external edges of U do not induce an invalid cut. Note that if $\beta \leq \frac{\Delta_U}{2}$, then either the conditions of Lemma 7.3 apply for $\epsilon' = \epsilon/2$ and $\beta' = 2\beta$, or the conditions of Lemma 7.4 apply. Also, note that if $\beta > \frac{\Delta_U}{2}$ then the conditions of Lemma 7.3 apply for $\epsilon' = \epsilon$ and $\beta' = \Delta_U$. We thus obtain the two ϵ-tests below. Note that whenever our tests use samples of edges incident with a vertex $u \in U$, the sampling is among all the edges incident with u, and not only internal edges.

Algorithm 7.5 *GEN-1($\overrightarrow{G}[U]$, α, β, ϵ)*

1. *If $\beta \leq \frac{\Delta_U}{2}$ then run GENERIC-1($\overrightarrow{G}[U]$, $p(U, \alpha, \beta, \epsilon/2, 2\beta)$, Δ_U), and otherwise run GENERIC-1($\overrightarrow{G}[U]$, $p(U, \alpha, \beta, \epsilon, \Delta_U)$, Δ_U). In both cases, $p(U, \alpha, \beta, \epsilon', \beta')$ is the lower bound given in Lemma 7.3.*

2. *If $\beta \leq \frac{\Delta_U}{2}$ then run GENERIC-1($\overrightarrow{G}[U]$, $p'(U, \alpha, \beta)$, Δ_U), where $p'(U, \alpha, \beta)$ is the lower bound given in Lemma 7.4.*

3. *Reject if at least one of the tests has rejected, and accept otherwise.*

Algorithm 7.6 *MULTI-2($\overrightarrow{G}[U]$, α, β, ϵ)*

1. *If $\beta \leq \frac{\Delta_U}{2}$ then run MULTISTAGE-2($\overrightarrow{G}[U]$, $p(U, \alpha, \beta, \epsilon/2, 2\beta)$, Δ_U), and otherwise run MULTISTAGE-2($\overrightarrow{G}[U]$, $p(U, \alpha, \beta, \epsilon, \Delta_U)$, Δ_U). In both cases, $p(U, \alpha, \beta, \epsilon', \beta')$ is the lower bound given in Lemma 7.3.*

2. *If $\beta \leq \frac{\Delta_U}{2}$ then run MULTISTAGE-2($\overrightarrow{G}[U]$, $p'(U, \alpha, \beta)$, Δ_U), where $p'(U, \alpha, \beta)$ is the lower bound given in Lemma 7.4.*

3. *Reject if at least one of the tests has rejected, and accept otherwise.*

Combining Lemma 7.3 and Lemma 7.4 with Lemma 4.4 and Lemma 4.6, we obtain the following lemmas, which will be used in Section 8.

Lemma 7.7 *GEN-1($\overrightarrow{G}[U], \alpha, \beta, \epsilon$) is a 1-sided ϵ-test for an (α, β)-expander subgraph $\overrightarrow{G}[U]$, assuming that the external edges of U are known and do not induce an invalid cut. The query complexity of the test is $O\left(\frac{\Delta_U \log m_U}{\epsilon\alpha} + \frac{\beta \cdot \min\{\beta, \Delta_U\}}{\epsilon}\right)$, where $m_U = |E(U)|$ and $\Delta_U = \max\{\deg(u) \mid u \in U\}$.*

Lemma 7.8 *MULTI-2($\overrightarrow{G}[U], \alpha, \beta, \epsilon$) is a 2-sided ϵ-test for an (α, β)-expander subgraph $\overrightarrow{G}[U]$, assuming that the external edges of U are known and do not induce an invalid cut. The query complexity of the test is $\widetilde{O}\left(\frac{\sqrt{\Delta_U} \log m_U}{\epsilon\alpha} + \frac{\beta \cdot \sqrt{\min\{\beta, \Delta_U\}}}{\epsilon}\right)$, where $m_U = |E(U)|$ and $\Delta_U = \max\{\deg(u) \mid u \in U\}$.*

8 General tests based on chopping

In this section we use our results from Section 7 to provide a 1-sided test and a 2-sided test as follows. Given an orientation $\overrightarrow{G}$ of an Eulerian graph G, we show how to decompose $\overrightarrow{G}$ into a collection of (α, β)-expanders with a relatively small number of edges that are outside the (α, β)-expanders, called henceforth *external edges*. We will find this "chopping" adaptively while querying external edges only. If we do not find a witness showing that $\overrightarrow{G}$ is not Eulerian already during the chopping procedure, then we sample a few (α, β)-expanders and test them using GEN-1 (Algorithm 7.5) or using MULTI-2 (Algorithm 7.6), obtaining a 1-sided test or a 2-sided test respectively.

Lemma 8.1 (The chopping lemma) *Given an orientation $\overrightarrow{G}$ as input and parameters $\alpha, \beta > 0$, we can either find a witness showing that $\overrightarrow{G}$ is not Eulerian, or find non-empty induced subgraphs $\overrightarrow{G}_i = (V_i, \overrightarrow{E}_i = \overrightarrow{E}(V_i))$ of $\overrightarrow{G}$ (where $i = 1, \ldots, k$ for some k), which we call (α, β)-*components *(or simply* components*), that satisfy the following:*

1. *The vertex sets $V_1, \ldots, V_k$ of the components are mutually disjoint.*
2. *$|\overrightarrow{E}_i| \geq \beta$ for $i = 1, \ldots k$.*
3. *All the components $\overrightarrow{G}_i$ are (α, β)-expanders.*
4. *The total number of external edges satisfies*

$$|\overrightarrow{E} \setminus \bigcup_{i=1,\ldots,k} \overrightarrow{E}(V_i)| = O(\alpha m^2 \log m / \beta).$$

During the chopping procedure, we query only external edges, i.e., edges that are not in any component G_i. The query complexity is of the same order also if we find a witness that $\overrightarrow{G}$ is not Eulerian.

Proof. The chopping procedure proceeds as follows. At first, we define $\overrightarrow{G} = \overrightarrow{G}[V]$ as our single component. Then, in each step, we decompose a component $\overrightarrow{G}[U]$ into two separate components $\overrightarrow{G}[A]$ and $\overrightarrow{G}[B]$, such that (A, B) is a β-cut of U and

$$|E(A, B)| - \left| |\overrightarrow{E}(V \setminus U, A)| - |\overrightarrow{E}(A, V \setminus U)| \right| < 2\alpha |E(A)|. \tag{26}$$

When decomposing, we query the edges of the cut (A, B) and mark them as external edges. Note that we need not query any additional edges to decide on cutting a component, as all the required information is given by the domain graph G and the orientation of the external edges that were queried in previous steps. After each stage, we check whether the orientations of the edges queried so far invalidate any of the cuts in the graph (see Section 3), in which case we conclude that $\overrightarrow{G}$ is not Eulerian and return the invalid cut.

The procedure terminates once there is no cut of any component that satisfies the chopping conditions. The components are clearly disjoint throughout the procedure. Since we only chopped components across β-cuts, every final component contains at least β edges. Moreover, note that a component is always chopped by the procedure unless all its β-cuts satisfy Inequality (17). Hence, if the algorithm terminates without finding a witness that $\overrightarrow{G}$ is not Eulerian, then every $\overrightarrow{G}_i$ is an (α, β)-expander.

It remains to prove the upper bound for the number of external edges and the query complexity of the chopping procedure. Consider a component U and a β-cut (A, B) of U that was queried in some step of the lemma. Suppose that the cut $(A, V \setminus A)$ is valid. Then

$$|\overrightarrow{E}(A, B)| - |\overrightarrow{E}(B, A)| + |\overrightarrow{E}(A, V \setminus U)| - |\overrightarrow{E}(V \setminus U, A)| = 0. \tag{27}$$

From the chopping condition (26) we have

$$|\overrightarrow{E}(A, B)| + |\overrightarrow{E}(B, A)| - \Big||\overrightarrow{E}(V \setminus U, A)| - |\overrightarrow{E}(A, V \setminus U)|\Big| < 2\alpha |E(A)|. \tag{28}$$

Combining Equations (27) and (28) and considering two cases depending on whether $|\overrightarrow{E}(A, V \setminus U)| - |\overrightarrow{E}(V \setminus U, A)|$ is positive or negative, we obtain that

$$\min\Big\{|\overrightarrow{E}(A, B)|, |\overrightarrow{E}(B, A)|\Big\} < \alpha |E(A)|. \tag{29}$$

Hence, if Inequality (29) does *not* hold, then, in any case, after querying the edges in (A, B) we discover an invalid cut in the graph (which could be (A, B) or another cut). We next compute the query complexity in the case where our knowledge graph contains no invalid cuts throughout the procedure, and later show how to modify our analysis for the case where an invalid cut is detected after querying the last cut.

For every β-cut (A, B) that we use for partitioning, we refer to the edges in the minimal cut among $\overrightarrow{E}(A, B)$ and $\overrightarrow{E}(B, A)$ as *rare* edges, and to the edges in the other direction as *common* edges. Let us first compute the cost of querying the rare edges only. For every partition of a β-cut (A, B), we "charge" a cost of α on every edge in $E(A)$. From Inequality (29), the sum of charges is larger than the number of rare edges queried. Since by our definition $|E(A)| \le |E(B)|$, every edge can belong to "Side" A of a partitioned β-cut (A, B) at most $\log m$ times throughout the entire procedure. Hence, the sum of charges is at most $\alpha \log m$ for every edge in the graph and at most $\alpha m \log m$ in total. We complete the proof by showing that the ratio between the number of external edges and the number of rare external edges is $O(m/\beta)$.

Consider the component multigraph $\overrightarrow{G}_{comp}$, whose vertices are the components G_i and whose edges are the external edges of $\overrightarrow{G}$. By our assumption that the knowledge graph contains no invalid cuts, $\overrightarrow{G}_{comp}$ is Eulerian (as a directed multigraph), and hence it is an edge-disjoint union of simple directed cycles. However, $\overrightarrow{G}_{comp}$ does not contain a directed cycle of common edges. This can be proved by induction as follows. When starting the chopping procedure there is only one component and no external edges. Then at every step we partition one component into two sets, and add the common edges, if any, from one of the sets to the other. One can see that this cannot create directed cycles of common edges. Therefore, $\overrightarrow{G}_{comp}$ is an edge-disjoint union of simple directed cycles, where each cycle contains at least one rare edge. As the number of vertices in $\overrightarrow{G}_{comp}$ is at most m/β, we conclude that the number of external edges in $\overrightarrow{G}$ is at most m/β times the number of rare edges. From the discussion above, it follows that the number of external edges in $\overrightarrow{G}$ is $O(\alpha m^2 \log m/\beta)$.

Regarding the case where querying the edges of a β-cut (A, B) reveals a violation of a cut in the graph, recall that as long as the knowledge graph does not induce an invalid cut, there exists an Eulerian orientation extending the knowledge graph (see Lemma 3.2). Thus, the edges of (A, B) have a "good" orientation that does not violate any cut in the graph. Clearly, the total query complexity after we query the edges of (A, B) and terminate is no higher than in the case where we would have queried the edges of (A, B) and discovered the good orientation. ■

We are now ready to present our 1-sided test. In the following, let $\alpha, \beta > 0$ be parameters.

Algorithm 8.2 *CHOP-1($\overrightarrow{G}, \epsilon, \alpha, \beta$):*

1. *Use Lemma 8.1 (the chopping lemma) for finding (α, β)-components $\overrightarrow{G}_1, \ldots, \overrightarrow{G}_k$ and querying their external edges, or reject and terminate if an invalid cut is found in the process.*

2. *Sample $3 \ln 3/\epsilon$ (α, β)-components $\overrightarrow{G}_i$ randomly and independently, where the probability of selecting a component $\overrightarrow{G}_i$ in each sample is proportional to $m_i \stackrel{\text{def}}{=} |E(V_i)|$.*

3. *Test every selected component $\overrightarrow{G}_i$ using GEN-1($\overrightarrow{G}_i, \alpha, \beta, \epsilon/2$) (Algorithm 7.5). Reject if the test rejects for at least one of the components selected.*

4. *Accept if the input was not rejected by any of the above steps.*

Lemma 8.3 *If $\overrightarrow{G}$ is Eulerian then CHOP-1 accepts $\overrightarrow{G}$ with probability 1.*

Proof. If $\overrightarrow{G}$ is Eulerian then $\overrightarrow{G}$ has no invalid cuts and therefore CHOP-1 does not reject in Step 1. In addition, every (α, β)-component of $\overrightarrow{G}$ is Eulerian, and since GEN-1 is 1-sided (see Lemma 7.7), CHOP-1 does not reject any of the tested (α, β)-components. ■

Lemma 8.4 *If the external edges induce an invalid cut, or if more than an $\epsilon/2$-fraction of the internal edges are in (α, β)-components that are $\epsilon/2$-far from being Eulerian, then CHOP-1 rejects with probability at least $2/3$.*

Proof. If the external edges induce an invalid cut then the algorithm rejects in Step 1 while trying to perform the chopping. Otherwise, every (α, β)-component sampled in Step 2 is $\epsilon/2$-far from being Eulerian with probability at least $\epsilon/2$. By Lemma 7.7, for every (α, β)-component that is $\epsilon/2$-far from being Eulerian, the rejection probability is at least $2/3$. Thus, the probability of rejecting one sampled component is at least $\epsilon/3$. Since the components are selected independently, the probability of accepting all the components is at most $(1 - \epsilon/3)^{3 \ln 3/\epsilon} < e^{-\ln 3} = 1/3$. ■

Lemma 8.5 *If the external edges do not induce an invalid cut and at most an $\epsilon/2$-fraction of the internal edges are in (α, β)-components that are $\epsilon/2$-far from being Eulerian, then $\overrightarrow{G}$ is ϵ-close to being Eulerian.*

Proof. By Item 1 of Lemma 7.1, every (α, β)-component $\overrightarrow{G}_i$ in $\overrightarrow{G}$ can be made Eulerian by inverting internal edges of $\overrightarrow{G}_i$. We thus orient the (α, β)-components that are $\epsilon/2$-far from being Eulerian so as to make them Eulerian. These components consist of at most $\epsilon m/2$ edges. In addition, we invert a minimum number of edges in each of the (α, β)-components that are $\epsilon/2$-close to being Eulerian, so as to make them Eulerian too. This requires at most $\epsilon m/2$ more alterations. We thus obtain an Eulerian orientation that is ϵ-close to $\overrightarrow{G}$. ■

Theorem 8.6 *CHOP-1 is a 1-sided test for being Eulerian with query complexity*

$$O\left(\frac{\alpha m^2 \log m}{\beta} + \frac{\Delta \log m}{\epsilon^2 \alpha} + \frac{\beta \cdot \min\{\beta, \Delta\}}{\epsilon^2}\right).$$

In particular, for $\alpha = \frac{(\Delta \log m)^{1/3}}{(\epsilon m)^{2/3}}$ and $\beta = \frac{(\epsilon m \log m)^{2/3}}{\Delta^{1/3}}$, the query complexity is

$$O\left(\frac{(\Delta m \log m)^{2/3}}{\epsilon^{4/3}}\right) = O\left(\left(\frac{\Delta}{\epsilon}\right)^{4/3} (n \log n)^{2/3}\right).$$

Proof. The correctness of the test follows from Lemmas 8.3, 8.4, and 8.5. The query complexity follows from Lemmas 7.7 and 8.1 (the chopping lemma). ■

We note that Theorem 8.6 provides a sub-linear algorithm for every graph of maximum degree $\Delta = o\left(\frac{\epsilon^2\sqrt{m}}{\log m}\right)$. For nearly regular graphs, i.e. graphs with $m = \Omega(\Delta n)$, the algorithm is sub-linear for every $\Delta = o\left(\frac{\epsilon^4 n}{\log^2 n}\right)$.

We conclude with a similar 2-sided test which gives a sub-linear query complexity for all graphs. In the following, let $\alpha, \beta > 0$ be parameters.

Algorithm 8.7 *CHOP-2($\overrightarrow{G}, \epsilon, \alpha, \beta$):*

1. *Use Lemma 8.1 (the chopping lemma) for finding (α, β)-components $\overrightarrow{G}_1, \dots, \overrightarrow{G}_k$ and querying their external edges, or reject and terminate if an invalid cut is found in the process.*
2. *Sample $\frac{3}{\epsilon}$ (α, β)-components $\overrightarrow{G}_i$ independently, where the probability of selecting a component $\overrightarrow{G}_i$ is proportional to $m_i \stackrel{\text{def}}{=} |E(V_i)|$.*
3. *Test every selected component $\overrightarrow{G}_i$ for being Eulerian $12\ln(9/\epsilon)$ times independently using MULTI-2($\overrightarrow{G}_i, \alpha, \beta, \epsilon/2$) (Algorithm 7.6). Reject if there is a component $\overrightarrow{G}_i$ which was rejected by at least half of its tests.*
4. *Accept if the input was not rejected in a previous step.*

Lemma 8.8 *If $\overrightarrow{G}$ is Eulerian then CHOP-2 accepts with probability at least $2/3$.*

Proof. If $\overrightarrow{G}$ is Eulerian then $\overrightarrow{G}$ has no invalid cuts and therefore CHOP-2 does not reject in Step 1. In addition, all the (α, β)-components $\overrightarrow{G}_i$ are Eulerian. Thus, by Lemma 7.8, every run of MULTI-2 on a component $\overrightarrow{G}_i$ rejects with probability at most $1/3$. By standard large deviation arguments, the probability of rejecting a component $\overrightarrow{G}_i$ by at most half of its tests is at most $\epsilon/9$. Applying the union bound for the $3/\epsilon$ (α, β)-components sampled, the probability of rejecting $\overrightarrow{G}$ is at most $1/3$. ∎

Lemma 8.9 *If the external edges induce an invalid cut, or if more than $\epsilon/2$-fraction of the internal edges are in (α, β)-components that are $\epsilon/2$-far from being Eulerian, then Algorithm 8.7 rejects with probability at least $2/3$.*

Proof. If the external edges induce an invalid cut then the algorithm rejects in Step 1 while trying to perform the chopping. Otherwise, the probability of not sampling any (α, β)-component that is $\epsilon/2$-far from being Eulerian is at most $(1-\frac{\epsilon}{2})^{3/\epsilon} < \frac{1}{4}$. Suppose that we have sampled at least one (α, β)-component that is ϵ-far from being Eulerian. By Lemma 7.8, the acceptance probability of a single test of this component is at most $1/3$. Using standard large deviation arguments, the probability of accepting in at most half of the tests of this component is smaller than $1/12$. We conclude that the probability of accepting $\overrightarrow{G}$ in this case is at most $1/3$. ∎

Observation 8.10 *Lemma 8.5 is true for CHOP-2 as well as for CHOP-1.* ∎

Theorem 8.11 *Algorithm 8.7 is a 2-sided test for being Eulerian with query complexity*

$$O\left(\frac{\alpha m^2 \log m}{\beta}\right) + \widetilde{O}\left(\frac{\sqrt{\Delta}\log m}{\epsilon^2 \alpha} + \frac{\beta \cdot \sqrt{\min\{\beta, \Delta\}}}{\epsilon^2}\right).$$

In particular, if $\Delta \le (\epsilon m)^{4/7}$, then for $\alpha = \frac{\Delta^{1/6}}{(\epsilon m)^{2/3}}$ and $\beta = \frac{(\epsilon m)^{2/3}}{\Delta^{1/6}}$ the query complexity is $\widetilde{O}\left(\frac{\Delta^{1/3} m^{2/3}}{\epsilon^{4/3}}\right) = \widetilde{O}\left(\frac{m^{6/7}}{\epsilon^{8/7}}\right)$. If $(\epsilon m)^{4/7} < \Delta \le m$, then for $\alpha = \frac{\Delta^{5/16}}{(\epsilon m)^{3/4}}$ and $\beta = \Delta^{1/8}\sqrt{\epsilon m}$ the query complexity is $\widetilde{O}\left(\frac{\Delta^{3/16} m^{3/4}}{\epsilon^{5/4}}\right) = \widetilde{O}\left(\frac{m^{15/16}}{\epsilon^{5/4}}\right)$.

Proof. The correctness of the test follows from Lemmas 8.8, 8.9, and 8.5. The query complexity follows from Lemmas 7.8 and 8.1 (The chopping lemma). ∎

9 Lower bounds for bounded-degree graphs

9.1 A 2-sided lower bound

In this subsection we prove the following theorem.

Theorem 9.1 *For every $0 < \epsilon \le 1/64$, every non-adaptive (2-sided) ϵ-test for Eulerian orientations of bounded degree graphs must use $\Omega\left(\sqrt{\frac{\log m}{\log \log m}}\right)$ queries. Consequently, every adaptive test requires $\Omega(\log \log m)$ queries.*

The main idea of the proof uses Yao's principle [27]. Namely, for infinitely many natural numbers ℓ, we define a graph G_ℓ with $m = 2\ell^2$ edges and two distributions over the orientations of G_ℓ. The first distribution, $\mathcal{P}_\ell$, contains only Eulerian orientations of G_ℓ, while the second distribution, $\mathcal{F}_\ell$, contains orientations that are with high probability 1/64-far from being Eulerian. We then show that any non-adaptive deterministic algorithm which makes $o\left(\sqrt{\frac{\log m}{\log\log m}}\right)$ queries cannot distinguish between the distributions $\mathcal{P}_\ell$ and $\mathcal{F}_\ell$ with probability higher than 1/5.

All our underlying graphs G_ℓ are two dimensional *tori*, which are 4-regular graphs having a highly symmetric structure (the exact definition is given below). We exploit this symmetry to construct distributions $\mathcal{P}_\ell$ and $\mathcal{F}_\ell$ such that, for any fixed set Q of $o\left(\sqrt{\frac{\log m}{\log\log m}}\right)$ edges, with high probability, the orientation of every pair of edges in Q has either no correlation in any of the distributions, or a correlation that is identical in both distributions.

We build the orientations in $\mathcal{P}_\ell$ and $\mathcal{F}_\ell$ from repeated "patterns" of varying sizes and show that, in order to distinguish between the distributions, a deterministic algorithm must be approximately "synchronized" with the (unknown) size of these patterns.

9.1.1 Preliminaries

For $i, j \in [\ell]$ we let $i \oplus j$ denote addition modulo ℓ, that is:

$$i \oplus j = \begin{cases} i + j & , \quad i + j \le \ell \\ i + j - \ell & , \quad i + j > \ell \end{cases} .$$

Given a graph $G = (V, E)$ and two vertices $u, v \in V$, we define the *distance* between u and v (or shortly $\mathrm{dist}(u, v)$) as the the walking distance in G between u and v. Given two edges $e_1, e_2 \in E$, we define the *distance* between e_1 and e_2 (or shortly $\mathrm{dist}(e_1, e_2)$) as the minimal distance between an endpoint of e_1 and an endpoint of e_2. For an edge $e = \{u, v\} \in E$ and a vertex $w \in V$, we define the *distance* of e from w (or shortly $\mathrm{dist}(e, w)$) as the minimum of $\mathrm{dist}(u, w)$ and $\mathrm{dist}(v, w)$. We stress that even when we consider an orientation $\overrightarrow{G}$, the distances between edges and vertices are still measured on the underlying undirected graph G for the purpose of the following proofs.

Definition 9.2 (Torus) *A torus is a two dimensional cyclic grid. Formally, an $\ell \times \ell$* torus *is the graph $T = (V, E)$ on $n = \ell^2$ vertices $V = \{v_{i,j} : i, j \in [\ell]\}$ and $m = 2\ell^2$ edges $E = E_H \cup E_V$, where $E_H = \Big\{\{v_{i,j_1}, v_{i,j_2}\} : j_2 = j_1 \oplus 1\Big\}$ and $E_V = \Big\{\{v_{i_1,j}, v_{i_2,j}\} : i_2 = i_1 \oplus 1\Big\}$. We refer to E_H as the set of* horizontal edges *and to E_V as the set of* vertical edges. *Two edges $e_1, e_2 \in E$ are said to be* perpendicular *if one of them is horizontal and the other is vertical, and otherwise they are called* parallel.

Given an orientation $\overrightarrow{T}$ of T, we say that a horizontal edge $e = \{v_{i,j}, v_{i,j\oplus 1}\}$ is directed to the right *if $v_{i,j}$ is the start-point of e, and otherwise we say that e is* directed to the left. *Similarly, we say that a vertical edge $e = \{v_{i,j}, v_{i\oplus 1,j}\}$ is* directed upwards *if $v_{i,j}$ is the start-point of e, and otherwise we say it is* directed downwards.

To simplify the presentation, we assume throughout this section that ℓ is even. We now define a graph operation that will be used later in the construction of the distributions $\mathcal{P}_\ell$ and $\mathcal{F}_\ell$.

Definition 9.3 (**(a,b)-shifting**) *Let $\overrightarrow{T}$ be an orientation of an $\ell \times \ell$* torus *T, and let $a,b \in [\ell]$. We define the (a,b)*-shifting of *$\overrightarrow{T}$ to be the orientation $\overrightarrow{T_{a,b}}$ of T, which is a transformation of the orientation $\overrightarrow{T}$ a units upwards and b units rightward. Namely, for every edge e of T, if $e = \{v_{i,j}, v_{i',j'}\}$ is directed from $v_{i,j}$ to $v_{i',j'}$ in $\overrightarrow{T}$ then $e_{a,b} \stackrel{\text{def}}{=} \{v_{i\oplus a,j\oplus b}, v_{i'\oplus a,j'\oplus b}\}$ is directed from $v_{i\oplus a,j\oplus b}$ to $v_{i'\oplus a,j'\oplus b}$ in $\overrightarrow{T_{a,b}}$.*

9.1.2 Defining auxiliary distributions

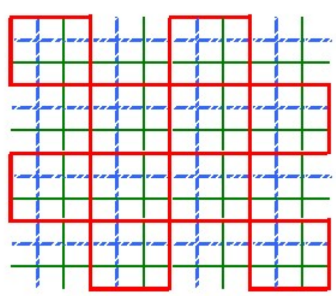
Let $H = (V,E)$ be an $\ell \times \ell$ torus, where $V = \{v_{i,j} \mid i,j \in [\ell]\}$, using the same indexing as in Definition 9.2. We define two simple distributions, $\mathcal{R}_\ell$ and $\mathcal{C}_\ell^{(k)}$, over the orientations of H. We later use these distributions to build the final distributions, $\mathcal{F}_\ell$ and $\mathcal{P}_\ell$.

The distribution $\mathcal{R}_\ell$ is simply a random orientation of H's edges. Namely, in $\overrightarrow{H} \sim \mathcal{R}_\ell$ the orientation of every edge $e \in E$ is chosen uniformly at random, independently of the other edges.

Lemma 9.4 *Let $\overrightarrow{H}$ be an orientation of H, chosen according to the distribution $\mathcal{R}_\ell$. Then with probability $1 - o(1)$, there are at least $\ell^2/4$ unbalanced vertices in $\overrightarrow{H}$.*

Proof. Define a subset $I = \{v_{i,j} \mid i + j \text{ is even}\} \subseteq V$ of $\ell^2/2$ vertices. Observe that I is an independent set in H (i.e. it has no internal edges), and so every vertex $v_i \in I$ is balanced with probability $x_i = \binom{4}{2}\left(\frac{1}{2}\right)^4 = 3/8$ independently from all other vertices in I. By Chernoff's inequality, the probability that at least half of the vertices in I are balanced is bounded by $\exp(-\ell^2/64)$. Namely, with probability $1 - o(1)$ there are at least $\ell^2/4$ unbalanced vertices in I. ■

Figure 4: *Partitioning H into edge-disjoint $4k$-cycles. Here $\ell = 12$ and $k = 3$.*



For a parameter k, $\mathcal{C}_\ell^{(k)}$ is a distribution over Eulerian orientations of H. We assume that $2k$ divides $\ell = \sqrt{m/2}$. To construct an orientation according to $\mathcal{C}_\ell^{(k)}$, we first partition the edges of H into edge-disjoint "square-shaped" $4k$-cycles as follows (see Figure 4). For every $0 \le i < \ell$, $0 \le j < \ell/2k$, we let $v_{i,2kj\oplus i}$ be a "lower-left corner" of a cycle C. The other "corner" vertices of C are $v_{i\oplus k,2kj\oplus i}$, $v_{i\oplus k,2kj\oplus i\oplus k}$, and $v_{i,2kj\oplus i\oplus k}$. The four corner vertices are connected by two paths of k horizontal edges and two paths of k vertical edges. One can see that this indeed forms a partition of all of H's edges into edge-disjoint cycles. Then, for every cycle C, we randomly and independently choose one of C's two possible Eulerian orientations. Let $\overrightarrow{H}'$ denote the orientation of H at this stage. Finally, $a, b \in [\ell]$ are chosen uniformly at random, and $\overrightarrow{H}$ is set to be the (a, b)-shifting of $\overrightarrow{H}'$.

In what follows, for a pair of edges $e_i, e_j \in E$ and an orientation $\overrightarrow{H}$ of H, we say that e_i and e_j are *independent* if either $\overrightarrow{H} \sim \mathcal{R}_\ell$, or $\overrightarrow{H} \sim \mathcal{C}_\ell^{(k)}$ and the edges e_i and e_j reside in different $4k$-cycles C_i and C_j. A set $Q \subseteq E$ is called *independent* if all the pairs $e_1, e_2 \in Q$ are independent. Observe that if Q is independent, then the orientation of every $e \in Q$ is distributed uniformly at random, independently of the orientation of all other members of Q. Clearly, if $\overrightarrow{H}$ is distributed according to $\mathcal{R}_\ell$ then every set $Q \subseteq E$ is independent. In the following lemmas we prove that, under some conditions, the set Q is independent with high probability also if $\overrightarrow{H}$ is distributed according to $\mathcal{C}_\ell^{(k)}$.

Lemma 9.5 *Let $e_1, e_2 \in E$ be two perpendicular edges of H. Let $\overrightarrow{H}$ be an orientation of H distributed according to $\mathcal{C}_\ell^{(k)}$, for an integer k that divides $\ell/2$. Then the probability that e_1 and e_2 are independent is at least $1 - \frac{1}{2k}$.*

Proof. Suppose that e_1 and e_2 are not independent. Hence, they both reside in the same cycle C in the partition of H's edges into $4k$-cycles. Note that, in such a case, e_1 and e_2 define a unique square-shaped $4k$-cycle in which they both reside, and hence, they define a unique vertex in H that must be the lower-left corner of this cycle. By the definition of the partition into $4k$-cycles, it is easy to see that the fraction of vertices in H that are corner vertices is $\frac{1}{2k}$. The lemma follows since in the last stage of constructing an orientation from $\mathcal{C}_\ell^{(k)}$, the partition into $4k$-cycles is randomly shifted. ■

Lemma 9.6 *Let k be an integer that divides $\ell/2$. Let $Q \subseteq E$ be a set of $o(\sqrt{k})$ edges such that for every pair $e_1, e_2 \in Q$, either $\mathrm{dist}(e_1, e_2) > 2k$, or e_1 and e_2 are perpendicular. Then for an orientation $\overrightarrow{H}$ of H distributed according to $\mathcal{C}_\ell^{(k)}$, the probability that Q is independent is $1 - o(1)$.*

Proof. Fix a pair $e_1, e_2 \in Q$. If $\mathrm{dist}(e_1, e_2) > 2k$ then e_1 and e_2 must reside in different $4k$-cycles, and hence they are independent. Otherwise, e_1, e_2 are perpendicular, and by Lemma 9.5 they are independent with probability at least $1 - \frac{1}{2k}$. The proof is now completed by applying the union bound for all $o(k)$ pairs $e_1, e_2 \in Q$. ■

9.1.3 Defining the main distributions

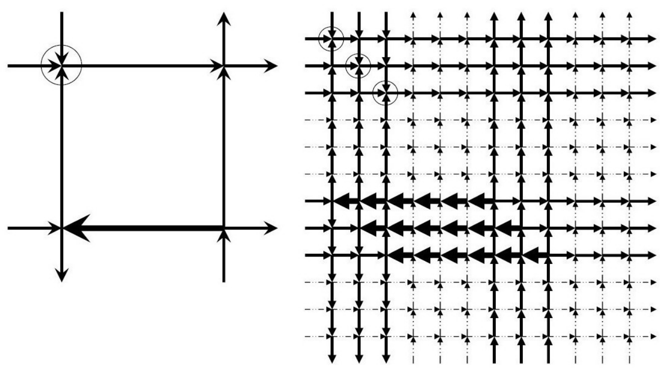
We now give our two main distributions of torus orientations. First, we need to define the following operation.

Definition 9.7 (t-tiling) *Let $\ell, t > 0$. Let $\overrightarrow{H} = (V(H), \overrightarrow{E}(H))$ be an $\ell \times \ell$ directed torus where $V(H) = \{v_{i,j} \mid i, j \in [\ell]\}$. Let $T = (V, E)$ be a $2t\ell \times 2t\ell$ torus where $V = \{u_{i,j} \mid i, j \in [2t\ell]\}$.*

We define the t-tiling of *$\overrightarrow{H}$ as the orientation $\overrightarrow{H}^t$ of T which is constructed as follows. First, partition T into ℓ^2 disjoint $2t \times 2t$ grids $\{G_{i,j}\}_{i,j \in [\ell]}$, where every grid $G_{i,j}$ is associated with the vertex $v_{i,j} \in V(H)$. Formally, For every $i, j \in [\ell]$ the grid $G_{i,j}$ is the induced subgraph of T whose set of vertices is $V_{i,j} \stackrel{\text{def}}{=} \{u_{i',j'} : 2t(i-1) < i' \leq 2ti,\ 2t(j-1) < j' \leq 2tj\}$. The upper left $t \times t$ grid of every $G_{i,j}$ is denoted by $R_{i,j}$ and is called the* representative grid *of the vertex $v_{i,j} \in V(H)$.*

The orientation $\overrightarrow{H}^t$ of T is defined as follows. For every $v_{i,j} \in V(H)$, let $r_{i,j}^1, r_{i,j}^2, \ldots, r_{i,j}^t \in V$ be the t vertices on the main diagonal of the representative grid $R_{i,j}$. For every edge $e = \{v_{i,j}, v_{i',j'}\} \in$

Figure 5: A directed 2×2 torus $\overrightarrow{H}$ (left) and its corresponding 3-tiling, $\overrightarrow{T}$. *The vertex* $v_{2,1}$ *is encircled in* $\overrightarrow{H}$, *and its corresponding vertices* $r^1_{2,1}, r^2_{2,1}$ *and* $r^3_{2,1}$ *are encircled in* $\overrightarrow{T}$. *In addition, the edge* $\{v_{1,2}, v_{1,1}\}$ *is emphasized in* $\overrightarrow{H}$, *and its corresponding edges are emphasized in* $\overrightarrow{T}$. *Note the circular orientation of the padding edges in* $\overrightarrow{T}$, *marked with dashed arrows.*



$\overrightarrow{E}(H)$ *directed from* $v_{i,j}$ *to* $v_{i',j'}$ *and every* $h \in [t]$, *we orient the edges on the shortest path from* $r^h_{i,j}$ *to* $r^h_{i',j'}$ *in a way that forms a directed path from* $r^h_{i,j}$ *to* $r^h_{i',j'}$. *For every edge* $e' \in E$ *that participates in this path, we call* e *the* originating edge *of* e', *and use the notation* $\mathrm{org}(e') \stackrel{\text{def}}{=} e$. *The edges* e' *of* T *originated in this manner are called* representative edges, *whereas the remaining edges are called* padding edges. *Next, all the horizontal padding edges are directed to the right, and all the vertical padding edges are directed upwards (see Definition 9.2). See an example in Figure 5. For every padding edge* e *we define* $\mathrm{org}(e) \stackrel{\text{def}}{=} \emptyset$, *since they have no origin in* H.

The next lemma states that a tiling of an Eulerian torus is also Eulerian, while on the other hand, a tiling of a torus with many unbalanced vertices results with a torus that is far from being Eulerian.

Lemma 9.8 *Let* $\overrightarrow{H} = (V(H), \overrightarrow{E}(H))$ *be a directed* $\ell \times \ell$ *torus and let* $\overrightarrow{H}^t = (V, \overrightarrow{E})$ *be the* t*-tiling of* $\overrightarrow{H}$ *for some natural number* t. *Then,*

- *If* $\overrightarrow{H}$ *is Eulerian, then* $\overrightarrow{H}^t$ *is also Eulerian.*
- *For every* $0 < \delta < 1$, *if* $\overrightarrow{H}$ *contains* $\delta\ell^2$ *unbalanced vertices, then* $\overrightarrow{H}^t$ *is* $\frac{\delta}{16}$*-far from being Eulerian.*

Proof. The first statement of the lemma follows easily from Definition 9.7. Assume now that $\overrightarrow{H}$ has $\delta\ell^2$ unbalanced (spring or drain) vertices. According to the definition of a t-tiling, for every unbalanced vertex $v_{i,j} \in V(H)$ we have exactly t unbalanced vertices $r^1_{i,j}, r^2_{i,j}, \ldots, r^t_{i,j}$ on the main diagonal of $v_{i,j}$'s representative grid $R_{i,j}$ in $\overrightarrow{H}^t$, so the number of unbalanced vertices in $\overrightarrow{H}^t$ is $\delta\ell^2 t$. In addition, whenever $v_{i,j}$ is a spring (respectively drain) vertex in $\overrightarrow{H}$, the vertices $r^1_{i,j}, r^2_{i,j}, \ldots, r^t_{i,j}$

are also springs (respectively drains) in $\overrightarrow{H}^t$, so every pair of spring-drain vertices must reside in different grids $R_{i,j}$ and $R_{i',j'}$. This implies that (due to the orientation of the padding edges) the distance from any spring vertex to any drain vertex in $\overrightarrow{H}^t$ is at least t. Consequently, every correction path in $\overrightarrow{H}^t$ must be of length at least t. Since every correction path in $\overrightarrow{H}^t$ can balance at most two unbalanced vertices, and since the length of every such path is at least t, we conclude that $\overrightarrow{H}^t$ is $\frac{t\delta\ell^2 t/2}{|E(\overrightarrow{H}^t)|} = \frac{\delta\ell^2 t^2/2}{8\ell^2 t^2} = \frac{\delta}{16}$-far from being Eulerian. ■

Lemma 9.9 *Let $\overrightarrow{H}^t$ be a t-tiling of a randomly oriented $\ell \times \ell$ torus $\overrightarrow{H} \sim \mathcal{R}_\ell$. Then with probability $1 - o(1)$, $\overrightarrow{H}^t$ is $1/64$-far from being Eulerian.*

Proof. Follows by combining Lemma 9.8 (with $\delta = 1/4$) and Lemma 9.4. ■

We are now ready to define the distributions $\mathcal{P}_\ell$ and $\mathcal{F}_\ell$ over the orientations of an $\ell \times \ell$ torus $T = (V, E)$. To avoid divisibility concerns, we assume that $\ell = 2^k$ and $k = 2^b$ for some natural number $b > 1$. It is easy to verify that the same proof works also for general values of ℓ and k by using rounding as appropriate.

Distribution $\mathcal{P}_\ell$: Choosing $\overrightarrow{T} \sim \mathcal{P}_\ell$ is done according to the following steps.

- Choose s uniformly at random from the range $[k/4, k/2]$. Let $t = 2^s$, that is, t can take $\frac{\log \ell}{4}$ values in the range $[\ell^{1/4}, \ell^{1/2}]$.
- For an $\frac{\ell}{2t} \times \frac{\ell}{2t}$ torus H, choose a random orientation $\overrightarrow{H}$ of H according to the distribution $\mathcal{C}^{(k)}_{\ell/2t}$.
- Set $\overrightarrow{T}'$ to be the t-tiling of $\overrightarrow{H}$.
- Choose $a, b \in [\ell]$ uniformly at random, and set $\overrightarrow{T}$ to be the (a, b)-shifting of $\overrightarrow{T}'$ (see Definition 9.3).

Distribution $\mathcal{F}_\ell$: Choosing $\overrightarrow{T} \sim \mathcal{F}_\ell$ is done according to the following steps.

- Choose s uniformly at random from the range $[k/4, k/2]$ and set $t = 2^s$.
- For an $\frac{\ell}{2t} \times \frac{\ell}{2t}$ torus H, choose a random orientation $\overrightarrow{H}$ of H according to the distribution $\mathcal{R}_{\ell/2t}$.
- Set $\overrightarrow{T}'$ to be the t-tiling of $\overrightarrow{H}$.
- Choose $a, b \in [\ell]$ uniformly at random, and set $\overrightarrow{T}$ to be the (a, b)-shifting of $\overrightarrow{T}'$.

9.1.4 Bounding the variation distance

Let $T = (V, E)$ be an $\ell \times \ell$ torus. According to Lemma 9.8, every orientation $\overrightarrow{T} \sim \mathcal{P}_\ell$ of T is Eulerian. According to Lemma 9.9, an orientation $\overrightarrow{T} \sim \mathcal{F}_\ell$ of T is $\frac{1}{64}$-far from being Eulerian with high probability. Our aim is to show that any non-adaptive deterministic algorithm that makes

$o\left(\sqrt{\frac{\log \ell}{\log\log \ell}}\right)$ queries will fail to distinguish between the orientations that are distributed according to $\mathcal{P}_\ell$ and those that are distributed according to $\mathcal{F}_\ell$.

Let $Q \subseteq E$ be a fixed set of at most $\frac{1}{10}\sqrt{\frac{\log \ell}{\log\log \ell}}$ edges queried by a non-adaptive deterministic algorithm. Let $\overrightarrow{H} = (V(H), \overrightarrow{E}(H))$ be the $\frac{\ell}{2t} \times \frac{\ell}{2t}$ torus (oriented according to either $\mathcal{C}^{(k)}_{\ell/2t}$ or $\mathcal{R}_{\ell/2t}$) that has been used to create an orientation $\overrightarrow{T}$ of T. By Definition 9.7, the orientations of the padding edges of $\overrightarrow{T}$ are identical in $\mathcal{P}_\ell$ and $\mathcal{F}_\ell$, and the orientations of the other edges are determined by those of their originating edges. We thus focus on the set $\mathrm{org}(Q) \stackrel{\mathrm{def}}{=} \{\mathrm{org}(e) : e \in Q\} \subseteq E(H)$ of Q's originating edges.

If $\overrightarrow{T}$ is distributed according to $\mathcal{F}_\ell$, then the set $\mathrm{org}(Q) \subseteq E(H)$ is independent in H, and hence the distribution of the orientations of the edges in $\mathrm{org}(Q) \subseteq E(H)$ is uniform. We henceforth assume that $\overrightarrow{T}$ is distributed according to $\mathcal{P}_\ell$. In the next lemmas we show that, with probability at least $4/5$, the set $\mathrm{org}(Q)$ is independent in H also in this case, and thus our algorithm cannot distinguish between the two distributions.

Lemma 9.10 *Let $e_1, e_2 \in E$ be two edges within distance x. Let $\overrightarrow{T}$ be a random orientation of T chosen according to the distribution $\mathcal{P}_\ell$. Then*

$$\Pr_s\left[\frac{t}{4k} \le x \le 4tk\right] < \frac{8(\log\log \ell + 2)}{\log \ell}$$

where $t = 2^s$.

Proof. Observe that there are at most $2\log k + 4 = 2\log\log \ell + 4$ values of s for which a fixed number x can satisfy $2^{s-\log k-2} = \frac{t}{4k} \le x \le 4tk = 2^{s+\log k+2}$. Moreover, s is distributed uniformly among its $\frac{k}{4} = \frac{\log \ell}{4}$ possible values, so the lemma follows. ■

For a set $Q \subseteq E$ and an orientation $\overrightarrow{T}$ of T, let I_Q be the following event: For all pairs $e_1, e_2 \in Q$, either $\mathrm{dist}(e_1, e_2) < \frac{t}{4k}$ or $\mathrm{dist}(e_1, e_2) > 4tk$ in T.

Lemma 9.11 *Let $Q \subseteq E$ be a fixed set of $\frac{1}{10}\sqrt{\frac{\log \ell}{\log\log \ell}}$ edges, and let $\overrightarrow{T}$ be a random orientation of T, according to distribution $\mathcal{P}_\ell$. Then the event I_Q occurs with probability at least $9/10$.*

Proof. Follows by applying the union bound on the inequality from Lemma 9.10 for all pairs $e_1, e_2 \in Q$. ■

Lemma 9.12 *With probability $1 - o(1)$ conditioned on the event I_Q, every two edges $e_1, e_2 \in Q$ such that $\mathrm{dist}(e_1, e_2) < \frac{t}{4k}$ satisfy one of the following: (1) e_1, e_2 are perpendicular; (2) at least one of e_1, e_2 has no origin in H; (3) $\mathrm{org}(e_1) = \mathrm{org}(e_2)$.*

Proof. Fix two edges $e_1, e_2 \in Q$ such that $\mathrm{dist}(e_1, e_2) < \frac{t}{4k}$. If one of them has no originating edge in H then we are done. Otherwise, since $\mathrm{dist}(e_1, e_2) < \frac{t}{4k}$, by the definition of the t-tiling, $\mathrm{org}(e_1)$ and $\mathrm{org}(e_2)$ must have a common endpoint in H, say $v_{i,j}$. If e_1 and e_2 are perpendicular, then again we are done. On the other hand, if e_1 and e_2 are parallel, then in order to have different origins they must be separated by the the main diagonal of $R_{i,j}$ (the representative grid of the common vertex $v_{i,j}$). Note that this may happen only if the distance of both e_1 and e_2 from the main diagonal is at most $\frac{t}{4k}$. But the probability that an edge is within that distance from the main diagonal of some representative grid is at most $2\ \frac{t}{4k}\ \frac{t}{t^2} = \frac{1}{2k} = o\left(\frac{\log\log\ell}{\log\ell}\right)$. Now the proof is completed by applying the union bound for all pairs $e_1, e_2 \in Q$. ■

Lemma 9.13 *Let $Q \subseteq E$ be a fixed set of $\frac{1}{10}\sqrt{\frac{\log\ell}{\log\log\ell}}$ edges, and let $\overrightarrow{T}$ be a random orientation of T, chosen according to the distribution $\mathcal{P}_\ell$. Then* $\mathrm{org}(Q)$ *is independent with probability at least* $4/5$.

Proof. By Lemma 9.11, with probability at least $9/10$, the event I_Q happens, that is, for all pairs $e_1, e_2 \in Q$ we have either $\mathrm{dist}(e_1, e_2) < \frac{t}{4k}$ or $\mathrm{dist}(e_1, e_2) > 4tk$ in T. Assume that I_Q occurs. Then, by Lemma 9.12, with probability $1 - o(1)$, all the pairs $e_1, e_2 \in Q$ with $\mathrm{dist}(e_1, e_2) < \frac{t}{4k}$ in T are perpendicular or have no more than one originating edge. Conditioned on this event, every two edges in $\mathrm{org}(Q)$ are perpendicular or are at distance larger than $2k$ in H. Recall that $|Q| = o(\sqrt{\log\ell}) = o(\sqrt{k})$ and hence $|\mathrm{org}(Q)| = o(\sqrt{k})$. Therefore, from Lemma 9.6, $\mathrm{org}(Q)$ is independent in this case with probability $1 - o(1)$.

Summing up, we have that $\mathrm{org}(Q)$ is independent with probability at least $\frac{9}{10} - o(1) > 4/5$ for ℓ large enough, where the probabilities are taken over $\mathcal{P}_\ell$. ■

Proof of Theorem 9.1. Let $Q \subseteq E$ be the fixed set of $\frac{1}{10}\sqrt{\frac{\log\ell}{\log\log\ell}}$ edges queried by a deterministic non-adaptive algorithm. For every fixed t, a, and b, let $\mathcal{P}_\ell^{t,a,b}$ be $\mathcal{P}_\ell$ conditioned on t, a, and b and let $\mathcal{F}_\ell^{t,a,b}$ be $\mathcal{F}_\ell$ conditioned on t, a, and b. Note that t, a, and b fully define the set $\mathrm{org}(Q)$ of originating edges and, in particular, it is the same set for orientations drawn according to $\mathcal{P}_\ell^{t,a,b}$ and according to $\mathcal{F}_\ell^{t,a,b}$. It follows that, for every t, a, and b, if $\mathrm{org}(Q)$ is independent then the restriction of $\mathcal{P}_\ell^{t,a,b}$ to Q is identical to that of $\mathcal{F}_\ell^{t,a,b}$. Now, recall that for every t, a, and b, if $\overrightarrow{T}$ is distributed according to $\mathcal{F}_\ell^{t,a,b}$ then $\mathrm{org}(Q)$ is independent. On the other hand, by Lemma 9.13, $\mathrm{org}(Q)$ is independent with probability at least $4/5$ also for $\mathcal{P}_\ell^{t,a,b}$, taken over the choice of t, a, and b. Summing over all the possible choices of t, a, and b, we obtain that the variation distance between the restriction of $\mathcal{P}_\ell^{t,a,b}$ to Q and to that of $\mathcal{F}_\ell^{t,a,b}$ is at most $1/5$. Hence, distinguishing between the two distributions with probability larger than $1/5$ requires more than $\frac{1}{10}\sqrt{\frac{\log\ell}{\log\log\ell}} = \Omega\left(\sqrt{\frac{\log m}{\log\log m}}\right)$ queries. ■

9.2 A 1-sided lower bound

In this subsection we prove the following theorem.

Theorem 9.14 *For every $0 < \epsilon \leq 1/16$, every non-adaptive 1-sided ϵ-test for Eulerian orientations of bounded degree graphs must use at least $\frac{1}{100}m^{1/4}$ queries. Consequently, every adaptive 1-sided test requires $\Omega(\log m)$ queries.*

As opposed to 2-sided testers, a 1-sided tester is not allowed to reject the input unless a negative witness was found. In our case, as claimed in Lemma 3.1, the only possible witness that an orientation is not Eulerian is an invalid cut, i.e. a (possibly partial) cut that cannot be made balanced under any orientation of the non-queried edges.

Following this observation, we prove Theorem 9.14 using the distribution $\mathcal{F}_\ell$ defined in Subsection 9.1.3. First, we define a distribution $\mathcal{F}'_\ell$ that is similar to the distribution $\mathcal{F}_\ell$, except that t is fixed to be $\ell/16$, and the orientation $\overrightarrow{H}$ of an 8×8 torus H is fixed to be one that makes all 64 vertices fully unbalanced. Then we show that for orientations that are distributed according to $\mathcal{F}'_\ell$, any non-adaptive deterministic algorithm that makes $o(\sqrt{\ell}) = o(\sqrt{m})$ orientation queries cannot find an invalid cut (a negative witness) with probability larger than $1/5$. This will imply that there exists an orientation that is $\frac{1}{16}$-far from being Eulerian on which any randomized tester fails with probability at least $4/5$.

The main idea is as follows. A cut can be invalid (and hence unbalanced) only if both its components contain unbalanced vertices. Let us now fix a cut (A, B) of an $\ell \times \ell$ torus $T = (V, E)$, and let $\overrightarrow{T}$ be an orientation of T chosen according to $\mathcal{F}'_\ell$. Suppose that indeed both A and B contain unbalanced vertices, and let Q be a subset of the edges in the cut (A, B) that witness its invalidity. Using basic properties of tori, we show that either Q contains $\Omega(m^{1/4})$ edges, or otherwise, one of the edges $e \in Q$ must be within distance $O(m^{1/4})$ from one of the unbalanced vertices of $\overrightarrow{T}$. Since the number of unbalanced vertices in $\overrightarrow{T} \sim \mathcal{F}'_\ell$ is $O(\ell) = o(m^{1/4})$, and since they are grouped into 64 diagonals of length $\ell/32$, the number of edges that are within distance $O(m^{1/4})$ from these unbalanced vertices is bounded by $O(m^{3/4})$. Finally, since the last step in building $\overrightarrow{T}$ is a random shift, the probability that a set Q of size $o(m^{1/4})$ contains any such edge tends to zero.

We first give a formal definition of the distribution $\mathcal{F}'_\ell$ of orientations over a torus.

Distribution $\mathcal{F}'_\ell$: Choosing $\overrightarrow{T} \sim \mathcal{F}'_\ell$ is done according to the following steps.

- Set $t = \ell/16$.
- Fix the orientation $\overrightarrow{H}$ of the $\frac{\ell}{2t} \times \frac{\ell}{2t} = 8 \times 8$ torus H, such that all 64 vertices of H are fully unbalanced in $\overrightarrow{H}$ (i.e. no vertex has both incoming and outgoing edges).
- Set $\overrightarrow{T}'$ to be the t-tiling of $\overrightarrow{H}$.
- Pick $a, b \in [\ell]$ uniformly at random and set $\overrightarrow{T}$ to be the (a, b)-shifting of $\overrightarrow{T}'$.

Lemma 9.15 *Let $T = (V, E)$ be an $\ell \times \ell$ torus and let $\overrightarrow{T}$ be an orientation of T distributed according to $\mathcal{F}'_\ell$. Let Q be a fixed set of $\frac{1}{100} m^{1/4}$ edges from E. Then the probability (over $\overrightarrow{T} \sim \mathcal{F}'_\ell$) that any of the edges in Q is within distance at most $\sqrt{\ell}$ from a vertex $v \in V$ that is unbalanced in $\overrightarrow{T}$ is at most $1/5$.*

Proof. Let U denote the set of unbalanced vertices in $\overrightarrow{T} \sim \mathcal{F}'_\ell$. Observe that $|U| = 64t = 4\ell = 4\sqrt{m/2}$, and recall that the vertices in $U \subseteq V$ are grouped into 64 diagonals of length t (see Definition 9.7). Thus, the number of vertices $v \in V$ that are within distance at most $\sqrt{\ell}$ from some vertex $u \in U$ is bounded by $64 \cdot (t + 2\sqrt{\ell}) \cdot 2\sqrt{\ell} \leq 10\ell^{3/2}$. Hence, the probability of a single edge $e \in Q$ satisfying $\mathrm{dist}(e, u) \leq \sqrt{\ell}$ for some $u \in U$ is bounded by $20m^{-1/4}$ and the lemma follows. ■

We establish the proof of Theorem 9.14 using a few lemmas, in which we point out some significant properties of the torus. But first, we give a general lemma about witnesses for not being Eulerian.

Lemma 9.16 *Let $G = (V, E)$ be a graph and let $\overrightarrow{G} = (V, \overrightarrow{E})$ be an orientation of G. If a set $Q \subseteq E$ is a witness that $\overrightarrow{G}$ is not Eulerian then Q contains more than half of the edges of some invalid cut (A, B) in $\overrightarrow{G}$, where both A and B are connected sets of vertices.*

Proof. Recall that, by Lemma 3.1, Q contains more than half of the edges of an invalid cut, say (A', B'). Without loss of generality we assume that $|\overrightarrow{E}(A', B')| > \frac{1}{2}|E(A', B')|$. Hence, Q contains more than $\frac{1}{2}|E(A', B')|$ edges going from A' to B'. Let $A_1, \ldots, A_r$ be the connected components of A'. Note that (A', B') is a disjoint union of $(A_1, B'), \ldots, (A_k, B')$. Using averaging calculations, we obtain that there exists a connected component A_i such that Q contains more than $\frac{1}{2}|E(A_i, B')|$ edges going from A_i to B'. Note in addition that there are no edges between A_i and other connected components A_k's of A', and thus $|E(A_i, B')| = |E(A_i, V \setminus A_i)|$. We conclude that Q contains more than $\frac{1}{2}|E(A_i, V \setminus A_i)|$ edges going from A_i to $V \setminus A_i$. Now, let $B_1, \ldots, B_s$ be the connected components of $V \setminus A_i$. Note that $(A_i, V \setminus A_i)$ is a disjoint union of $(A_i, B_1), \ldots, (A_i, B_s)$. Using averaging calculations, we obtain that there exists a connected component B_j such that Q contains more than $\frac{1}{2}|E(A_i, B_j)|$ edges going from A_i to B_j. Note in addition that there are no edges between B_j and other connected components B_k's of $V \setminus A_i$, and therefore $|E(A_i, B_j)| = |E(V \setminus B_j, B_j)|$. We conclude that Q contains more than $\frac{1}{2}|E(V \setminus B_j, B_j)|$ edges going from $V \setminus B_j$ to B_j, and hence, Q is a witness to the invalidity of $(V \setminus B_j, B_j)$. B_j is clearly connected. To complete the proof, we need to show that $V \setminus B_j$ is connected as well. Recall that $V \setminus B_j$ is a union of A_i and all the connected components B_k's of $V \setminus A_i$ for $k \neq j$. The B_k's are not connected to each other. However, the torus T is a connected graph, and therefore, every B_k must be connected to A_i. Since A_i is connected, $V \setminus B_j$ is connected. To conclude, we set $A = V \setminus B_j$ and $B = B_j$. ■

In the following, we let $T = (V, E)$ be an $\ell \times \ell$ torus and use the notation of Definition 9.2. For every $i \in [\ell]$, define the i^{th} *row* of T as $R_i \stackrel{\text{def}}{=} \{v_{i,j} \in V \mid j \in [\ell]\}$. For every $j \in [\ell]$, define the j^{th} *column* of T as $C_j \stackrel{\text{def}}{=} \{v_{i,j} \in V \mid i \in [\ell]\}$. Given a set $A \subseteq V$, let $R(A)$ be the set of rows R_i of T such that $A \cap R_i \neq \emptyset$, and let $C(A)$ be the set of columns C_j of T such that $A \cap C_j \neq \emptyset$.

Given a cut (A, B) of V we say that a row R_i is *mixed* if $R_i \subseteq R(A) \cap R(B)$, that is, if R_i includes vertices in A as well as vertices in B. Similarly, we say that a column C_j is *mixed* if $C_j \subseteq C(A) \cap C(B)$. Let r_{mix} be the number of mixed rows with respect to (A, B) and let c_{mix} be the number of mixed columns with respect to (A, B).

Observation 9.17 $|E(A, B)| \geq 2(r_{mix} + c_{mix})$.

Proof. Looking at the cycle of vertical edges connecting all the vertices in every mixed column, it is easy to see that every mixed column has at least two vertical edges in (A, B). Similarly, it can be shown that every mixed row has at least two horizontal edges in (A, B). ■

Observation 9.18

1. *If $|R(A)| < \ell$ then $c_{mix} = |C(A)|$.*
2. *If $|C(A)| < \ell$ then $r_{mix} = |R(A)|$.*

The analogous claims also hold for B.

Proof. We give the proof of the first item as the proof of the second item is identical. Let R_i be a row of T that is not in $R(A)$. Then $v_{i,j} \in B$ for every $j \in [\ell]$. Hence, every column $C_j \in C(A)$ has a vertex in A as well as a vertex in B (namely, $v_{i,j}$), which proves the claim. ∎

Observation 9.19

1. *If* $|R(A)| = \ell$ *then* $r_{mix} = |R(B)|$.
2. *If* $|C(A)| = \ell$ *then* $c_{mix} = |C(B)|$.

Proof. We give the proof of the first item as the proof of the second item is identical. Suppose that $|R(A)| = \ell$. Then every row includes a vertex in A. Let $R_i \in R(B)$. Then R_i includes a vertex in A as well as a vertex in B, which completes the proof. ∎

We say that a set $A \subseteq V$ of vertices in T is *grid-bounded* if $|R(A)| < \ell$ and $|C(A)| < \ell$.

Lemma 9.20 *Let* (A, B) *be a cut of* V *where* $|E(A, B)| < 2\ell$. *Then at least one of* A *and* B *is grid-bounded.*

Proof. From Observation 9.17 we have that $r_{mix} + c_{mix} < \ell$. Note that $|R(A)| + |R(B)| - r_{mix} = \ell$ and $|C(A)| + |C(B)| - c_{mix} = \ell$. Hence $|R(A)| + |C(A)| + |R(B)| + |C(B)| \leq 2\ell + r_{mix} + c_{mix} < 3\ell$, and thus, at most two of the sets $R(A), C(A), R(B), C(B)$ are of size ℓ. Assuming that both A and B are not grid-bounded, we have $\max(|R(A)|, |C(A)|) = \ell$ and $\max(|R(B)|, |C(B)|) = \ell$. Therefore, we must have $|R(A)| = \ell$ and $|C(A)| < \ell$ or $|R(A)| < \ell$ and $|C(A)| = \ell$. We complete the proof for the case where $|R(A)| = \ell$ and $|C(A)| < \ell$ as the proof for the other case is identical. From Observation 9.19 we have $|R(B)| = r_{mix} < \ell$, and thus, from Observation 9.18, $c_{mix} = |C(B)|$. Now, $|C(B)| = \ell$, as otherwise B is grid-bounded. We hence obtain $c_{mix} = \ell$, a contradiction. ∎

Observation 9.21 *If* A *is connected then there exists a row index* $i^* \in [\ell]$ *such that* $R(A) = \{R_{i^*}, R_{i^*\oplus 1}, \ldots, R_{i\oplus(s-1)}\}$ *where* $s = |R(A)|$, *and there exists a column index* $j^* \in [\ell]$ *such that* $C(A) = \{C_{j^*}, C_{j^*\oplus 1}, \ldots, C_{j\oplus(t-1)}\}$ *where* $t = |R(B)|$. *Hence,* A *is contained in a subgraph* G *of* T *which is an* $|R(A)| \times |R(B)|$ *grid. Renaming* i^* *as* 1 *and* j^* *to* 1, *we have that* G *is a grid with the vertex set* $V_G = \{v_{i,j} \mid i \in [s], j \in [t]\}$.

Proof. Let R_{i_1}, R_{i_2} be rows in $R(A)$. Hence, both R_{i_1} and R_{i_2} include at least one vertex in A. Since A is connected, there exists a path of vertices in A between R_{i_1} and R_{i_2}. Clearly, for every edge in the path, the endpoints are in the same row (in case of a horizontal edge) or in subsequent rows (in case of a vertical edge). We thus conclude that $R(A)$ is a set of successive rows in the torus. Similarly, $C(A)$ is a set of successive columns in the torus. ∎

Lemma 9.22 *Let* $T = (V, E)$ *be an* $\ell \times \ell$ *torus, and let* $\overrightarrow{T}$ *be a non-Eulerian orientation of* T. *Let* $U \subseteq V$ *denote the set of unbalanced vertices with respect to* $\overrightarrow{T}$. *Let* $Q \subseteq E$ *be a set of edges forming a witness that* $\overrightarrow{T}$ *is not Eulerian, where* $|Q| < \frac{1}{2}\ell$. *Let* q *denote the minimal distance of an edge in* Q *to an unbalanced vertex, that is,* $q \stackrel{\text{def}}{=} \min_{e\in Q,\ u\in U}\{\text{dist}(e, u)\}$. *Then* $|Q| \geq q$.

Proof. By Lemma 9.16, we may assume without loss of generality that Q contains more than half of the edges of an invalid cut (A, B), where both A and B are connected. Since $|Q| < \frac{1}{2}\ell$, we have $|E(A, B)| < \ell$, and hence, from Lemma 9.20, one of the sets A and B is grid-bounded. Assume without loss of generality that A is grid-bounded. Let $s = |R(A)|$ and $t = |C(A)|$. Then $s, t < \ell$. Since A is connected, from Observation 9.21, A is contained in an $s \times t$ grid G.

Suppose that $|Q| < q$. Then $|E(A, B)| < 2q$, and from Observation 9.17 we have $r_{mix}+c_{mix} < q$. Let $e = (w_A, w_B)$ be an edge in $Q \cap E(A, B)$, where $w_A \in A$ and $w_B \in B$. Since A is invalid, there exists an unbalanced vertex $u \in A$. By the definition of q we have $\mathrm{dist}(e, u) \geq q$ and hence $\mathrm{dist}(u, w_A) \geq q$. Using the notation of Observation 9.21, we denote $u = v_{i_1,j_1}$ and $w_A = v_{i_2,j_2}$, where $i_1, i_2 \in [s]$ and $j_1, j_2 \in [t]$. Then clearly $|i_1 - i_2| + |j_1 - j_2| \geq q$. We thus have $s = |R(A)| \geq |i_1 - i_2| + 1$ and $t = |C(A)| \geq |j_1 - j_2| + 1$. Since A is grid-bounded, from Observation 9.18 we obtain $|C(A)| + |R(A)| = r_{mix} + c_{mix} \geq |i_1 - i_2| + 1 + |j_1 - j_2| + 1 > q$. Finally, Observation 9.17 gives that $|E(A, B)| > 2q$, a contradiction. ∎

Proof of Theorem 9.14. Let $T = (V, E)$ be an $\ell \times \ell$ torus and let $\overrightarrow{T} \sim \mathcal{F}'_\ell$ be an orientation of T. Let $Q \subseteq E$ be the fixed set of $\frac{1}{100}m^{1/4}$ edge queries that a deterministic non-adaptive algorithm makes on $\overrightarrow{T}$. By Lemma 9.22, in order to form a witness that $\overrightarrow{T}$ is not Eulerian, one of the edges in Q must be within distance at most $|Q| = \frac{1}{100}m^{1/4} < \sqrt{\ell}$ from an unbalanced vertex in $\overrightarrow{T}$. But according to Lemma 9.15, the probability of the above is at most $1/5$. We thus conclude that discovering a witness that $\overrightarrow{T} \sim \mathcal{F}'_\ell$ is not Eulerian with probability larger than $1/5$ requires more than $\frac{1}{100}m^{1/4}$ nonadaptive queries. ∎

10 Concluding comments and open problems

We have shown a test that has a sub-linear number of queries for all graphs. The test procedure is surprisingly involved considering the problem statement. However, (possibly excepting the special cases of dense graphs and expander graphs), this should be only considered as a first step for this problem, as many questions still remain open.

First, the question arises as to whether we can reduce the computational complexity of the test. While having a sub-linear number of queries, the exponential in m time complexity makes this test unrealistic for implementation. Also, to make the test truly attractive, not only its computation time needs to be polynomial in m, but most of the calculations should be performed in a preprocessing stage, where the amount of calculations done while making the queries should ideally be also sub-linear in m.

Related to the preprocessing question is the unresolved question of adaptivity. The current test is adaptive, but we would like to think that a sub-linear query complexity *non-adaptive* test also exists for the same class of graphs. Other adaptive versus non-adaptive gaps, such as the one concerning the 2-sided lower bounds, need also be addressed.

Acknowledgements

We thank the anonymous referees for their helpful comments.

References

[1] N. Alon and A. Shapira, Testing subgraphs in directed graphs, *J. Comput. Syst. Sci.* 69(3): 354–382, 2004 (a preliminary version appeared in Proc. of STOC 2003, 700-709).

[2] L. Babai, On the diameter of Eulerian orientations of graphs, *Proceedings of the* 17^{th} *SODA*, (2006), 822–831.

[3] M. Bender and D. Ron, Testing properties of directed graphs: Acyclicity and connectivity, *Random Structures and Algorithms (2002)*, 184–205.

[4] E. Ben-Sasson, P. Harsha and S. Raskhodnikova, Some 3CNF properties are hard to test, *SIAM Journal on Computing*:35(1), 1–21, 2005 (a preliminary version appeared in Proc. 35^{th} STOC, 2003).

[5] M. Blum, M. Luby and R. Rubinfeld, Self-testing/correcting with applications to numerical problems. *Journal of Computer and System Sciences* 47 (1993), 549–595 (a preliminary version appeared in Proc. 22^{nd} STOC, 1990).

[6] G. R. Brightwell and P. Winkler, Counting Eulerian circuits is #P-complete, *Proc. 7th ALENEX and 2nd ANALCO 2005 (Vancouver BC)*, C. Demetrescu, R. Sedgewick and R. Tamassia, eds, SIAM Press, 259–262.

[7] S. Chakraborty, E. Fischer, O. Lachish, A. Matsliah and I. Newman: Testing *st*-Connectivity, *Proceedings of the* 11^{th} *RANDOM and the* 10^{th} *APPROX (2007)*: 380–394.

[8] E. Fischer, The art of uninformed decisions: A primer to property testing, *Current Trends in Theoretical Computer Science: The Challenge of the New Century, G. Paun, G. Rozenberg and A. Salomaa (editors), World Scientific Publishing* (2004), Vol. I 229-264.

[9] E. Fischer and O. Yahalom, Testing convexity properties of tree colorings, *Proc. of the 24th International Symposium on Theoretical Aspects of Computer Science (STACS 2007)*, LNCS 4393, Springer-Verlag 2007, 109–120.

[10] H. Fleishcner, *Eulerian graphs and related topics*, Part 1. Vol. 1. Annals of Discrete Mathematics 45, 1990.

[11] H. Fleishcner, *Eulerian graphs and related topics*, Part 1. Vol. 2. Annals of Discrete Mathematics 50, 1991.

[12] O. Goldreich, S. Goldwasser and D. Ron, Property testing and its connection to learning and approximation, JACM 45(4): 653-750 (1998).

[13] O. Goldreich and D. Ron, Property testing in bounded degree graphs, *Algorithmica*, (2002), 32(2):302–343

[14] S. Halevy, O. Lachish, I. Newman and D. Tsur, Testing properties of constraint-graphs, *Proceedings of the* 22^{nd} *IEEE Annual Conference on Computational Complexity (CCC 2007)*, 264–277.

[15] S. Halevy, O. Lachish, I. Newman and D. Tsur, Testing orientation properties, technical report, *Electronic Colloquium on Computational Complexity (ECCC)*, (2005), 153.

[16] T. Ibaraki, A. V. Karzanov and H. Nagamochi, A fast algorithm for finding a maximum free multiflow in an inner Eulerian network and some generalizations, *Combinatorica 18(1)* (1988), 61–83.

[17] T. Kaufman, M. Krivelevich and D. Ron, Tight bounds for testing bipartiteness in general graphs, *SICOMP*, (2004), 33(6):1441–1483.

[18] L. Lovász, On some connectivity properties of Eulerian graphs, *Acta Math. Hung. 28* (1976) 129–138.

[19] M. Mihail, P. Winkler, On the number of Eulerian orientations of a graph, *Algorithmica 16(4/5)* (1996), 402–414.

[20] I. Newman, Testing of Function that have small width Branching Programs, *SIAM Journal on Computing*, 31(5):1557–1570, 2002 (a preliminary version appeared in Proc. 41^{st} FOCS, 2000).

[21] M. Parnas and D. Ron, Testing the diameter of graphs, *Random Struct. and Algorithms*, (2002), 20(2):165–183.

[22] P. A. Pevzner, H. Tang and M. S. Waterman, An Eulerian path approach to DNA fragment assembly, *Proc. Natl. Acad. Sci. USA 98*, (2001), 9748–9753.

[23] R. W, Robinson, Enumeration of Euler graphs, In *Proof Techniques in Graph Theory* (Ed. F. Harary), New York: Academic Press, 1969, 147–153.

[24] D. Ron, Property testing (a tutorial), In: *Handbook of Randomized Computing* (S. Rajasekaran, P. M. Pardalos, J. H. Reif and J. D. P. Rolim eds), Kluwer Press (2001), Vol. II Chapter 15.

[25] R. Rubinfeld and M. Sudan, Robust characterization of polynomials with applications to program testing, *SIAM Journal on Computing* 25 (1996), 252–271 (first appeared as a technical report, Cornell University, 1993).

[26] W. T. Tutte, *Graph theory*, Addison-Wesley, New York, 1984.

[27] A. C. Yao, Probabilistic computation, towards a unified measure of complexity, In *Proceedings of the* 18^{th} *IEEE FOCS*: 222–227, 1977.