Università di Pisa

Dipartimento di Informatica

Technical Report : TR-99-04

Matrix-Valued Linear Positive Operators and Applications to Graph Optimization

Antonio Frangioni, Capizzano Stefano Serra

March 08, 1999

ADDR: Corso Italia 40,56125 Pisa,Italy. TEL: +39-50-887111. FAX: +39-50-887226

Matrix-Valued Linear Positive Operators and Applications to Graph Optimization

Antonio Frangioni[*] Stefano Serra Capizzano[†]

March 8, 1999

Abstract

In this paper, we extend some general results on linear positive operators to the case of matrix-valued operators taking values in certain spaces of Hermitian matrices. These results are then used to show that certain matrices are optimal preconditioners for the linear systems arising in Interior Point methods for Linear Programs. In the case of graph matrices, corresponding to Min Cost Flow problems, we are able to find superlinear preconditioners for a class of "local" graphs generalizing the grid graphs commonly used in applications. For general graphs, we are able to give a spectral characterization of the preconditioners that may have a theoretical interest in its own.

Keywords: Linear positive operators, Interior Point methods, graph matrices, conditioning and preconditioning, ergodic results.
AMS SC: 05C50, 15A12, 65F10, 47B65.

1 Introduction

Linear Positive Operators (LPOs) are a general tool that has been successfully employed to deal with structured matrices, such as Toeplitz matrices and matrices coming from the discretization by finite differences or finite elements of [semi]elliptic differential problems over bounded domains. Via

[*]Dipartimento di Informatica, Corso Italia 40, 56100 Pisa (ITALY); e-mail: frangio@di.unipi.it

[†]Dipartimento di Energetica, Via Lombroso 6/17, 50100 Firenze, Dipartimento di Informatica, Corso Italia 40, 56100 Pisa (ITALY); e-mail: serra@mail.dm.unipi.it

some related "equivalence" results, the spectral properties of preconditionings of these structured matrices can be studied in details, and therefore efficient preconditioners can be developed. [24, 19, 20].

We will adapt and use these tools to the case of Interior point methods for Linear [18] and Convex [14] Programming, which have grown a well-established reputation as efficient algorithms for large-scale problems. In these methods, at each step we have to solve linear systems of the form

$$E\Theta E^T \mathbf{x} = \mathbf{b}. \tag{1}$$

where E is a $n \times m$ matrix and $\Theta > 0$ is a $m \times m$ diagonal matrix. These linear systems are most often solved by means of direct methods, typically the Cholesky decomposition, possibly preceded by a heuristic reordering of the columns of E aimed at minimizing the "fill-in" [18]. We are interested in the possibility of using iterative methods instead, especially in cases when E is a sparse structured matrix, such as the node-arc incidence matrix of a graph. Therefore, we want to devise optimal and superlinear preconditioners for solving (1).

The paper is organized as follows: in Section 2 we give some general results on LPOs and in Section 3 we specialize them to the case of matrix-valued linear positive operators. These results are then used to show that EE^T is an optimal preconditioner for (1) for "non-degenerating" Θ. In Section 4 we restrict our attention to the case where E is the node-arc incidence matrix of a directed graph, and we study the spectral properties (extremal behaviour and conditioning) of EE^T and $E\Theta E^T$. Finally, in Section 5 we introduce the notion of "sequence of local graphs" and we discuss the preconditioning problem from the viewpoint of the "Locally X" matrices studied in [23], showing how superlinear preconditioners can be devised for this special case.

2 Preliminary results

In the following, the symbol $\Re$ denotes the real field and the symbol $\mathcal{F}(X)$ indicates the set of all the (real) linear spaces of functions acting on the set X. With these notations, we introduce the following definition.

Definition 2.1 *Given $\mathcal{A} \in \mathcal{F}(A)$ and $\mathcal{B} \in \mathcal{F}(B)$, the operator $\Phi : \mathcal{A} \to \mathcal{B}$ is a linear and positive operator (LPO) if*

1. *$\Phi(\alpha f + \beta g) = \alpha\Phi(f) + \beta\Phi(g)$, with $\alpha, \beta \in \Re$, $f, g \in \mathcal{A}$;*

2. $\Phi(f) \geq 0$ *for any nonnegative function* $f \in \mathcal{A}$.

A LPO Φ *is strong (SLPO) if* $\inf \Phi(f) > 0$ *when* f *is nonnegative and not identically zero. A LPO* Φ *is normally positive (LNPO) if the constant function* 1 *belongs to* $\mathcal{A}$ *and*

$$\inf \Phi(1) > 0 \quad \text{and} \quad \sup \Phi(1) < \infty$$

that is, Φ *bounded and strictly positive when applied to* 1.

The natural partial orderings of $\mathcal{A}$ and $\mathcal{B}$ are those induced by the ordering of $\Re$. The Cesaro and Bernstein operators are normally positive as well as the map from $f \in \mathbf{C}([0,1], \Re)$ to the constant $f(0)$. Moreover, the Cesaro operators are strong while the Bernstein operators are nonstrong.

For SLPOs, a theorem about the range of $\Phi(f)/\Phi(g)$ with g nonnegative and not identically zero can be easily proved.

Theorem 2.1 [24] *Let* $\Phi : \mathcal{A} \to \mathcal{B}$ *be a SLPO, and let* $f \in \mathcal{A}$ *and* $g \in \mathcal{A}$ *with* $g \geq 0$ *be given: if* f/g *is not constant, then*

$$r = \inf_C \frac{f}{g} < \frac{\Phi(f)}{\Phi(g)} < R = \sup_C \frac{f}{g}$$

with $C = \{x \in A : \ f(x) \neq 0 \ \text{ or } \ g(x) \neq 0\}$. *Otherwise,* $\Phi(f)/\Phi(g) \equiv f/g = k$ *for some constant* $k \in \Re$.

Theorem 2.1 can be generalized to the case of nonstrong operators; to this end, let us introduce the notion of equivalence between two nonnegative functions f and g, that will be denoted by $f \sim g$.

Definition 2.2 *Two nonnegative functions* f *and* g *are equivalent if there exist* $0 < r \leq R < \infty$ *such that*

$$r \cdot g \leq f \leq R \cdot g.$$

In the above definition, $\leq$ is the partial ordering between two functions; $\sim$ will be used for sequences also, considered as functions acting from $\mathcal{N}$.

Theorem 2.2 [24] *Let* $\Phi : \mathcal{A} \to \mathcal{B}$ *be a LPO, and let* $f \in \mathcal{A}$ *and* $g \in \mathcal{A}$ *with* $g \geq 0$ *be given: if* f/g *is not constant then*

$$r = \inf_C \frac{f}{g} \leq \frac{(\Phi(f))(x)}{(\Phi(g))(x)} \leq R = \sup_C \frac{f}{g} \qquad \forall x \in D$$

where $C = \{x \in A : \ f(x) \neq 0 \ \text{ or } \ g(x) \neq 0\}$ *and* $D = \{x \in B : \ (\Phi(f))(x) \neq 0 \ \text{ or } \ (\Phi(g))(x) \neq 0\}$. *Otherwise,* $\Phi(f)/\Phi(g) \equiv f/g = k$ *for some constant* $k \in \Re$. *If* r *and* R *are finite, then the set where* $\Phi(f)$ *vanishes contains the set where* $\Phi(g)$ *vanishes. Finally, if* $f \sim g$ *(that is,* r *and* R *are finite and strictly positive), then the sets where* $\Phi(f)$ *and* $\Phi(g)$ *vanish coincide.*

When both r and R are finite and positive, Theorem 2.2 tells us that $\Phi(f)$ vanishes where $\Phi(g)$ vanishes exactly as well as f vanishes where g vanishes. In particular, for any bounded $f \in \mathcal{A}$ we have $(\Phi(f))(x) = 0$ for any x that is a zero of $\Phi(1)$. This fact can be used in the analysis of the preconditioning of matrices coming from LPOs: see Lemma 3.1 and Theorem 3.1, as well as [22] for the case of finite differences matrices coming from discretizations of [semi]elliptic problems.

Theorem 2.3 and its subsequent Corollary 3.1 are useful in order to obtain asymptotic information on the extremal behaviour of sequences of LPOs. Actually, the result can be proved for a class of operators slightly more general than LPOs.

Definition 2.3 *Given* $\mathcal{A} \in \mathcal{F}(A)$ *and* $\mathcal{B} \in \mathcal{F}(B)$, *the operator* $\Phi : \mathcal{A} \to \mathcal{B}$ *is a monotone and homogeneous operator (MHO) if*

- $\Phi(f) \geq \Phi(g)$ *if* $f \geq g$ *(monotonicity);*
- $\Phi(kf) = k\Phi(f)$, $k \in \Re$ *(homogeneity).*

Φ *is linear with respect to the constants (Lc) if*

- $\Phi(f + 1) = \Phi(f) + \Phi(1)$.

Remark 2.1 *Each LPO is a MHO.*

Remark 2.2 *Each MHO applied to the zero constant function is identically zero (reproduces the zero constant). If it is also Lc and* $\Phi(1) = 1$, *then it reproduces the constants (*$\Phi(1) = 1$*).*

Remark 2.3 *Each MHO is a positive operator, that is from* $f \geq 0$ *and from the monotonicity and the homogeneity of* Φ *it follows that* $\Phi(f) \geq 0$.

Theorem 2.3 *Let* $\{\Phi_n\}$ *be a sequence of MHOs from* $\mathcal{A}$ *to* $\mathcal{B}$ *such that* $1 \in \mathcal{A}$ *and* $1 \in \mathcal{B}$. *Assume that* $f \in \mathcal{A}$, $g \in \mathcal{A}$, $g \geq 0$ *and* $f - m_f \sim g - m_g$ *where* m_f *and* m_g *are the* inf *of* f *and* g *respectively.*

- *If each Φ_n is Lc, then uniformly with respect to $x \in B$*

$$(\Phi_n(f))(x) - m_f(\Phi_n(1))(x) \sim (\Phi_n(g))(x) - m_g(\Phi_n(1))(x).$$

- *If each Φ_n is Lc and $\Phi_n(\cdot)$ reproduces the constants, then uniformly with respect to $x \in B$*

$$(\Phi_n(f))(x) - m_f \sim (\Phi_n(g))(x) - m_g.$$

- *If $m_f = m_g = 0$, then uniformly with respect to $x \in B$*

$$(\Phi_n(f))(x) \sim (\Phi_n(g))(x).$$

Proof $f - m_f \sim g - m_g$ means that two positive constants r and R exist such that

$$r \cdot (g(x) - m_g) \leq f(x) - m_f \leq R \cdot (g(x) - m_g) \quad \forall x \in A.$$

Therefore, from the monotonocity of Φ_n we infer

$$(\Phi_n(r \cdot (g - m_g)))(x) \leq (\Phi_n(f - m_f))(x) \leq (\Phi_n(R \cdot (g - m_g)))(x)$$

$\forall x \in B$. Recalling that each Φ_n is monotone and homogeneous and Lc, we obtain

$$\begin{aligned} r \cdot ((\Phi_n(g))(x) - m_g(\Phi_n(1))(x)) &\leq (\Phi_n(f))(x) - m_f(\Phi_n(1))(x) \qquad (2) \\ &\leq R \cdot ((\Phi_n(g))(x) - m_g(\Phi_n(1))(x)) \end{aligned}$$

uniformly with respect to $x \in B$. For the second part it is enough to plug $\Phi_n(1) = 1$ in (2). The rest of the theorem can be proved in a similar way. •

3 Matrix-valued LPOs

In this section we consider matrix-valued operators, that is operators having range in the (real) linear space of the $n \times n$ Hermitian matrices $\mathcal{H}_n$, and we give a "matrix version" of the results of the previous section. The only tool that we need to connect the two notions is the well-known Rayleigh quotient [11]. Then, we use these results to prove that EE^T is an optimal preconditioner for (1) in the sense of Definition 3.2.

Definition 3.1 *Given $\mathcal{A} \in \mathcal{F}(A)$, the matrix-valued operator $\Phi : \mathcal{A} \to \mathcal{H}_n$ is linear and positive (LPO) if*

1. *$\Phi(\alpha f + \beta g) = \alpha\Phi(f) + \beta\Phi(g)$, with $\alpha, \beta \in \Re$, $f, g \in \mathcal{A}$;*
2. *$\Phi(f)$ is nonnegative definite for any nonnegative function $f \in \mathcal{A}$.*

Φ is a SLPO if $\Phi(f)$ is positive definite when f is nonnegative and not identically zero. Φ is a LNPO if $\Phi(1)$ is positive definite.

Now we exploit the fact that the Rayleigh operator, $Ra(A)(\mathbf{x}) = \mathbf{x}^H A\mathbf{x}$ with $A \in \mathcal{H}_n$, gives a one-to-one correspondence between matrix-valued LPOs and LPOs whose space $\mathcal{B}$ is $BILH(\mathbf{S}^n, \Re)$, the linear space of the homogeneus, self-adjoint quadratic forms from $\mathbf{S}^n$ to $\Re$, with $\mathbf{S}^n$ the Euclidean unitary sphere of $\mathcal{C}^n$.

Proposition 3.1 *If $\Phi_n(\cdot)$ is a matrix-valued LPO, then*

$$\tilde{\Phi}_n(\cdot) \equiv Ra \circ \Phi_n(\cdot) : \mathcal{A} \to BILH(\mathbf{S}^n, \Re)$$

is a LPO. Moreover, $\tilde{\Phi}_n(\cdot)$ inherits from $\Phi_n(\cdot)$ both normal positivity and strong positivity. Finally, if $\Phi_n(1)$ is the identity matrix, then $\tilde{\Phi}_n(\cdot)$ reproduces the constants.

Proof The linearity of $\tilde{\Phi}_n(\cdot)$ is trivial. By standard linear algebra, the minimum of the quadratic form $Ra(\Phi_n(f))(\mathbf{x}) = \mathbf{x}^H\Phi_n(f)\mathbf{x}$ for $\mathbf{x} \in \mathbf{S}^n$ is the minimum eigenvalue of $\Phi_n(f)$, which must be nonnegative if f is, due to the positivity of $\Phi_n(\cdot)$. The rest of the proof is trivial. •

3.1 Localization results

In this subsection, we introduce the concept of sequences of preconditioners for matrix sequences $\{\Phi_n(f)\}$, where $\{\Phi_n(\cdot)\}$ is a fixed sequence of matrix-valued LPOs. The natural choice is $\{\Phi_n(g)\}$, where g is nonnegative and such that the system $\Phi_n(g)\mathbf{x} = \mathbf{b}$ is easy to solve. The following results precisely indicate how to choose the function g in order to obtain optimal preconditioners.

Definition 3.2 *Let $\{A_n\}$ and $\{B_n\}$ be two sequences of $n \times n$ nonnegative definite Hermitian matrices. $\{B_n\}$ are optimal preconditioners for $\{A_n\}$ if*

- $\ker(B_n^+ A_n) = \ker(A_n)$,
- *all the nonzero eigenvalues of* $B_n^+ A_n$ *belong to a bounded interval* $[r, R]$, *with* r *and* R *positive constants independent on* n.

Note that the preconditioner needs only to work in the subspace where A_n is invertible; in the orthogonal subspace, where A_n is the null matrix, the problem of finding a solution of the related linear system is ill-posed, and therefore it is convenient that B_n also acts like the null matrix. This definition coincides with the one given by Axelsson and Lindskög [5] when each A_n is nonsingular; here, "+" denotes the pseudo-inverse in the sense of Moore-Penrose [13, 15]. In general, $\{B_n\}$ are optimal preconditioners for $\{A_n\}$ if and only if the two sequences of matrices are equivalent in the following sense.

Definition 3.3 *Two sequences of* $n \times n$ *nonnegative definite matrices* $\{A_n\}$ *and* $\{B_n\}$ *are equivalent if there exist* $0 < r \leq R < \infty$ *such that*

$$r\mathbf{x}^T B_n \mathbf{x} \leq \mathbf{x}^T A_n \mathbf{x} \leq R\mathbf{x}^T B_n \mathbf{x} \quad \forall \mathbf{x}.$$

In view of the above discussion, we will denote by $\mathcal{P}_n(f; g)$ the preconditioned matrix $\Phi_n^+(g)\Phi_n(f)$, where g is a nonnegative function of $\mathcal{A}$ and each $\Phi_n(\cdot)$ is a matrix-valued LPO.

Lemma 3.1 *Let* f, $g \in \mathcal{A}$ *with* g *nonnegative and let* r *and* R *be the inf and the sup of* f/g *on the set* C *where it is well defined (the same set defined in Theorem 2.1). If* r *and* R *are finite, then the kernel of* $\Phi_n^+(g)$ *and* $\Phi_n(g)$ *is contained into the kernel of* $\Phi_n(f)$. *Moreover, if* $0 < r \leq R < \infty$ $(f \sim g)$ *then the kernels of* $\Phi_n(f)$ *and* $\Phi_n(g)$ *coincide.*

Proof Applying Theorem 2.2 to $\tilde{\Phi}_n(\cdot)$ as defined in Proposition 3.1, one obtains

$$r \leq \frac{\mathbf{x}^H \Phi_n(f)\mathbf{x}}{\mathbf{x}^H \Phi_n(g)\mathbf{x}} \leq R \quad \forall \mathbf{x} \in \mathcal{C}^n \backslash \ker(\Phi_n(g)). \tag{3}$$

If r and R are finite, then the kernel of $\Phi_n(g)$ is clearly contained in the kernel of $\Phi_n(f)$, for otherwise a sequence $\mathbf{x}_k \in \mathcal{C}^n \backslash \ker(\Phi_n(g))$ would exist that would drive the ratio in (3) to ∞. Analogously, if in addiction $r > 0$ then the kernel of $\Phi_n(f)$ must be contained in the kernel of $\Phi_n(g)$, for otherwise a sequence $\mathbf{x}_k \in \mathcal{C}^n \backslash \ker(\Phi_n(g))$ would exist that would drive the ratio in (3) to 0. •

Theorem 3.1 *If g is nonnegative, then each nonzero eigenvalue of $\mathcal{P}_n(f;g)$ belongs to $[r,R]$, where r and R are those of Lemma 3.1.*

Proof By the first part of Lemma 2.3 in [21] we know that $\mathcal{P}_n(f;g)$ and $\Pi_n(f;g) = (\Phi_n^+(g))^{1/2}\Phi_n(f)(\Phi_n^+(g))^{1/2}$ have the same characteristic polynomial; therefore, each eigenvalue of $\mathcal{P}_n(f;g)$ belongs to the range of the Rayleigh quotient of $\Pi_n(f;g)$ over $\mathbf{S}^n$. Now, select any $\mathbf{z} \in \mathcal{C}^n \backslash \ker(\Phi_n(g))$ and plug $\mathbf{x} = ((\Phi_n^+(g))^{1/2})\mathbf{z}$ in (3) to obtain

$$r \leq \frac{\mathbf{z}^H \Pi_n(f;g)\mathbf{z}}{\mathbf{z}^H[((\Phi_n^+(g))^{1/2})(\Phi_n(g))^{1/2}][((\Phi_n(g))^{1/2})(\Phi_n^+(g))^{1/2}]\mathbf{z}} \leq R. \qquad (4)$$

Since $\mathbf{z}/\|\mathbf{z}\|_2 \in \mathbf{S}^n$, from the properties of the pseudo-inverse and (4) we infer that the Rayleigh quotient of $\Pi_n(f;g)$ over $\mathbf{S}^n$ is contained in the interval $[\min\{0,r\},\max\{0,R\}]$. Hence, the thesis has to be proved only if 0 does not belong to $[r,R]$; we will restrict to the case $r > 0$, since the case of $R < 0$ is analogous.

We must distinguish two subcases: R finite and $R = \infty$.
When R is finite, we have that $0 < r \leq R < \infty$ and therefore f and g are equivalent. From Lemma 3.1 we know that the kernels of $\Phi_n(f)$ and $\Phi_n(g)$ coincide; furthermore, they also coincide with the kernels of $\Phi_n^+(f)$ and $\Phi_n^+(g)$, as a consequence of the classical Schur decomposition. In addition, since $\ker(\cdot) = \mathrm{im}(\cdot)^\perp$ in the case of normal matrices, it is easy to prove by a contradiction argument that they also coincide with the kernel K of $\mathcal{P}_n(f;g)$ and that $\mathrm{im}(\mathcal{P}_n(f;g)) = K^\perp$ (notice that the latter relation is not trivial, since $\mathcal{P}_n(f;g)$ may fail to be normal [11]). Let us now consider a generic vector $\overline{\mathbf{u}} = \overline{\mathbf{u}}_1 + \overline{\mathbf{u}}_2$, such that $\overline{\mathbf{u}}_1 \in K$ and $\overline{\mathbf{u}}_2 \in K^\perp$; due to the nonnegative definiteness of $\Phi_n(f)$ and $\Phi_n(g)$, if $\overline{\mathbf{u}}_2 \neq 0$ then clearly

$$\overline{\mathbf{u}}^H \Phi_n(f)\overline{\mathbf{u}} = \overline{\mathbf{u}}_2^H \Phi_n(f)\overline{\mathbf{u}}_2 \neq 0$$

$$\overline{\mathbf{u}}^H \Phi_n(g)\overline{\mathbf{u}} = \overline{\mathbf{u}}_2^H \Phi_n(g)\overline{\mathbf{u}}_2 \neq 0.$$

In addition, it is well known that

$$\overline{\mathbf{u}}_2 = (\Phi_n^+(g))^{\frac{1}{2}}(\Phi_n(g))^{\frac{1}{2}}\overline{\mathbf{u}}_2 \qquad (5)$$

since $\overline{\mathbf{u}}_2$ exactly belongs to the subspace where $\Phi_n^+(g)$ acts like the inverse of $\Phi_n(g)$.

Now, we prove our thesis using a contradiction argument. Suppose that there exists an eigenvalue $\hat{\lambda}$ of $\mathcal{P}_n(f;g)$, with corresponding eigenvector $\hat{\mathbf{u}}$, such that $\hat{\lambda} \in (0,r)$: $\hat{\mathbf{u}}$ must belong to $K^\perp$. In fact, assume (again, by contradiction) that $\hat{\mathbf{u}}$ does not belong to $K^\perp$: one would have

$$\hat{\lambda}\hat{\mathbf{u}} = \mathcal{P}_n(f;g)\hat{\mathbf{u}} = (\Phi_n^+(g)\Phi_n(f))\hat{\mathbf{u}}. \tag{6}$$

But since the kernels of $\Phi_n^+(g)$ and $\Phi_n(f)$ both coincide with K, then $(\Phi_n^+(g)\Phi_n(f))\hat{\mathbf{u}}$ must belong to $K^\perp$ while $\hat{\mathbf{u}}$ must not, and we obtain the contradiction since $\hat{\lambda}$ is nonzero.

Now, multiply the both sides of (6) for $(\Phi_n(g))^{\frac{1}{2}}$ to obtain

$$(\Phi_n^+(g))^{\frac{1}{2}}\Phi_n(f)\hat{\mathbf{u}} = \hat{\lambda}(\Phi_n(g))^{\frac{1}{2}}\hat{\mathbf{u}}$$

which, combining $\hat{\mathbf{u}} \in K^\perp$ and (5), gives

$$(\Phi_n^+(g))^{\frac{1}{2}}\Phi_n(f)(\Phi_n^+(g))^{\frac{1}{2}}(\Phi_n(g))^{\frac{1}{2}}\hat{\mathbf{u}} = \hat{\lambda}(\Phi_n(g))^{\frac{1}{2}}\hat{\mathbf{u}}. \tag{7}$$

Now, defining $\hat{\mathbf{v}} = (\Phi_n(g))^{\frac{1}{2}}\hat{\mathbf{u}} \neq 0$, left-multiplying both sides of (7) by $\hat{\mathbf{v}}^H$, and dividing the result by $\hat{\mathbf{v}}^H\hat{\mathbf{v}}$, one obtains

$$\hat{\lambda} = \frac{\hat{\mathbf{v}}^H(\Phi_n^+(g))^{\frac{1}{2}}\Phi_n(f)(\Phi_n^+(g))^{\frac{1}{2}}\hat{\mathbf{v}}}{\hat{\mathbf{v}}^H\hat{\mathbf{v}}} = \frac{\hat{\mathbf{u}}^H\Phi_n(f)\hat{\mathbf{u}}}{\hat{\mathbf{u}}^H\Phi_n(g)\hat{\mathbf{u}}}.$$

But the l.h.s. is $\geq r$ (Lemma 3.1), hence the contradiction has been found.

The case $R = \infty$ is similar; we use the first (weaker) part of Lemma 3.1 to deduce that

$$K_f = \ker(\Phi_n(f)) \subset \ker(\Phi_n(g)) = K_g$$

so that $K_g^\perp \subset K_f^\perp$. Then, we proceed as in the previous case, with the only difference that now

$$r\overline{\mathbf{u}}^H\Phi_n(g)\overline{\mathbf{u}} \leq \overline{\mathbf{u}}^H\Phi_n(f)\overline{\mathbf{u}} \leq R = \infty,$$

which creates no problem, since the contradiction is based on $r > 0$. •

In the following, we will denote by $\lambda_j(A)$ the j-th eigenvalue of the Hermitian nonnegative definite matrix A ordered nonincreasingly; that is, $\lambda_1(A) = \lambda_{\max}(A)$ is the largest eigenvalue and $\lambda_n(A) = \lambda_{\min}(A)$ is the smallest eigenvalue. Hence, the spectral condition number [6] of A is

$$\kappa_2(A) = \lambda_1(A)/\lambda_p(A)$$

where p =rank$(A) \leq n$. The analogous notation $\sigma_j(A)$ will be used for the j-th singular value of the generic matirx $A \in \mathcal{C}^{n\times m}$: in this case, $\kappa_2(A) = \sigma_1(A)/\sigma_p(A)$ where p =rank$(A) \leq \min(m,n)$.

Theorem 3.2 *The same statements hold as in Theorem 3.1 with* (r,R) *instead of* $[r,R]$*, and without the exception of the zero eigenvalue, if* $\tilde{\Phi}_n(\cdot)$ *is a SLPO (that is, if it maps a nonnegative and not identically zero function into a positive definite quadratic form).*

Proof This is just a special case of the more general Theorem 2.1 (for a more detailed discussion, refer to [24] and [21]). Indeed, we have

$$\lambda_{\min}(\mathcal{P}_n(f;g)) = \min_{\mathbf{u}\in\mathbf{S}^n} \frac{\mathbf{u}^H\Phi_n(f)\mathbf{u}}{\mathbf{u}^H\Phi_n(g)\mathbf{u}} = \min_{\mathbf{u}\in\mathbf{S}^n} \frac{(\tilde{\Phi}_n(f))(\mathbf{u})}{(\tilde{\Phi}_n(g))(\mathbf{u})}$$

$$\lambda_{\max}(\mathcal{P}_n(f;g)) = \max_{\mathbf{u}\in\mathbf{S}^n} \frac{\mathbf{u}^H\Phi_n(f)\mathbf{u}}{\mathbf{u}^H\Phi_n(g)\mathbf{u}} = \max_{\mathbf{u}\in\mathbf{S}^n} \frac{(\tilde{\Phi}_n(f))(\mathbf{u})}{(\tilde{\Phi}_n(g))(\mathbf{u})}.$$

and we can directly apply Theorem 2.1 to $\tilde{\Phi}_n(\cdot)$ as in Lemma 3.1; the continuity of the Rayleigh quotient over $\mathbf{S}^n$ and the compactness of $\mathbf{S}^n$ guarantee that the strict inequalities are preserved under min and max. •

3.2 Extremal asymptotic behaviour

As a consequence of Theorem 2.3 and of the subsequent remarks, the following corollary can be easily proved.

Corollary 3.1 *Let* $\{\Phi_n(\cdot)\}$ *be a sequence of matrix-valued LPOs defined over* $\mathcal{A}$ *with* $1 \in \mathcal{A}$*. Assume that* $f \in \mathcal{A}$*,* $g \in \mathcal{A}$ *and* $f - m_f \sim g - m_g$ *where* m_f *and* m_g *are the* inf *of* f *and* g*, respectively.*

- *If* $\Phi_n(1) = I_n$ *(*$\tilde{\Phi}_n(\cdot)$ *reproduces the constants), then*

  $$\lambda_j(f) - m_f \sim \lambda_j(g) - m_g;$$

  furthermore, if $m_f = m_g = 0$*, then* $\kappa_2(\Phi_n(f)) \sim \kappa_2(\Phi_n(g))$*.*

- *If* $\Phi_n(1) \neq I_n$ *(*$\tilde{\Phi}_n(\cdot)$ *does not reproduce the constants) and* $m_f = m_g = 0$*, then*

  $$\lambda_j(f) \sim \lambda_j(g), \quad \text{and} \quad \kappa_2(\Phi_n(f)) \sim \kappa_2(\Phi_n(g)).$$

Proof For any j, $\lambda_j(\Phi_n(\cdot))$ is a MHO since $\Phi_n(\cdot)$ is a LPO. Indeed, the homogeneity is a trivial consequence of the linearity, while the monotonicity can be easily proved by exploiting the positivity of $\Phi_n(\cdot)$ and the minimax characterization [9] of the eigenvalues of a Hermitian matrix $A \in \mathcal{H}_n$:

$$\lambda_j(A) = \min_{\mathcal{V}\subset\mathcal{C}^n:\ \dim(\mathcal{V})=n+1-j}\ \max_{\mathbf{x}\in\mathcal{V}\cap\mathbf{S}^n} \mathbf{x}^H A\mathbf{x}.$$

As a consequence, it is easy to see that $A \le B$ implies that $\lambda_j(A) \le \lambda_j(B)$ since $\mathbf{x}^H A\mathbf{x} \le \mathbf{x}^H B\mathbf{x}$ holds for any $\mathbf{x} \in \mathcal{C}^n$; therefore, the linear positivity of $\Phi_n(\cdot)$ induces the mononicity of $\Phi_n(\cdot)$ (by Remark 2.1 and via the auxiliary operator $\tilde{\Phi}_n(\cdot)$), which leads to the monotonicity of $\lambda_j(\cdot)$.

The corollary now follows from the three parts of Theorem 2.3 applied to $\lambda_j(\Phi_n(\cdot))$, noting that $\Phi_n(1) = I_n \Leftrightarrow \tilde{\Phi}_n(1) = 1 \Leftrightarrow \lambda_j(1) = 1$, $\forall j$. •

3.3 Application to Interior Point methods

Let us now specialize the above results to the case of linear systems (1) arising in Interior Point methods. Let θ be a positive m-vector (considered as a function $\{1,\dots,m\} \to \Re_+$), and let Θ be the $m \times m$ diagonal matrix having the components of the θ as diagonal elements: if $\{m = m(n)\}$ is a fixed increasing sequence of positive integers, it is clear that $\Phi_n(\theta) = E_n\Theta E_n^T$ is a sequence of LPOs in a matrix sense. Φ_n is a LNPO if E_n is full-rank: this can always assumed to be the case, since lineraly dependent rows can always be discarded from Linear Programs. Φ_n is not strong because the vanishing of one component of θ can diminish the rank of $\Phi_n(\theta)$; moreover, the related LPOs $\tilde{\Phi}_n(\cdot) = Ra(\Phi_n(\cdot))$ do not reproduce the constants, except for some academical examples.

Therefore, all the theorems concerning the localization of the eigenvalues and the spectral extremal behaviour of $\{\Phi_n(\cdot)\}$ can be applied; more precisely, we have the following proposition.

Proposition 3.2 *Let $\{m(n)\}$ be an increasing sequence of integers, and let $\{\theta_n\}$ be a nondegenerating sequence of positive $m(n)$-vectors (such that $0 < r \le (\theta_n)_i \le R < \infty$). Then*

1. *the sequence $\{\Phi_n(\mathbf{1}_{m(n)}) = E_nE_n^T\}$ is an optimal preconditioning sequence for $\{\Phi_n(\theta_n) = E_n\Theta_nE_n^T\}$ in the sense of Definition 3.2;*
2. *$\lambda_j(\theta_n) \sim \lambda_j(\mathbf{1}_{m(n)})$ and $\kappa_2(\Phi_n(\theta_n)) \sim \kappa_2(E_nE_n^T)$.*

Proof Use Theorem 3.1 and Corollary 3.1. •

Therefore, $\{E_n E_n^T\}$ can be used as preconditioners and the asymptotic behaviour of their spectra describes the asymptotic behaviour of the spectra of all the matrix sequences $\{\Phi_n(\theta_{m(n)}))\}$ for any choice of nondegenerating vector-sequences $\{\theta_{m(n)}\}$. Indeed, a PCG strategy based on a blending of our preconditioning technique and of a (classical) $O(\sqrt{n})$ adaptive updating has been recently used in [3] to obtain a $O\left(\frac{n^3}{\log n} L\right)$ Interior Point method for Linear Programming.

Analogous results have been used in order to characterize the asymptotical behaviour of the condition numbers of Toeplitz matrices, both in the Hermitian [19, 20] and non-Hermitian case [20]. The same theory can also be used to study the asymptotical behaviour of the conditioning of matrices coming from the discretization of [semi]elliptic boundary value problems when the discretization procedures are finite differences [22], as well as Faedo-Ritz-Galerkin methods [24] such as finite element methods or weak spectral methods. Other applications directly related to Theorem 2.3 are the asymptotical evaluation of the extremes of sequences of LPOs approximating the identity in infinity norm: among them, we mention the De La Vallée Poussin integral operators and the well known Cesaro sums [8].

Yet, little can be said in general about the basic matrices $E_n E_n^T$; the following sections are therefore dedicated to the analysis of their spectral and computational features in the special case when E_n is the node-arc incidence matrix of a graph.

4 Graph matrices

Let $H \equiv H_n = (\mathcal{U}_n, \mathcal{V}_n)$ be a directed graph with n nodes $\mathcal{U}_n = \{u_1, \ldots, u_n\}$ and $m = m(n)$ arcs $\mathcal{V}_n = \{v_1, \ldots, v_m\}$: its node-arc incidence matrix $E_n = E(H)$ is the $n \times m$ matrix such that $(E_n)_{i,j} = 1$ if v_j emanates from u_i, $(E_n)_{i,j} = -1$ if v_j terminates at u_i and $(E_n)_{i,j} = 0$ otherwise.

Computations involving this kind of matrices are of practical interest for the solution of Min-Cost Flow problems [1]; promising practical implementations of Interior Point methods for the linear Min-Cost Flow problem have already been developed [16, 17], but with little theoretical analysis.

Here, we complement the results of the previous section with the analysis of the spectral properties of the $\{E_n E_n^T\}_n$ sequences of matrices in the graph case. This analysis can have an interest of its own, as demonstrated by the literature on the subject [10]. However, the usual approach to this problem has most often been of strongly combinatorial flavour, and for a fixed graph size n; our analysis focuses instead on asymptotical results, for which little or no previous work seems to have been published.

4.1 Conditioning of trees

We start by studying the special case when H is a tree. We do not require the arcs to have a specific "direction", since the matrix $E(H^{'})$, corresponding to the *directed* tree obtained from H by re-orienting the arcs, can be obtained from $E(H)$ by right multiplication for a $m \times m$ diagonal $\{-1,+1\}$-matrix. By the SVD decomposition [6, 11], $E(H)$ and $E(H^{'})$ have the same set of singular values; therefore, from the spectral viewpoint the directed tree $H^{'}$ can be considered a special representative of an equivalence class. In this case, $m(n) = n-1$; we can always reorder the nodes and the arcs in such a way that the square submatrix $S(H_n)$ made of the first $n-1$ rows of $E(H_n)$ is nonsingular (in fact, lower triangular with unimodular determinant). Since we can always assume to work with full-rank matrices (Cf. 3.3), we are allowed to discard the last row of $E(H_n)$; therefore, in the following we analyze the spectral behaviour of the matrix sequences $\{S(H_n)\}$.

Proposition 4.1 [10] *$S(H_n)$ is nonsingular and totally unimodular, i.e. the determinant of each square submatrix belongs to $\{\pm 1, 0\}$.*

Proposition 4.2 [10] *The entries of $[S(H_n)]^{-1}$ belong to $\{\pm 1, 0\}$.*

Theorem 4.1 *The conditioning of $S(H_n)$ and $E(H_n)$ satisfy*

$$\overline{\delta}(n-1)^{1/2} \leq \kappa_2(S(H_n)) \leq \sqrt{n}(n-1)$$

$$\overline{\delta} n^{1/2} \leq \kappa_2(E(H_n)) \leq n^{3/2}.$$

where $\overline{\delta} = \sqrt{\delta(H_n)}$ and $\delta(H_n)$ is the maximum degree (number of incident arcs) among all nodes in H_n. Indeed

R1 $\sigma_{\min}(S(H_n)) \leq (n-1)^{-1/2}$;

R2 $\sigma_{\min}(S(H_n)) \geq (n-1)^{-1}$;

R3 $\sigma_{\max}(S(H_n)) \leq n^{1/2}$;

R4 $\sigma_{\max}(S(H_n)) \geq \overline{\delta} \geq \sqrt{2}$.

Proof
Part R1: by the SVD decomposition of A

$$\sigma_{\min}(A) = \inf_{\|\mathbf{x}\|_2>0} \frac{\|\mathbf{x}^H A\|_2}{\|\mathbf{x}\|_2};$$

therefore, being $\mathbf{e}$ the vector of all ones, we have

$$\sigma_{\min}(S(H_n)) \leq \frac{\|\mathbf{e}^H S(H_n)\|_2}{\|\mathbf{e}\|_2} = \frac{1}{(n-1)^{1/2}}.$$

Part R2: by Proposition 4.2 $|[S(H_n)^{-1}]_{i,j}| \leq 1$, hence the entries of $B_n = S(H_n)^{-H} S(H_n)^{-1}$ cannot exceed $n-1$. Therefore, $\|B_n\|_1 \leq (n-1)^2$; since B_n is positive definite, its maximal eigenvalue coincides with its spectral norm and is less than its $\|\cdot\|_1$ norm, hence

$$\lambda_{\max}(B_n) \leq (n-1)^2,$$

but $\lambda_{\max}(B_n) = \lambda_{\max}(S(H_n)^{-1})^2 = (1/\lambda_{\min}(S(H_n)))^2$.

Part R3: From Theorem 1 in [2] it is known that

$$\sigma^2_{\max}(E(H_n)) = \lambda_{\max}(E(H_n)[E(H_n)]^T) \leq n,$$

but $\sigma_{\max}(S(H_n)) \leq \sigma_{\max}(E(H_n))$.

Part R4: by the SVD decomposition of A

$$\sigma_{\max}(A) = \sup_{\|\mathbf{x}\|_2>0} \frac{\|\mathbf{x}^H A\|_2}{\|\mathbf{x}\|_2}.$$

By re-orienting the arcs in such a way that the node with largest degree becomes the root, and choosing $\mathbf{x} = e_1$ (the first vector fo the canonical basis), one obtains $\|\mathbf{x}^H H_n\|_2 = \sqrt{\delta(H_n)}$ and $\|\mathbf{x}\|_2 = 1$; note that $\delta(H_n) \geq 2$ since H_n is connected.

The bounds on the condition numbers of $S(H_n)$ are simple consequences of **R1**–**R4** above. For the minimal and the maximal singular values of $E(H_n)$, as well as its asymptotic conditioning, we observe that $S(H_n)$ is a submatrix of $E(H_n)$ which, in turn, is a submatrix of $S(H_{n+1})$ where H_{n+1} it a new tree obtained by H_n by adding a new arc emanateing from the node that was deleted by $E(H_n)$ in order to define $S(H_n)$ and terminating in a new leaf. Therefore, $E(H_n)$ is "embedded" between $S(H_n)$ and $S(H_{n+1})$, and we can apply a rewording of the Cauchy interlace theorem that holds for non–Hermitian matrices [9]. •

The estimates **R1**–**R4** are, up to positive constants, tight: the following special structures are the "extremes" that prove it.

4.1.1 Linear trees

H_n is a linear tree if its directed representative $H_n^{'}$ is a path (that is, each node has at most one son). By choosing the natural ordering of the nodes suggested by the linearity of the structure, we obtain a bidiagonal matrix $E(H_n)$; in particular, $S(H_n^{'})$ is the $(n-1)\times(n-1)$ square Toeplitz matrix generated by the symbol $f(x) = 1 - e^{\mathbf{i}x}$ [19, 7]. $f(x)$ is weakly sectorial [7] and has a zero of order 1: therefore, in light of the analysis given in [7], we deduce that

$$\sigma_{\min}(S(H_n^{'})) \sim n^{-1} \quad \text{and} \quad \lim_{n\to\infty} \sigma_{\max}(S(H_n^{'})) = \|f\|_\infty = 2$$

with $\sigma_{\max}(S(H_n^{'})) \le \|f\|_\infty = 2$. Hence, **R2** and **R4** are tight (up to suitable multiplicative constants) for linear trees.

Remark 4.1 *Observe that the constant $\|f\|_\infty = 2$ is exactly the maximum degree of this kind of linear trees. From* **R4** *$\sigma_{\max}(S(H_n)) \ge \overline{\delta} \ge \sqrt{2}$: therefore, in the case of linear trees of large size this lower bound is not tigth and can be replaced by $\delta(H_n)$ minus a positive quantity which is asymptotically small with n.*

4.1.2 Star trees

In the opposite direction, we have "concentrated" trees, the most concentrated one being the "star" tree where the root has $n-1$ sons. Let us choose any ordering for the nodes where the first node is the root, and let us order the arcs according to the chosen order of the nodes; the resulting $E(H_n)$ is not lower triangular, but the corresponding $S(H_n)$ has the following interesting structure:

$$S(H_n) = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & & & \vdots \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & -1 & 0 \end{bmatrix}. \tag{8}$$

This structure is close to the one of the Frobenius matrices [6], and it is easy to prove that the characteristic polynomial is $p(\lambda) = (1+\lambda^n)/(1+\lambda)$; however, since $S(H_n)$ is "highly nonnormal" [6], the fact that all the eigenvalues have unitary modulus does not tell anything about the conditioning of the matrix. To this end, we study the matrix

$$S(H_n)^H S(H_n) = I_{n-1} + [\mathbf{f}|\mathbf{e}_1] \cdot [\mathbf{e}_1|\mathbf{g}]^H$$

where $\mathbf{e}_j$ is the j-th vector of the canonical basis, $\mathbf{g} = -\sum_{j=2}^{n-1} \mathbf{e}_j$ and $\mathbf{f} = (n-2)\mathbf{e}_1 + \mathbf{g}$. Since the matrix $[\mathbf{f}|\mathbf{e}_1] \cdot [\mathbf{e}_1|\mathbf{g}]^H$ has rank two, the complete set of the eigenvalues of $S(H_n)^H S(H_n)$ is made up by 1 with multiplicity $n-3$ plus two other values that can be explicitly calculated. The nonzero eigenvalues of $[\mathbf{f}|\mathbf{e}_1] \cdot [\mathbf{e}_1|\mathbf{g}]^H$ are those of the 2×2 matrix [6]

$$[\mathbf{e}_1|\mathbf{g}]^H \cdot [\mathbf{f}|\mathbf{e}_1] = \begin{bmatrix} n-2 & 1 \\ n-2 & 0 \end{bmatrix};$$

direct calculation yields

$$\lambda_{\max}(S(H_n)^H S(H_n)) = n - 2 + O(n^{-1})$$

$$\lambda_{\min}(S(H_n)^H S(H_n)) = 2/(n-2) + O(n^{-2}).$$

Therefore

$$\sigma_{\max}(S(H_n)) = \sqrt{n - 2 + O(n^{-1})}$$

$$\sigma_{\min}(S(H_n)) = \sqrt{\frac{2}{n-2} + O(n^{-2})}.$$

which prove that **R1** and **R3** are tight up to suitable multiplicative constants.

Remark 4.2 *In the case of a "star" tree, the root has degree $n-1$. Hence, $\sigma_{\max}(S(H_n)) = \sqrt{n-2+O(n^{-1})} = \sqrt{\delta(H_n)} + O(1)$ proves that* **R4** *cannot be relaxed any further. For instance, the square root of the maximum degree cannot be replaced by the maximum degree as "substantially" done for linear trees (see Remark 4.1).*

Remark 4.3 *Tightness of* **R1**–**R4** *does not imply that the estimates concerning the spectral conditioning in Theorem 4.1 are tight. In fact, both "linear" and "star" trees have $O(n)$ condition numbers, as opposed to the $O(n^{3/2})$ bound given in the Theorem.*

4.2 Conditioning of graphs

The results of the previous section are now used in order to evaluate the extremal behaviour of the singular values of $E(H_n)$ when $\{H_n\}$ is a sequence of generic graphs with n nodes.

If H_n has k maximal connected components, then $E(H_n)$ is a block diagonal matrix with k blocks; the minimal[maximal] singular value of $E(H_n)$ is the minimum[maximum] among the minimal[maximal] singular values of the submatrices associated to the connected components, and $E(H_n)[E(H_n)]^T$ has exactly k zero eigenvalues.

Hence, we can restrict ourselves to connected graphs: since trees have been analyzed before, we will reconduce the case of connected graphs to the case of trees.

Proposition 4.3 *Let $H_n = (\mathcal{U}_n, \mathcal{V}_n)$ be a connected graph which is not a tree: then $\sigma_{\min}(E(H_n)) = 0$ and*

$$\sigma_{n-1}(E(H_n)) \geq \max_{T \in T(H_n)} \sigma_{\min}(S(T))$$

where $T(H_n)$ is the set of the spanning trees of H_n.

Proof Since H_n is connected but it is not a tree, it has at least as many columns as rows: summing all the rows of H_n gives the null vector, and this proves that $\sigma_{\min}(E(H_n)) = 0$.

Now, let T be a generic spanning tree of H_n: reorder the nodes and the arcs of H_n in such a way that T is represented by the first $n-1$ columns of $E(H_n)$ and $S(T)$ is represented by the first $n-1$ rows and columns (the application of permutation matrices does not modify the singular values). Therefore, we have

$$\begin{aligned}
\sigma_{n-1}(E(H_n)) &= \sup_{\mathrm{dim U}=n-1} \inf_{\mathbf{x} \in U} \frac{\|\mathbf{x}^H E(H_n)\|_2}{\|\mathbf{x}\|_2} \\
&\geq \inf_{\mathbf{y} \in \mathcal{C}^{n-1}} \frac{\|[\mathbf{y}^H, 0] E(H_n)\|_2}{\|[\mathbf{y}, 0]\|_2} \\
&= \inf_{\mathbf{y} \in \mathcal{C}^{n-1}} \frac{\sqrt{\|\mathbf{y}^H S(T)\|_2^2 + \|\mathbf{w}\|_2^2}}{\|\mathbf{y}\|_2} \\
&\geq \inf_{\mathbf{y} \in \mathcal{C}^{n-1}} \frac{\|\mathbf{y}^H S(T)\|_2}{\|\mathbf{y}\|_2} = \sigma_{\min}(S(T)).
\end{aligned}$$

•

¿From Proposition 4.3 and part **R2** of Theorem 4.1 we deduce the inequality $\sigma_{n-1}(E(H_n)) \geq n^{-1}$. On the other hand, it is known $\sigma_{\max}(E(H_n)) \leq n^{1/2}$ [2]. As a consequence, the following theorem holds.

Theorem 4.2 *$\kappa_2(E(H_n))$ grows at most as $n^{3/2}$.*

These bounds are asymptotically tight: linear trees realize the former, while the latter is realized by complete graphs. In fact, the matrix

$$E(H_n)[E(H_n)]^T$$

corresponding to a complete graph H_n is the circulant matrix $nI - \mathbf{e}\mathbf{e}^T$, whose maximal eigenvalue is n with multiplicity $n-1$ (all the nonzero vectors orthogonal to $\mathbf{e}$ are eigenvectors associated to the eigenvalue n) and whose minimal eigenvalue is zero.

The bound on the condition number is asymptotically realized by a sequence of graphs H_n with two components T_n^1 and T_n^2, where T_n^1 is a star tree with $n/2$ nodes and T_n^2 is a linear tree with $n/2$ nodes. The maximal singular value of $E(H_n)$ coincides with the one of $E(T_n^1)$, growing as $(n/2)^{1/2}$, while the minimal nonzero singular value of $E(H_n)$ coincides with the one of $E(T_n^2)$, collapsing to zero as $(n/2)^{-1}$. Therefore, the spectral condition number behaves as $n^{3/2}$.

The former example suggests one a further example which answer in the positive to the question raised in Remark 4.3. Consider a sequence of trees $\hat{H}_n$ formed by the union of T_n^1 and T_n^2 with a new node u and the two arcs (u, u_1) and (u, u_2), where u_1 is the root of T_n^1 (the center of the star) and u_2 is the root of T_n^2. The corresponding $(n+1) \times n$ matrix $E(\hat{H}_n)$ has full-rank: therefore,

$$\sigma_{\min}(E(\hat{H}_n)) = \sigma_n(E(\hat{H}_n)) = \inf_{\|\mathbf{x}\|_2=1} \|E(\hat{H}_n)\mathbf{x}\|_2.$$

Let the ordering of the nodes and of the arcs of $\hat{H}_n$ be such that the first $n/2$ rows and columns are related to the linear tree T_n^2: it is easy to check that, by choosing $\mathbf{x} = \mathbf{y}/\|\mathbf{y}\|_2$ where $\mathbf{y}_j = \sin\left(\frac{2j\pi}{n}\right)$ for $j = 1, \ldots n/2$ and zero otherwise, one has $\|E(\hat{H}_n)\mathbf{x}\|_2 \sim n^{-1}$.

5 Locally X matrices

In this section, we introduce the notion of sequences of "Locally X" matrices, where $X(\cdot) = \{X_n(\cdot)\}$ is a given fixed sequence of LPOs. For this kind of matrices, introduced in [23], it is possible to define "good" preconditioners and to analyze in detail the distribution of the spectra of the preconditioned matrices. In the following, we first define the "Locally X" matrices and then we construct nontrivial classes of graphs $\{H_n\}$ such that the associated sequence of LPOs $X(\cdot) = \{X_n(\cdot)\} = \{E(H_n)\Theta_n(\cdot)[E(H_n)]^T\}$ generates a class of "Locally X" matrices for a proper choice of the (sequence of) functions $\{\Theta_n(\cdot)\}$ mapping points in the reference space into arc weights. Such a sequence of graphs will be called a "local-graph sequence".

5.1 Definitions and premises

We first define the notion of (spectral) cluster: this is useful in order to give more general versions of the concept of "optimal preconditioner" provided by Definition 3.2.

Definition 5.1 *Consider a sequence $S = \{S_n\} = \{\{S_i^n\}_{i=1}^n\}$ of real vectors (of increasing size) and a set M in the real line. Denote by M_ϵ the ϵ-extension of M, that is the union of all balls of radius ϵ centered at points of M. For any n, let $\gamma_n(\epsilon)$ count those S_i^n that do not belong to M_ϵ. If*

- *$\forall\epsilon > 0 \quad \gamma_n(\epsilon) = o(n)$ as $n \to \infty$, then S is generally clustered at M (M is a general cluster for S);*
- *$\forall\epsilon > 0 \quad \exists c(\epsilon)$ s.t. $\gamma_n(\epsilon) \le c(\epsilon)\ \forall n$, then S is properly clustered at M (M is a proper cluster for S).*

In the special case where $M = \{p\}$, we say that S is generally/properly clustered at p. The same notion applies to sequences of Hermitian/general matrices $A = \{A_n\}_n$, by taking as S_n the complete set of the eigenvalues/singular values of A_n [25]*: in this case, we say that A is spectrally (generally/properly) clustered at M.*

The notion of proper clustering and the notion of optimal preconditioners given in Definition 3.2 are related to results concerning the convergence rate of PCG methods with similar statements, but requiring quite different techniques for the proofs, as detailed in the two following remarks.

Remark 5.1 *If $\{B_n\}$ is an optimal preconditioning sequence for $\{A_n\}$ in the sense of Definition 3.2, then the number of PCG iterations required in order to reach the solution (in the least squares sense) within a preassigned accuracy is bounded by a constant independent of n [5]. In this case, we say that the PCG method is optimal.*

Note that the conditions in Definition 3.2 are stronger than the requirement that $\{B_n^+ A_n\}$ be spectrally clustered at $[r, R]$: indeed, the optimality of the PCG method still holds with a slightly stronger definition of proper clustering requiring all the "bad" singular values outside M to be on the right side of the interval [5].

Remark 5.2 *If $\{B_n\}$ is such that the preconditioned sequence $\{B_n^+ A_n\}$ is spectrally properly clustered at $p > 0$ and the minimal singular value of $B_n^+ A_n$ not related the the kernel of A_n does not go to zero too fast, then the corresponding PCG method is optimal and the "quality" of the convergence is superlinear [5].*

In the following, we will denote by $\Omega \subset \Re^d$ a fixed bounded domain, and by $K = \bar{\Omega}$ its closure. We will also denote by $B_K(x_0, \rho)$ any ball of centre x_0 and radius ρ contained in K, i.e. $B_K(x_0, \rho) = \{x \in \Re^d : \ \|x - x_0\|_\infty \leq \rho\} \subset K$, and by $\mathrm{meas}_K(X)$ the measure of the set X relative to K, i.e.

$$\mathrm{meas}_K(X) = \mathrm{meas}(X)/\mathrm{meas}(K).$$

Definition 5.2 *A n-split of K is any collection $\mathcal{P}_n = \{\mathcal{P}_i^n\}_{i=1}^n$ of n disjoint connected Peano-Jordan* [12, pp. 28–29] *subsets of K; a n-split is a n-grid if each subset is pointed, i.e. $\mathcal{P}_i^n = \{x_i^n\}$, while it is a n-partition if it covers all K, i.e. $\bigcup_{i=1}^n \mathcal{P}_i^n = K$.*

Definition 5.3 *A sequence $\mathcal{P} = \{\mathcal{P}_n\}$ of splits of K is:*

- *quasi–uniform if for any x_0 and $\rho > 0$*

  $$\#\{\mathcal{P}_i^n \subset Y\} = n \cdot \mathrm{meas}_K(Y) + o(n), \quad Y \in \{B_K(x_0, \rho), K \backslash B_K(x_0, \rho)\};$$

- *uniform if for any x_0 and $\rho > 0$*

  $$\#\{\mathcal{P}_i^n \subset Y\} = n \cdot \mathrm{meas}_K(Y) + O(n^{\frac{d-1}{d}}), \ Y \in \{B_K(x_0, \rho), K \backslash B_K(x_0, \rho)\};$$

- *$*$-uniform if for any connected set $S \subset K$ with Lipschtiz continuous boundaries*

$$\#\{\mathcal{P}_i^n \subset Y\} = n \cdot \text{meas}_K(Y) + O(n^{\frac{d-1}{d}}), \quad Y \in \{S, K \backslash S\}.$$

Uniform sequences of partitions will be instrumental in defining the concept of "locality" that is central to our developement: in fact, locality is defined on K by using the $\mathcal{P}_i^n$ subsets as the "neighbourhood structure", and then transported into an appropriate matrix space. Although uniform sequences of grids are not strictly necessary for our developement, they often arise naturally in applications, as clearly shown in the following Sections.

As a first application of the above definition, let us define the class $PC(K)$ of all the piecewise-continuous functions over K: f belongs to $PC(K)$ if an only if an n-partition $\mathcal{P}_n$ exists (for some n) such that $f|_{\mathcal{P}_j^n} \in C(\mathcal{P}_j^n)$, i.e. f is continuous if restricted to each $\mathcal{P}_j^n$.

5.2 The asymptotic notion of "Locality"

We now define the concept of "Locally X" sequences of matrices. Note that this definition does not apply to a single matrix, but only to sequences of matrices of increasing order: in fact, this notion describes asymptotical properties.

Let $X(\cdot) = \{X_n(\cdot)\}$ be a sequence of $n \times n$ Hermitian matrix–valued LPOs defined over $PC(K)$; given this "basic" sequence, we define

$$\mathcal{LL}_n(X) = \bigcup_{a \in PC(K)} X_n(a), \qquad \mathcal{LL}(X) = \bigcup_{n \in \mathcal{N}} \mathcal{LL}_n(X)$$

i.e. the class of all possible matrices obtainable from X. Under proper "locality" assumptions on X, the spectra of all the matrices in the sequence $\{X_n(a)\}$ can be precisely characterized in terms of the weight function a; this also allows to define "good" preconditioners.

Let $\mathcal{P}$ be a sequence of partitions of K: for each n, we can "decompose" $X_n(a)$ by means of the following sequence of matrix–valued LPOs

$$V_n(a) = \{V_i^n(a)\}_{i=1}^n = \{X_n(a|_{\mathcal{P}_i^n})\}_{i=1}^n.$$

In fact, an easy consequence of the linearity of $X_n(\cdot)$ is

$$X_n(a) = \sum_{i=1}^n V_i^n(a);$$

therefore, the sequence of matrices

$$V_n = \{V_i^n\}_{i=1}^n = \{X_n(I_{\mathcal{P}_i^n})\}_{i=1}^n$$

can be viewed as a "base" for $\mathcal{LL}_n(X)$ (I_A being the characteristic function of a set A).

Roughly speaking, what is required for $\mathcal{LL}(X)$ to have a "locality" structure is that is that each V_i^n covers an asymptotically small part of the range of $X_n(\cdot)$, and that the V_i^ns only interact according to the locality relationships dictated by $\mathcal{P}_n$. To make these notions more precise, let $Z = \{\mathbf{Z}_n\} = \{\{\mathbf{z}_i^n\}_{i=1}^n\}$ be a sequence of orthonormal bases of $\mathcal{C}^n$, and let $\mathcal{N} = \{\mathcal{N}_n\}$ be a sequence of bijections from $\{\mathcal{P}_n\}$ onto $\{\mathbf{z}_n\}$: we will require the following two conditions to hold

$$\#\{j : [\mathbf{z}_j^n]^H V_i^n \mathbf{z}_j^n \neq 0\} = O(1) \quad [\text{rank}(V_i^n) = O(1)] \quad \forall i \leq n \tag{9}$$

$$[\mathbf{z}_j^n]^H V_i^n \mathbf{z}_j^n \neq 0 \Rightarrow \text{dist}(\mathcal{P}_i^n, \mathcal{P}_j^n) = o(1) \quad \forall i, j \leq n. \tag{10}$$

In other words, (9) and (10) say that each V_i^n must be made of a *constant* number of vectors $\mathbf{z}_j^n$, choosen according to the neighbourhood structure dictated by the $\mathcal{P}_i^n$.

Definition 5.4 *If there exist a quasi–uniform sequence of partitionings $\mathcal{P}$ of K, a sequence of hortonormal bases Z and a sequence of bijections $\mathcal{N} = \{\mathcal{N}_n : \mathcal{P}_n \mapsto \mathbf{Z}_n\}$ such that both (9) and (10) hold, then $\mathcal{LL}(X)$ is said to be "Locally X". If $\mathcal{P}$ is $*$-uniform and (10) holds with $o(1)$ replaced by $O(n^{-1/d})$, then $\mathcal{LL}(X)$ is "Locally X" in a strong sense.*

We stress here that, in Definition 5.4, a "Locally X" class of matrices is intimately tied with a sequence of partitionings $\mathcal{P}$: the choice of $\mathcal{P}$ does not change the class $\mathcal{LL}(X)$ itself, but it is necessary to prove its "Locally X"-ness. The same of course holds for Z and $\mathcal{N}$, although their choice is usually easier in applications. Thus, whenever we speak of a "Locally X" class of matrices, we should see some sequence of partitionings $\mathcal{P}$ attached to it (together with Z and $\mathcal{N}$) - specifically, the one(s) that serves for proving its "Locally X"-ness. Note however that different $\mathcal{P}$s can serve the purpose for the same $\mathcal{LL}(X)$.

5.3 The preconditioners

For "Locally X" matrices, asymptotic results can be given that relate the spectra of $X_n(a)$ with the behaviour of the function a over K [23]. Here we focus our attention on spectral results concerning the preconditioned matrix-sequences. Let us define

$$\begin{aligned}
A_+(a) &= \{x \in K : \ a(x) > 0\}, \quad \alpha_+(a) = \mathrm{meas}_K(A_+(a)), \\
A_-(a) &= \{x \in K : \ a(x) < 0\}, \quad \alpha_-(a) = \mathrm{meas}_K(A_-(a)), \\
A_0(a) &= \{x \in K : \ a(x) = 0\}, \quad \alpha_0(a) = \mathrm{meas}_K(A_0(a)).
\end{aligned}$$

Fix a and b in the space $C(K)$ so that b is positive: we can think of using the "standard" preconditioning of Section 3.1

$$\mathcal{P}_n(a;b) = X_n^{-1}(b)X_n(a).$$

However, in the case of "Locally X" matrices alternative choices are possible. Let $\mathbf{Z}_n$ be the (hortonormal) matrix having as columns the vectors $\mathbf{z}_j^n$, and let $D_n(a)$ be the diagonal matrix obtained extracting the main diagonal of $\mathbf{Z}_n^H X_n(a)\mathbf{Z}_n$. Now, let D_p be the diagonal matrix obtained extracting the main diagonal of $X_n(1)$ (in the case of graphs, $X_n(1) = E_n E_n^T$ and D_p is the diagonal matrix having on the the degree of node i in position (i,i)), and $\hat{D}_n(a) = \mathbf{Z}_n D_n^{1/2}(a) D_p^{-1/2}\mathbf{Z}_n^H$; an alternative preconditioning is

$$\mathcal{P}_n^*(a) = [\hat{D}_n(a)X_n(1)\hat{D}_n(a)]^{-1}X_n(a).$$

Finally, assume that for each partitioning $\mathcal{P}_n$ we select a "compatible" grid $\mathcal{G}_n$ such that $x_i^n \in \mathcal{P}_i^n$, let $\Delta_n(a)$ be the diagonal matrix such that $(\Delta_n(a))_{i,i} = a(x_i^n)$ and let $\hat{\Delta}_n(a) = \mathbf{Z}_n\Delta_n^{1/2}(a)D_p^{-1/2}\mathbf{Z}_n^H$: a third possible preconditioning is given by

$$\mathcal{P}_n^{**}(a) = [\hat{\Delta}_n(a)X_n(1)\hat{\Delta}_n(a)]^{-1}X_n(a).$$

For these new sequences of matrices we proved spectral results [23] in the style of the Szegö relations [25] concerning the asymptotic distribution of the eigenvalues: these results are summarized in the following theorems.

Theorem 5.1 [23] *Let $\mathcal{LL}(X)$ be "Locally X", and let a and b be two equivalent functions such that $A_0(a - sb)$ is Peano-Jordan for all $s \in T$ with T*

dense in $\Re$ and $0 \in T$. Then, for any continuous function F with bounded support

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} F(\lambda_i(\mathcal{P}_n(a;b)) = \int_{A_+(b)} F((a/b)(x))dx + \mathrm{meas}\{A_0(b)\}F(0). \quad (11)$$

This is a quite powerful "ergodic" result, meaning that the spectra of $\{\mathcal{P}_n(a;b)\}$ essentially "samples" the function

$$\eta(x) = \{a(x)/b(x) \ \text{ if } \ b(x) > 0, \ 0 \ \text{ otherwise}\}$$

in n points, so that the sample becomes better and better as n increases. As a direct consequence, we can state a weaker localization result of asymptotic type (recall Theorem 3.1).

Corollary 5.1 *Under the hypotheses of Theorem 5.1, the topological closure of the set of the nonzero eigenvalues of all the matrices $\{\mathcal{P}_n(a;b)\}$ is contained in the usual range $[r, R]$ of a/b (over the set where b is nonzero); moreover, the eigenvectors related to the null eigenvalue belong, for any n, to the null space of $X_n(a)$.*

When the a and b are not equivalent, $X_n(b)$ is not a good preconditioner for $X_n(a)$ since a "lot" (at least ϕn with fixed $\phi > 0$) of eigenvalues of $\mathcal{P}_n(a;b)$ not related to the kernel of $X_n(a)$ are in any neighbourhood of zero and/or of infinity; in the light of the Axelsson Linskog analysis [5], this leads to a substantial deterioration of the performances of the related PCG method. Hence, in this case the alternative preconditioners may be useful.

Theorem 5.2 *Let $\mathcal{LL}(X)$ be "Locally X", let a be a nonnegative function with $A_0(a)$ being Peano-Jordan and let $\{X_n(1)\}$ be sparsely vanishing i.e.*

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{\#\{\lambda_i(X_n(1)) \le \epsilon\}}{n} = 0.$$

Then, for any continuous function F with bounded support

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} F(\lambda_i(\mathcal{P}_n^{*[*]}(a))) = \int_{A_+(a)} F(1)dx + \mathrm{meas}(A_0(a))F(0). \quad (12)$$

Therefore, $\{\mathcal{P}_n^*(a)\}$ and $\{\mathcal{P}_n^{**}(a)\}$ are spectrally clustered at one if a is positive or with isolated zeros; when the cluster is proper, the matrices $\{\hat{D}_n(a)X_n(1)\hat{D}_n(a)\}$ and $\hat{\Delta}_n(a)X_n(1)\hat{\Delta}_n(a)$ are superlinear preconditioners for $\{X_n(a)\}$ in view of Remark 5.2.

5.4 "Local graphs"

Here we define a notion of "local graphs": let us start with some examples.

A sequence of uniform grids over a bounded iper-rectangle of $\Re^d$, with a "local connection" of the nodes to a constant number of neighbours define a sequence of local graphs (d-dimensional ipercube graphs): these types of graphs characterize e.g. finite differences and finite elements matrices for differential equations [4].

One very simple instance of the above class is obtained when $d = 1$, that is K is a segment, and each node - point in the segment - is only linked to its predecessor and successor in the natural ordering (a linear tree, see Section 4.1). We will now show in detail what "Locally X" matrices look like in this particular case. Let $\mathbf{e}_j$ be the j-th vector of the canonical base, and E_n^j be the j-th column of E_n: in this simple case

$$E_n^j = e_j - e_{j+1}, \quad E_n^n = e_n.$$

Then, by setting

- $\mathcal{P}_i^n = [(i-1)/n, i/n)$, $i = 1, \dots, n-1$, and $\mathcal{P}_n^n = [(n-1)/n, 1]$;
- $X_n(a) = E_n \mathrm{diag}_i(a_i^n) E_n^T$ ($a_i^n = a(x_i^n)$ for some $x_i^n \in \mathcal{P}_i^n$)

we have $X_n(a) = \sum_{j=1}^n \tilde{a}_j^n E_n^j {E_n^j}^T$: the base matrices V_i^n in this case are just diads. By choosing $\mathbf{Z}_n = I_n$, it easily follows that (9) and (10) (with $o(1)$ replaced by $O(n^{-1/d})$) hold: since $\mathcal{P}$ is $*$-uniform and $X_n(\cdot)$ is a LPO for any n, we the class $\mathcal{LL}(X)$ is "Locally X" in the strong sense according to Definition 5.4.

The simplicity of this example should not lead to the conclusion that only "trivial" graphs can be local. Indeed, for $d > 1$ there is no "natural" ordering of the nodes as there was in the example, and the matrices V_i^n must be accurately choosen in such a way that they reflect the (complex) neighbourhood relationships dictated by the $\mathcal{P}_i^n$. However, even this simple example is not without interest of its own: in fact, it is easy to see that the matrix $X_n(a)$ has the form

$$\begin{bmatrix} \tilde{a}_1^n & -\tilde{a}_1^n & & \\ -\tilde{a}_1^n & \tilde{a}_1^n + \tilde{a}_2^n & -\tilde{a}_2^n & \\ & \ddots & \ddots & \ddots \\ & & -\tilde{a}_{n-1}^n & \tilde{a}_{n-1}^n + \tilde{a}_n^n \end{bmatrix}. \tag{13}$$

Therefore, this local graph is clearly related to the discretization by centered finite differences of precision order two of the following second order boundary value problem

$$\begin{cases} -\left(a(x)u_x\right)_x = f\,(x) & \text{on } \Omega = (0,1) \\ \text{Dirichlet B.C.} \quad \text{on } \partial\,\Omega \end{cases} \tag{14}$$

which leads to the matrices of the form

$$\begin{bmatrix} a_{\frac{1}{2}} + a_{\frac{3}{2}} & -a_{\frac{3}{2}} & & \\ -a_{\frac{3}{2}} & a_{\frac{3}{2}} + a_{\frac{5}{2}} & -a_{\frac{5}{2}} & \\ & \ddots & \ddots & \ddots \\ & & -a_{n-\frac{1}{2}} & a_{n-\frac{1}{2}} + a_{n+\frac{1}{2}} \end{bmatrix}, \quad a_t = a\left(\frac{t}{n+1}\right). \tag{15}$$

Having in mind these simple examples, we can now give a general definition.

Definition 5.5 *A sequence $H = \{H_n\} = \{(\mathcal{U}_n, \mathcal{V}_n)\}$ of graphs is "weakly local" if there exists a finite positive integer d, a bounded closed domain $K \subset \Re^d$ and a quasi-uniform sequence $\mathcal{P}$ of partitionings of K such that*

Hp1 $\#\mathcal{P}_n = \#\mathcal{U}_n = n$;

Hp2 *the degree of each node is bounded by an absolute constant;*

Hp3 $\sup_{(i,j)\in\mathcal{V}_n} \mathrm{dist}(\mathcal{P}_i^n, \mathcal{P}_j^n) = o(1)$.

H is "strongly local" if **Hp3** *holds with $O(n^{-1/d})$ in place of $o(1)$ and $\mathcal{P}$ is $*$-uniform.*

Needless to say, $\mathcal{P}$ will be the "reference" sequence of partitionings for the "Locally X" matrices associated with H: the corresponding Z and $\mathcal{N}$ will just be the sequence of the canonical bases and the mapping from each node i (subset $\mathcal{P}_i^n$) to the i-th vector of the canonical base $\mathbf{e}_i^n$.

However, a local sequence of graphs is not enough to properly define a sequence of LPOs out of their graph matrices: in fact, the usual definition would be

$$\{X_n(\theta) = E_n \Theta E_n^T\}$$

where Θ is the diagonal matrix $\Theta_{w,w} = \theta(w)$. But the definition of "Locally X" requires $X_n(\cdot)$ to be a function taking argments in $PC(K)$: therefore, we must find a meaningful way for constructing a vector of arc weights from a function taking arguments in K. One possible way of doing that is to choose a sequence of $m(n)$-grids of K

$$S_n = \{s_{i,j}^n\}_{(i,j)\in\mathcal{V}_n}$$

and to define the sequence of matrix–valued LPOs associated with H as

$$\{X_n(a) = E_n\Theta_n(a)E_n^T\} \quad \text{where} \quad \Theta_n(a)_{i,j} = a(s_{i,j}^n),\ a \in PC(K).$$

Here, a local graph is in some sense the skeleton of the bounded domain K, where n regions $\mathcal{P}_i^n$ are defined and for each arc (i,j) a "sample" is taken: in order to prove "locality" properties, we clearly have to relate the sample points $s_{i,j}^n$ with the regions. It turns out that the following property is all that is needed:

$$\text{either} \quad s_{i,j}^n \in \mathcal{P}_i^n \quad \text{or} \quad s_{i,j}^n \in \mathcal{P}_j^n. \tag{16}$$

In a slightly more general definition, $\Theta_n(a)_{i,j}$ may be taken as a convex combination of a constant number of samples of $a(\cdot)$ in points belonging either to $\mathcal{P}_i^n$ or to $\mathcal{P}_j^n$ (this happens e.g. in finite difference discretizations of partial differential equations): this would complicate the notation and it is avoided for the sake of simplicity.

5.5 "Local graphs" and "Locally X" matrices

In this section, we show that the sequence $X = \{X_n\}$ defined above generates a class of 'Locally X" matrices for sequences of "local" grahs and if (16) holds. This is done by just veryfying that all the necessary properties are satisfied.

- By construction, it is clear that each $X_n(\cdot)$ is a linear and normally positive matrix–valued operator.
- It is easy to see that $\Theta_n(I_{\mathcal{P}_i^n})$ is a diagonal matrix having a 1 on the diagonal entry corresponding to arc (u,v) if $s_{u,v}^n \in \mathcal{P}_i^n$, and a 0 otherwise. Therefore, $X_n(I_{\mathcal{P}_i^n}) = E_i^n E_i^{nT}$ where E_i^n is the submatrix of E_n containing only the columns corresponding to those arcs (u,v) such that $s_{u,v}^n \in \mathcal{P}_i^n$.
- As a consequence of the previous point, of (16) and of **Hp2**, each $X_n(I_{\mathcal{P}_i^n})$ can only have at most a constant number of columns: therefore, (9) is satisfied, i.e. $\text{rank}(X_n(I_{\mathcal{P}_i^n})) = O(1)$ holds.

- Finally, we show that (10) is satisfied. To do that, we simply have to remind that $X_n(I_{\mathcal{P}_i^n}) = E_i^n E_i^{nT}$, so that

$$[\mathbf{z}_j^n]^T X_n(I_{\mathcal{P}_i^n})\mathbf{z}_j^n = [\mathbf{e}_j^n]^T E_i^n E_i^{nT}\mathbf{e}_j^n = ||\mathbf{e}_j^{nT} E_i^n||_2^2.$$

  But $\mathbf{e}_j^{nT} E_i^n$ is just the j-th row of the submatrix E_i^n: therefore, it has nonzeroes in it only if some of the columns of E_i^n corresponds to arcs entering or leaving node j. But, by construction, the columns of E_i^n correspond to those $s_{u,v}^n \in \mathcal{P}_i^n$, and by (16) this can happen only if the corresponding arc either enters or leaves i. Therefore, $||\mathbf{e}_j^{nT} E_i^n||_2^2 \neq 0$ implies that there exist one arc entering (or leaving) i and leaving (or entering) j, i.e. that i and j are adjacent in H_n: we can now use **Hp3** to conclude that $\mathrm{dist}(\mathcal{P}_i^n, \mathcal{P}_j^n) = o(1)$ as required by (10).

- It is easy to modify the previous point to show that, if the sequence H of graphs is "strongly local", then the class of matrices generated by X is "locally X" in a strong sense.

- It is also easy to adapt the above developement to the case where $\Theta_n(a)_{i,j}$ is defined in the more general way alluded to at the end of the previous paragraph.

Therefore, we can apply the powerful results of 5.3 to the analysis of the linear systems generated by "local" graphs.

6 Acknowledgements

We are grateful to Claudio Gentile for his contribution to a part of Theorem 4.1.

References

[1] R.K. Ahuja, T.L. Magnanti and J.B. Orlin *Network Flows.* Prentice Hall, 1993.

[2] W.N. Anderson, T. Morley, "Eigenvalues of the Laplacian of a graph", *Linear and Multilin. Algebra*, **18** (1985), pp. 141–145.

[3] K.M. Anstreicher, "Linear Programming in $O(\frac{n^3}{\log n}L)$ operations " *private communication.*

[4] O. Axelsson, V. Barker, *Finite Element Solution of Boundary Value Problems, Theory and Computation.* Academic Press Inc., New York, 1984.

[5] O. Axelsson, G. Lindskög, "The rate of convergence of the preconditioned conjugate gradient method", *Numer. Math.*, **52** (1986), pp. 499–523.

[6] R. Bhatia, *Matrix Analysis.* Springer, New York, 1997.

[7] A. Böttcher, S. Grudsky, "On the condition numbers of large semidefinite Toeplitz matrices", *Linear Algebra Appl.*, **279** (1998), pp. 285–301.

[8] F. Costabile, S. Serra Capizzano, "Expansions of Komleva type and asymptotics for linear operators", *manuscript*, (1998).

[9] R. Courant, D. Hilbert, *Methods of Mathematical Physics, I.* Interscience, New York, 1953.

[10] D. Cvetkovic, M. Doob, H. Sachs, *Spectra of Graphs.* Academic Press, New York, 1979.

[11] G.H. Golub, C.F. Van Loan, *Matrix Computations.* The Johns Hopkins Univ. Press, Baltimore, 1983.

[12] C. Jordan, *Cours d'Analyse de l'Ecole Polytecnique: Vol. I.* Gauthier-Villars, Paris, France, 1909.

[13] E.H. Moore, *General Analysis. Part I.* Amer. Phil. Society, Philadelphia, 1935.

[14] Y. Nesterov, A. Nemirovskii, *Interior-point polynomial algorithms in convex programming* SIAM, Philadelphia, 1994

[15] R. Penrose, "A generalized inverse for matrices", *Proc. Cambridge Phil. Soc.*, **51** (1955), pp. 406–413.

[16] M.G.C. Resende, Geraldo Veiga "An Implementation of the dual affine scaling algorithm for minumum cost flow on bipartite uncapacited networks", *SIAM J. Optimization,* **3**(3) (1993), pp. 516–537.

[17] M.G.C. Resende, Geraldo Veiga "An Efficient Implementation of a Network Interior Point Method", *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, to appear.

[18] T. Roos, T. Terlaky and J.-Ph. Vial, *Theory and Algorithms for Linear Optimization : An Interior Point Approach.* John Wiley and Sons, Chichester, 1997.

[19] S. Serra, "On the extreme eigenvalues of Hermitian (block) Toeplitz matrices", *Linear Algebra Appl.*, **270** (1998), pp. 109–129.

[20] S. Serra Capizzano, "How bad can positive definite Toeplitz matrices be?", *Proc. Int. Conf. on Fourier Analysis and Applications*, Kuwait, may 3–8, (1998), (*Numer. Funct. Anal. Optimiz.*, under revision).

[21] S. Serra Capizzano, "An ergodic theorem for classes of preconditioned matrices", *Linear Algebra Appl.*, **282** (1998), pp. 161–183.

[22] S. Serra Capizzano, C. Tablino Possio, "Spectral and structural analysis of high order Finite Difference matrices for Elliptic Operators", *Linear Algebra Appl.*, in press.

[23] S. Serra Capizzano, "Locally X matrices, spectral distributions, preconditioning and applications", *SIAM J. Matrix Anal. Appl.*, under revision.

[24] S. Serra Capizzano, "Some theorems on linear positive operators and functionals and their applications", *TR nr. 26, LAN, Dept. of Mathematics - Univ. of Calabria*, (1997).

[25] E.E. Tyrtyshnikov, "A unifying approach to some old and new theorems on distribution and clustering", *Linear Algebra Appl.*, **232** (1996), pp. 1–43.