3.3 Between-ness centrality

In this section we discuss an example of a centrality measure that is *not* based on walks. Given two distinct nodes i and j, a *shortest* path from i to j is a path from i to j of minimal length among all the possible paths. (Note that open paths must have length at least 1 and at most $n-1$, where $n = \#V$; why?). If k is a third vertex, distinct from i and j, a *shortest path through* k from i to j is a shortest path from i to j that goes through k. Between-ness centrality is defined on the vertex i as

$$b_i = \sum_{j \neq i \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}}$$

where σ_{jk} is the number of shortest paths from j to k and $\sigma_{jk}(i)$ is the number of shortest paths from j to k through i.

Between-ness centrality was first proposed in 1977 by L. Freeman. Although we will not study it in detail, it is interesting to compare its computational complexity with walk-based centralities, because between-ness centralities is among the best purely graph theoretical centrality measure in terms of computational costs. Indeed, the current best algorithm (by U. Brandes in 2001) can compute between-ness centrality in $O(\#V\#E)$ operation for a general graph. A typical real-life network is very sparse, with $\#E = O(\#V) = O(n)$, so this means an $O(n^2)$. However, in some special cases this reduces to $O(n)$ (for a tree) or $O(n^{3/2})$ (for a planar graph). It is interesting to compare these asymptotic costs with those of computing Katz centrality. Let us again assume that we are studying a typical real-life network and that the adjacency matrix has $O(n)$ nonzero elements; iterative solvers can be used that rely on Krylov subspace methods, which require to compute a matrix-vector multiplication at each iteration. This requires an amount of multiplications and additions of the same order of the nonzero elements in the matrix, i.e., $O(n)$ in the case of our thought experiment. The number of needed iterations for a good accuracy of the approximate d solution is typically much smaller than n. With a clever use of preconditioning techniques, it can sometimes be brought down even to $O(\log n)$, leading to $O(n \log n)$ cost.

For example, if we were to study the graph of Facebook users, than n is of the order of a few billions. A typical modern laptop computer can perform several billions of basic mathematical operations per second: therefore, (at least by using state-of-the art algorithms) Katz centrality can be computed

on your laptop in a few seconds using the method described above. On the other hand, computing between-ness centrality via the Brandes method takes (unless some special feature of the graph can be exploited) takes on the same latpop a few billions of seconds, that is, a few decades. It is thus remarkable how matrix theory based tools can make even such large graph computations feasible.

3.4 Nonbacktracking centralities

The topic of walk-based centrality is a classical one in network theory. The original paper by Katz appeared in 1953, and the developments and variants described in the previous section have been thoroughly studied through the subsequent decades. However, the Benzi-Klymko theorem described some concrete limits on how walk-based centrality measures can differ: all somehow interpolate between degree centrality and eigenvector centrality. For this reason, some authors (including your lecturer) have recently proposed that, to get more substantially different centralities, rather than changing the weights of the walks of length k one could exclude some types of walks, based on features other than length, and stil having in mind the basic principle underlying Katz that the less useful a walk is to explore the network the less weight it should have in the computation of a centrality measure. In this section, we will describe centrality measure that are based on counting only a certain type of walks, called nonbacktracking. The resulting *nonbacktracking centrality* has first been proposed in 2018 [1] by P. Grindrod, D. Higham and V. Noferini.

3.4.1 Nonbacktracking Walks

The sum (1) includes some traversals that, intuitively, are less relevant than others. In particular, for every edge ij, (1) incorporates all walks that pass from i to j and immediately pass back to i, rather than exploring other parts of the network. Arguably, such traversals are less useful for the purpose of expoloring a graph, which is the intuition underlying walk-based centralities.

---

[1]Historically, note that this is in a post-Benzi-Klymko-theorem scenario. Indeed, we mention in passing the following result of Arrigo, Higham and Noferini (2020): variants of nonbacktracking centrality also follow a universality limiting behaviour, but the limit for high values of the parameter is *different* than the one described by the Benzi-Klymko theorem.

In the present subsection of the notes and in the ones following it, we are going to explore how to perform a more sophisticated counting of the walks, excluding the ones that behave as described above.

We start by giving a more precise definition.

Definition 3.2. *A* backtracking walk *is a walk that contains at least one node subsequence of the form uvu (a closed path of length 2), i.e., it visits u,v and then u in immediate succession.*

A nonbacktracking walk *is a walk that is not backtracking, i.e., it does not contain any subsequence of the form uvu.*

For brevity we will henceforth often replace the phrase nonbacktracking walk with NBTW. Our next idea is to refine Katz centrality by performing a weighted sum of NBTWs, as opposed to all possible walks.

Definition 3.3. *For an appropriate value of the real parameter $t > 0$, the NBTW centrality of node i is defined by*

$$1 + \sum_{j=1}^{n} \sum_{k=1}^{\infty} t^k \left(p_k(A) \right)_{ij},$$

where $\left(p_k(A) \right)_{ij}$ records the number of distinct NBTWs of length k from i to j.

In the case of Katz centrality, we showed that while the centrality was defined by a power series, for appropriate values of the Katz parameter it neatly converged to an expressions corresponding to the solution of a linear system. We do not (yet) know how to perform the same task, and indeed, while the matrices $p_k(A)$ are functions of the underlying graph (and hence of A), we cannot immediately express them as simply as A^k. Still, later on the course we will show how to compute NBTW centrality in terms of a certain matrix polynomial depending on the original adjacency matrix, A, and study the role of the parameter t. At this stage, we simply note that $0 < t < 1$ is a natural requirement, so that longer walks carry less weight, and we continue with an illustrative example that differentiates the new measure from Katz centrality.

Example 3.3. *Consider a star graph with n nodes, as illustrated for the case $n = 9$ in Figure 1. Here the central hub node has an undirected edge to each*

of the $n-1$ leaf nodes. The adjacency matrix has the form

$$A = \begin{bmatrix} & 1 & \cdots & 1 \\ 1 & & & \\ \vdots & & & \\ 1 & & & \end{bmatrix} \in \mathbb{R}^{n \times n}, \tag{5}$$

where a blank denotes a zero entry. The eigenvalues of A are $\pm\sqrt{n-1}$ and 0 (repeated $n-2$ times). Hence, Katz centrality is defined for $0 < \alpha < 1/\sqrt{n-1}$.

By symmetry the x_i values in the Katz system are equal for all $i \geq 2$, and the equations reduce to $x_1 - \alpha(n-1)x_2 = 1$ and $x_2 - \alpha x_1 = 1$. These solve to give

$$x_1 = \frac{1 + \alpha(n-1)}{1 - \alpha^2(n-1)} \quad \text{and} \quad x_i = \frac{1 + \alpha}{1 - \alpha^2(n-1)}, \quad \text{for } i \geq 2. \tag{6}$$

The ratio of hub centrality to leaf centrality is therefore, for $i \geq 2$,

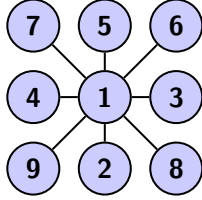$$\frac{x_1}{x_i} = \frac{1 + \alpha(n-1)}{1 + \alpha}. \tag{7}$$

Figure 1: A star graph with $n = 9$ vertices.

Turning to NBTWs, for the star graph it follows directly from Definition 3.2 that

- node 1 has $n-1$ NBTWs of length one and no NBTWs of length greater than one,

- node i for $i \geq 2$ has one NBTW of length one, $n-2$ NBTWs of length two, and no NBTWs of length greater than two.

Hence, in Definition 3.3 the NBTW centralities are

$$x_1 = 1 + (n-1)t \qquad \text{and} \qquad x_i = 1 + t + (n-2)t^2, \quad \text{for } i \geq 2. \qquad (8)$$

So the ratio of hub centrality to leaf centrality is, for $i \geq 2$,

$$\frac{x_1}{x_i} = \frac{1 + (n-1)t}{1 + t + (n-2)t^2}. \qquad (9)$$

We are interested in large systems, so consider the limit $n \to \infty$. In the Katz regime we require $\alpha < 1/\sqrt{n-1}$. If we take α to be a fixed proportion of this upper limit, say $0.9/\sqrt{n-1}$, then in (6) and (7) we have

$$x_1 = O(\sqrt{n}), \quad x_i = O(1), \quad x_1/x_i = O(\sqrt{n}). \qquad (10)$$

Similarly, for the NBTW version, using $t = 0.9/\sqrt{n-1}$ in (8) and (9) we obtain the same asymptotic behaviour. However, in this example the NBTW centrality measure is valid for any t. So we may consider the case where $t = O(1)$ as $n \to \infty$, e.g., $t = 1/2$, in which case

$$x_1 = O(n), \quad x_i = O(n), \quad x_1/x_i \to \frac{1}{t} = O(1). \qquad (11)$$

So, compared with Katz, the NBTW measure

- *has a much less severe restriction on the downweighting parameter, and*

- *for fixed t and large n, gives a less dramatic distinction between the hub and the leaves.*

3.4.2 Matrix Polynomials and the Deformed Graph Laplacian

We now provide some general background material on matrix polynomials before introducing and studying the deformed graph Laplacian. Recall that, given a field $\mathbb{F}$ (in these notes, $\mathbb{F}$ is either $\mathbb{R}$ or $\mathbb{C}$), the set of univariate polynomials in t with coefficients in $\mathbb{F}$ is denoted by $\mathbb{F}[t]$. Moreover, the set of square matrices of size n with entries in $\mathbb{F}[t]$ is denoted by $\mathbb{F}[t]^{n \times n}$.

For $j = 0, 1, \ldots, k$, let $A_j \in \mathbb{C}^{n \times n}$ be square matrices of the same size with $A_k \neq 0$. The matrix-valued function $P(t) = \sum_{j=0}^{k} A_j t^j \in \mathbb{C}[t]^{n \times n}$ is called a square matrix polynomial of degree k. If $\det P(t) \equiv 0$ then $P(t)$ is said to be *singular*, otherwise it is called *regular*. We now recall some basic definitions from the spectral theory of regular matrix polynomials.

31

The finite eigenvalues of a square regular matrix polynomial of degree k are the zeros of the scalar polynomial $\det P(t)$. Moreover, if $\deg \det P(t) < kn$, we say that ∞ is an eigenvalue of $P(t)$. If a finite eigenvalue has multiplicity 1 as a root of $\det P(t)$ it is called a *simple eigenvalue*; similarly, an infinite eigenvalue is simple if $\deg \det P(t) = kn - 1$. An eigenvalue which is not simple is called *multiple*. If $\lambda \in \mathbb{C}$ is a finite eigenvalue of $P(t)$, any nonzero vector $\mathbf{v} \in \mathbb{C}^n$ such that $P(\lambda)\mathbf{v} = \mathbf{0}$ is called an eigenvector associated with the eigenvalue λ. Similarly, if $P(t)$ has at least one infinite eigenvalue, then any nonzero vector $\mathbf{w} \in \mathbb{C}^n$ such that $A_k\mathbf{w} = \mathbf{0}$ is an eigenvector associated with the eigenvalue ∞.

The *algebraic multiplicity* of a finite (resp., infinite) eigenvalue of $P(t)$ is the multiplicity of the eigenvalue as a root of $\det P(t)$ (resp., the number $kn - \deg \det P(t)$). Moreover, we say that a finite eigenvalue $\lambda \in \mathbb{C}$ has *geometric multiplicity* $n - \text{rank} P(\lambda)$, and similarly the eigenvalue infinity has geometric multiplicity $n - \text{rank} A_k$. An eigenvalue has geometric multiplicity g if and only if one can find g linearly independent eigenvectors associated with it. If the algebraic and geometric multiplicities of an eigenvalue coincide, we say that the eigenvalue is *semisimple*; otherwise, it is *defective*.

It is easy to check that, by the definitions above, a regular matrix polynomial of size n and degree k has precisely kn eigenvalues, counted with their algebraic multiplicities and possibly including infinite eigenvalues. In this paper, we will focus on a real matrix polynomial, i.e., $A_i \in \mathbb{R}^{n \times n}$. Note that, even if $P(t) \in \mathbb{R}[t]^{n \times n}$, the variable $t \in \mathbb{C}$ is generally allowed to be complex, and a real matrix polynomial may have nonreal finite eigenvalues.

We now turn our attention to the *deformed graph Laplacian*: a special matrix polynomial associated with any graph. The deformed graph Laplacian has been studied in the engineering literature because of its applications to consensus algorithms in multi-agent systems and robotics. Here, we will analyze it more thoroughly using the spectral theory of matrix polynomials, and we will then focus on its connections to NBTW centrality.

Definition 3.4. *Let $A \in \mathbb{R}^{n \times n}$ be the adjacency matrix of an undirected graph. For any $t \in \mathbb{C}$, the associated* deformed graph Laplacian *is the Hermitian matrix polynomial*

$$M(t) = I - At + (\Delta - I)t^2 \in \mathbb{R}[t]^{n \times n}. \tag{12}$$

Observe that $M(1) = L$ is the graph Laplacian, $M(0) = I$ is the identity matrix, while $M(-1)$ is the so-called signless graph Laplacian.

3.4.3 Nonbacktracking Walk Centrality and the Deformed Graph Laplacian

In subsection 3.4.1 we gave a simple example where the NBTW centrality in Definition 3.3 could be computed from first principles. To obtain a general-purpose algorithm, we derive results that concern the combinatorics of NBTWs. Although originally derived from a pure mathematics viewpoint, these results turn out to be extremely useful from the perspective of matrix computations in network science, and they also highlight a connection between NBTWs and the deformed graph Laplacian. Lemma 3.4 gives a recurrence relation between NBTW counts of different lengths. Theorem 3.5 is an immediate corollary that gives an expression for the associated generating function.

Lemma 3.4. *Recall that Δ denotes the diagonal degree matrix and $p_r(A)$ has (i, j) element that counts the number of NBTWs of length r from i to j. Then $p_1(A) = A$, $p_2(A) = A^2 - \Delta$, and for $r > 2$*

$$Ap_{r-1}(A) = p_r(A) + (\Delta - I)p_{r-2}(A). \tag{13}$$

Proof. When $r = 1$, as all walks are nonbacktracking, the formula is easily verified. For $r = 2$, observe that the backtracking walks of length 2 are precisely the closed paths of length 2, and they correspond to the diagonal elements of A^2; this proves the formula.

For $r > 3$, observe first that

$$[Ap_{r-1}(A)]_{ij} = \sum_{k=1}^{n} A_{ik}[p_{r-1}(A)]_{kj}$$

is tantamount to the sum over all vertices k that are adjacent to node i, of the number of NBTWof length $r-1$ from node k to node j. This means that we are adding an additional first step to a NBTWof length $r-1$. There are two possibilities: either the resulting walk of length r is backtracking or not. As, by definition, the number of NBTWof length r from node i to node j, we conclude that the (i, j) element of $Ap_{r-1}(A) - p_r(A)$ counts the number of backtracking walks from i to j having the additional special property that they would become nonbacktracking if we dropped the first step.

It remains to count the latter walks. As they are backtracking in their first two steps, we note that they must begin as $iki \ldots j$, where k is a node

adjacent to i. Observe that there are precisely Δ_{ii} (the degree of node i) such nodes. However, not all of these nodes are allowed: indeed, let ℓ be the node (also adjacent to i) such that the considered walk goes on as $iki\ell \dots j$: if $k = \ell$, then this walk would remain backtracking even after dropping the first, since $\ell i\ell \dots j$ is backtracking. Hence, there are only $\Delta_{ii} - 1$ allowed choices for k.

Therefore, the number of backtracking walks from i to j that would become nonbacktracking if we dropped the first step have the form $iki\ell \dots j$, where $k \neq \ell$ and $i\ell \dots j$ is any NBTWof length $r - 2$ from $\ell to j$. can be computed as the product of $\Delta_{ii} - 1$ choices for k, the second vertex in the walk, times the number of NBTWof length $r - 2$ from ℓ to j, which is $p_{r-2}(A)_{\ell j}$. In formulae, since $\Delta - I$ is diagonal, this can equivalently be expressed as

$$\sum_{\ell=1}^{n} (\Delta - I)_{i\ell} p_{r-2}(A)_{\ell j} = [(\Delta - I)p_{r-2}(A)]_{ij};$$

this proves the statement. $\qquad\square$

Theorem 3.5. *Let $\Phi(A, t) := \sum_{r=0}^{\infty} p_r(A)t^r$, where, for convenience, we set $p_0(A) = I$, and recall that $M(t)$ denotes the deformed graph Laplacian associated with A. Suppose moreover that t is such that the power series converges. Then,*

$$M(t)\Phi(A, t) = (1 - t^2)I. \tag{14}$$

Before proving Theorem 3.5, we observe that it yields a computation method for NBTW centrality. By Definition 3.3 we see that the NBTW centrality x_i of node i may be computed via $\mathbf{x} = \Phi(A, t)\mathbf{e}$ where $\mathbf{e}$ is the vector of all ones. From Theorem 3.5 we see that this simplifies to the linear system

$$M(t)\mathbf{x} = (1 - t^2)\mathbf{e}. \tag{15}$$

Therefore, just as with Katz centrality, the (admittedly more complicated in this case) combinatorial analysis leads a method that involves the solution of a, possibly very sparse in practice, system of linear equations in order to compute NBTW centrality. Moreover, we note from (12) that, for any fixed value of t, $M(t)$ in (15) has the same sparsity structure as the coefficient matrix $I - \alpha A$ that appears in the original Katz system! Hence, NBTW centrality may be computed at the same cost as Katz centrality: a very good property in computational terms, and a somewhat suprising one given

the more complicated combinatorics involved for NBTWs. Indeed, there are even some subtleties involved in the that make the computation of NBTW centrality somewhat *more* efficient than Katz.

We next turn to a proof of Theorem 3.5.

Proof of Theorem 3.5. Since for all $r \geq 3$ $p_r(A) = Ap_{r-1}(A) + (I - \Delta)p_{r-2}(A)$, we have

$$\Phi(A, t) = I + At + (A^2 - \Delta)t^2 + tA\left(\sum_{r=3}^{\infty} t^{r-1} p_{r-1}(A)\right) + t^2(I - \Delta)\left(\sum_{r=3}^{\infty} t^{r-2} p_{r-2}(A)\right)$$

implying

$$\Phi(A, t) = I + At + (A^2 - \Delta)t^2 + tA(\Phi(A, t) - I - At) + t^2(I - \Delta)(\Phi(A, t) - I)$$

and hence

$$M(t)\Phi(A, t) = I + t(A - A) + t^2(A^2 - \Delta - A^2 - I + \Delta) = (1 - t^2)I,$$

proving the statement. ☐

3.4.4 Further Spectral Analysis of the Deformed Graph Laplacian

In this section, we discuss how the (finite and infinite) eigenvalues of the deformed graph Laplacian record features of the underlying graph. We will prove only some of the results we give. An individual study of all the proofs, together with additional results on this theme, is available as a possible student project.

Proposition 3.6 records some basic spectral properties of $M(t)$.

Proposition 3.6. *The following hold:*

1. *$M(t)$ is a regular matrix polynomial, and 0 is never an eigenvalue of $M(t)$;*

2. *1 is always an eigenvalue of $M(t)$, with geometric multiplicity equal to the number of connected components (or isolated nodes) of the graph of A;*

3. *the geometric multiplicity of ∞ as an eigenvalue of $M(t)$ is equal to the number of leaves, i.e., vertices of degree 1, in the graph of A (in particular, ∞ is an eigenvalue of $M(t)$ if and only if the graph of A has at least one leaf);*

4. −1 is an eigenvalue of $M(t)$ if and only if the graph of A has at least one bipartite component. In this case, the geometric multiplicity of −1 is equal to the number of bipartite components (including isolated nodes) of the graph of A.

Proof. 1. We have $\det M(0) = 1$, and therefore $\det(M(t))$ cannot be the zero polynomial; moreover, by the same argument, 0 is not an eigenvalue.

2. Observe that $M(1) = L$. As we know already, the nullity of L, and hence the geometric multiplicity of 1 as an eigenvalue of $M(t)$, is equal to the number of connected components plus the number of isolated nodes in the graph of A.

3. The geometric multiplicity of the infinite eigenvalue is the nullity of $\Delta - I$, which is equal to the number of leaves in the graph of A.

4. $M(t)$ has the eigenvalue $-1 \Leftrightarrow$ the signless graph Laplacian, defined as $A + \Delta$, is a singular matrix $\Leftrightarrow$ the graph of A has at least one bipartite component, where the last equivalence is a result by Cvektovic; moreover, again by a theorem by Cvektovic, the multiplicity of the eigenvalue 0 of the signless graph Laplacian, and hence the geometric multiplicity of -1 as an eigenvalue of $M(t)$, is equal to the number of bipartite components in the underlying graph (an isolated node is considered a bipartite component). □

The next proposition shows that, for disconnected graphs, it suffices to study the individual deformed graph Laplacians associated with each connected component (or isolated node).

Proposition 3.7. *Let $A \in \mathbb{R}^{n \times n}$ be the adjacency matrix of a disconnected graph having c connected components (or isolated nodes) with adjacency matrices A_i, for $i = 1, \ldots, c$; let $M(t)$ be the deformed graph Laplacian associated with A and $M_i(t)$ be the deformed graph Laplacians associated with A_i, for $i = 1, \ldots, c$; and let $\lambda \in \mathbb{C} \cup \{\infty\}$. Then, λ is an eigenvalue of $M(t)$ if and only if it is an eigenvalue of $M_i(t)$ for some value of $i = 1, \ldots, c$. Moreover, denote by $\gamma(\lambda)$ (resp. $\gamma_i(\lambda)$) the geometric multiplicity of λ as an*

eigenvalue of $M(t)$ (resp. $M_i(t)$). Similarly, let $\alpha(\lambda)$ and $\alpha_i(\lambda)$ denote the corresponding algebraic multiplicities. Then, it holds

$$\gamma(\lambda) = \sum_{i=1}^{c} \gamma_i(\lambda), \qquad \alpha(\lambda) = \sum_{i=1}^{c} \alpha_i(\lambda).$$

We omit a rigorous proof of Proposition 3.7, but we note that one can be obtained along the same lines of the argument we gave to show that the spectrum of A is, in the case of several connected components, just the union of the spectra of the adjacency matrices associated with each connected component (incljuding isolated nodes)

We now give (without proof) a powerful auxiliary result.

Theorem 3.8. *Let A be the adjacency matrix of a simple, undirected, connected graph. Denote by $\widetilde{A}$ the adjacency matrix, possibly of smaller size, such that the graph of $\widetilde{A}$ is obtained by removing from the graph of A all the leaves, if any, and the edges connecting these leaves to the rest of the graph. Suppose that the graph of $\widetilde{A}$ is not empty, i.e., it contains at least one node. Let $M(t)$, $\widetilde{M}(t)$ be the deformed graph Laplacians associated with $A, \widetilde{A}$ respectively.*

Then, $\lambda \in \mathbb{C}$ is a finite eigenvalue of $M(t)$ if and only if it is a finite eigenvalue of $\widetilde{M}(t)$. Moreover, the algebraic and geometric multiplicities of λ as an eigenvalue of $M(t)$ and $\widetilde{M}(t)$ are the same.

Theorem 3.8 says that to compute the finite eigenvalues of $M(t)$ we are allowed to remove all the leaves of the underlying graph, and iterate the process until we are left with a graph with no leaves. (As a consequence, if the underlying graph is a forest, then the only finite eigenvalues are ± 1, which must be both semisimple. This observation is recorded as Corollary 3.11 in subsection 3.4.5, with an alternative proof based on the connection with NBTWs.)

As our first application of Theorem 3.8, we show that the deformed graph Laplacian can never have finite eigenvalues of modulus larger than 1.

Theorem 3.9. *Let $M(t)$ be the deformed graph Laplacian associated with a simple undirected graph. Suppose that $\lambda \in \mathbb{C}$ is a finite eigenvalue of $M(t)$. Then, $|\lambda| \leq 1$.*

Proof. By Proposition 3.7 and Theorem 3.8, we may assume with no loss of generality that the graph of A is connected and that it does not have any leaves.

If $\lambda \in \mathbb{C}$ is a finite eigenvalue of $M(t)$, then there exists a nonzero $\mathbf{v} \in \mathbb{C}^n$ such that $M(\lambda)\mathbf{v} = 0$ holds. Without loss of generality we take $\|\mathbf{v}\|_2 = 1$. Premultiplying the eigenvalue eigenvector equation $M(\lambda)\mathbf{v} = 0$ by $\mathbf{v}^*$, we obtain

$$\alpha\lambda^2 - \beta\lambda + 1 = 0, \tag{16}$$

where $\alpha = \mathbf{v}^*\Delta\mathbf{v} - 1$ and $\beta = \mathbf{v}^*A\mathbf{v}$. Denoting the degree of the ith node by $\deg_i$, we have $\deg_i \geq 2$ for all i, and hence $\alpha = \sum_{i=1}^n \deg_i |v_i|^2 - 1 \geq 2 - 1 = 1$.

There are two cases. If $\lambda \notin \mathbb{R}$, then λ^* is also an eigenvalue of $M(t)$ and a root of (16). It follows that $1 \leq \alpha = |\lambda|^{-2} \Leftrightarrow |\lambda| \leq 1$. Suppose now $\lambda \in \mathbb{R}$. Using also the fact that $\Delta \pm A$ are both positive semidefinite matrices, which implies $|\beta| \leq \alpha + 1$, we have $0 \leq \beta^2 - 4\alpha \leq (\alpha - 1)^2$, and hence, $|\lambda| \leq (|\beta| + \sqrt{\beta^2 - 4\alpha})/(2\alpha) \leq (\alpha + 1 + \alpha - 1)/(2\alpha) = 1$. $\qquad\square$