Supplementary material to the the paper: "The ω-Condition Number: Applications to Preconditioning and Low Rank Generalized Jacobian Updating" *†

Woosuk L. Jung‡ David Torregrosa-Belén§ Henry Wolkowicz‡

Revising as of January 29, 2025, 9:52am

Key words and phrases: $\kappa, \omega, \omega^{-2}$-condition numbers, preconditioning, generalized Jacobian, iterative methods, clustering of eigenvalues

AMS subject classifications: 15A12, 65F35, 65F08, 65G50, 49J52, 49K10, 90C32

Contents

List of Tables

*Emails resp.: w2jung@uwaterloo.ca, david.torregrosa@ua.es, hwolkowicz@uwaterloo.ca

†This report is available at URL: www.math.uwaterloo.ca/~hwolkowi/henry/reports/ABSTRACTS.html

‡Department of Combinatorics and Optimization, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada

§Centro de Modelamiento Matemático, Centro de Modelamiento Matemático (CNRS IRL2807), Universidad de Chile, Beauchef 851, Santiago, Chile

List of Figures

1 Further ω-Optimal Preconditioners

In this section we derive expressions for ω-optimal preconditioner matrices in different forms. The first one of them is a lower triangular two diagonal preconditioner. The second is a diagonal + upper triangular preconditioner. The proofs of both results proceed similarly to Claim 1 in Theorem 2.7 of the main paper. Therefore, we will not reproduce the complete proofs and limit ourselves to highlight the main steps.

1.1 Lower Triangular, Two Diagonal Preconditioning

In this section, we extend the ω-optimal diagonal scaling to an ω-optimal *lower triangular two diagonal scaling*. We define Diags_2 and $\mathrm{diags}_2 = \mathrm{Diags}_2{}^*$ in obvious ways to construct the lower triangular two diagonal matrix from a vector and its adjoint. Specifically, for a matrix $L = (L_{ij})_{i,j=1}^n \in \mathbb{R}^{n\times n}$, we get that

$$\mathrm{diags}_2(L) = \begin{pmatrix} L_{1,1} \\ L_{2,2} \\ \dots \\ L_{n,n} \\ L_{2,1} \\ L_{3,2} \\ L_{4,3} \\ \dots \\ L_{n,n-1} \end{pmatrix} =: \begin{pmatrix} \bar{l} \\ \hat{l} \end{pmatrix} \in \mathbb{R}^{n+(n-1)},$$

while, given vectors $\bar{d} = (\bar{d}_1, \dots, \bar{d}_n)^T \in \mathbb{R}^n$ and $\hat{d} = (\hat{d}_1, \dots, \hat{d}_{n-1}) \in \mathbb{R}^{n-1}$, we have

$$\mathrm{Diags}_2(\bar{d}, \hat{d}) = \begin{bmatrix} \bar{d}_1 & 0 & \dots & \dots & \dots & 0 \\ \hat{d}_1 & \bar{d}_2 & 0 & \dots & \dots & 0 \\ 0 & \hat{d}_2 & \bar{d}_3 & \vdots & \vdots & 0 \\ \vdots & \dots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & \dots & \hat{d}_{n-1} & \bar{d}_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \hat{d}_{n-1} & \bar{d}_n \end{bmatrix}.$$

Note that $\mathrm{Diags}_2 : \mathbb{R}^{2n-1} \to \mathbb{R}^{n\times n}$ and $\langle \mathrm{Diags}_2(\bar{d},\hat{d}), L\rangle = \left\langle \begin{pmatrix}\bar{d}\\ \hat{d}\end{pmatrix}, \mathrm{diags}_2(L)\right\rangle$, for any squared matrix $L \in \mathbb{R}^{n\times n}$.

Theorem 1.1. *Let $W \in \mathbb{S}^n_{++}$ and set*

$$\bar{d}^*_i = \begin{cases} \left(W_{i,i} - \frac{W^2_{i,i+1}}{W_{i+1,i+1}}\right)^{-1/2} = \left(\frac{W_{i,i}W_{i+1,i+1}-W^2_{i,i+1}}{W_{i+1,i+1}}\right)^{-1/2}, & \text{if } i \in [n-1];\\ W_{n,n}^{-1/2}, & \text{if } i = n \end{cases}$$

and

$$\hat{d}^*_i = -\frac{W_{i,i+1}}{W_{i+1,i+1}}\bar{d}^*_i, \quad i \in [n-1].$$

Then the ω-optimal lower triangular two diagonal scaling of W is given by

$$(\bar{d}^*, \hat{d}^*) = \operatorname*{argmin}_{(\bar{d},\hat{d})\in\mathbb{R}^n_{++}\times\mathbb{R}^{n-1}} \omega(\bar{d},\hat{d}), \tag{1.1}$$

where $\omega(\bar{d},\hat{d}) := \omega\left(\mathrm{Diags}_2(\bar{d},\hat{d})^T W\, \mathrm{Diags}_2(\bar{d},\hat{d})\right)$.

Proof. First we note, since the 2×2 principal minors for $W \succ 0$ are all positive, the definitions of the optimal d^* are well defined. Let $\bar{d} \in \mathbb{R}^n_{++}$ and $\hat{d} \in \mathbb{R}^{n-1}$. Define the ω-condition number, f and g as functions of a pair $(\bar{d},\hat{d}) \in \mathbb{R}^n_{++}\times\mathbb{R}^{n-1}$. This is

$$\omega(\bar{d},\hat{d}) = \frac{f(\bar{d},\hat{d})}{g(d,\hat{d})} := \frac{\mathrm{tr}\big(\mathrm{Diags}_2(\bar{d},\hat{d})^T W\,\mathrm{Diags}_2(\bar{d},\hat{d})\big)/n}{\det(W)^{1/n}\prod_{i=1}^n(\bar{d}_i)^{2/n}}.$$

Differentiating the pseudoconvex ω and equating to 0, we get the optimality condition

$$(\mathrm{diags}_2\, W\, \mathrm{Diags}_2)\,(\bar{d},\hat{d}) = \begin{pmatrix}\bar{d}^{-1}\\ 0_{n-1}\end{pmatrix} \tag{1.2}$$

Solving (1.2) for $(\bar{d},\hat{d})$, results in

$$\bar{d}_i = \begin{cases} \left(W_{i,i} - \frac{W^2_{i,i+1}}{W_{i+1,i+1}}\right)^{-1/2} = \left(\frac{W_{i,i}W_{i+1,i+1}-W^2_{i,i+1}}{W_{i+1,i+1}}\right)^{-1/2}, & \text{if } i \in [n-1];\\ W_{n,n}^{-1/2}, & \text{if } i = n; \end{cases}$$

and

$$\hat{d}_i = -\frac{W_{i,i+1}}{W_{i+1,i+1}}\bar{d}_i, \quad i \in [n-1].$$

■

1.2 Upper Triangular D_{+k} Diagonal Preconditioning

We note that the ω-optimal lower triangular two diagonal preconditioner in Theorem 1.1 is sparse but its inverse though still lower triangular is not necessarily as sparse, i.e., the two diagonal structure can be lost completely, sparsity can be lost. We now consider the diagonal with upper triangular elements that maintain the same structure in the inverse, i.e., maintain sparsity for the inverse. Recall that the triangular number $t(k) = k(k+1)/2$ and define the transformation $D_{+k} : \mathbb{R}^{n+t(k)} \to \mathbb{R}^{n\times n}$:

$$
\begin{array}{rcl}
D_{+k}(d,\alpha) & = & \operatorname{Diag}(d) + \left[\begin{bmatrix} 0_{n\times n-k} \end{bmatrix} \;\middle|\; \begin{bmatrix} \begin{bmatrix} \operatorname{Triu}(\alpha) \end{bmatrix} \\ \begin{bmatrix} 0_{n-k\times k} \end{bmatrix} \end{bmatrix} \right] \\
& = & \operatorname{Diag}(d) + \operatorname{Triu_k}(\alpha) = \begin{bmatrix} \operatorname{Diag} & \operatorname{Triu_k} \end{bmatrix} \begin{pmatrix} d \\ \alpha \end{pmatrix} \\
& = & \begin{pmatrix}
d_1 & 0 & \dots & 0 & \dots & \alpha_{1,n-k+1} & \alpha_{1,n-k+2} & \dots & \alpha_{1,n} \\
0 & d_2 & \dots & 0 & \dots & 0 & \alpha_{2,n-k+2} & \dots & \alpha_{2,n} \\
\vdots & \vdots & \ddots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \dots & d_k & \dots & 0 & 0 & 0 & \alpha_{k,n} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \dots & 0 & \dots & d_{n-k+1} & 0 & 0 & 0 \\
0 & 0 & \dots & 0 & \dots & 0 & d_{n-k+2} & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 & d_n
\end{pmatrix}
\end{array}
\tag{1.3}
$$

where $d \in \mathbb{R}^n$ and $\alpha := (\alpha_{1,n-k+1}, \alpha_{1,n-k+2}, \alpha_{2,n-k+2}, \dots, \alpha_{1,n}, \dots, \alpha_{k,n})^T \in \mathbb{R}^{t(k)}$. Then the optimal upper triangular $D_{+k}(d,\alpha)$ diagonal preconditioner is given by solving the following optimization problem:

$$
(\bar{d},\bar{\alpha}) := \operatorname*{argmin}_{(d,\alpha)\in\mathbb{R}^n_{++}\times\mathbb{R}^{t(k)}} \omega\big(D_{+k}(d,\alpha)^T W D_{+k}(d,\alpha)\big). \tag{1.4}
$$

Theorem 1.2. *Let $W \in \mathbb{S}^n_{++}$ be given and let $(\bar{d},\bar{\alpha}) \in \mathbb{R}^{n+t(k)}$ such that*

$$
\bar{d}_i = W_{i,i}^{-1/2}, \quad i \in [n-k] \tag{1.5}
$$

and the following hold for each $i \in [n-k+1,n]$:

$$
\begin{array}{rcl}
W_{i,i}\bar{d}_i + \sum_{\ell=1}^{i-n+k} \bar{\alpha}_{\ell,i} W_{\ell,i} & = & 1/\bar{d}_i, \\
W_{i,j}\bar{d}_i + \sum_{\ell=1}^{i-n+k} \bar{\alpha}_{\ell,i} W_{\ell,j} & = & 0, \qquad j \in [i-n+k].
\end{array}
\tag{1.6}
$$

Then, $(\bar{d},\bar{\alpha})$ is the optimal solution of (1.4).

Proof. Define the transformations (isometries) $\operatorname{Triu} : \mathbb{R}^{t(k)} \to \mathbb{R}^{k\times k}$ and $\operatorname{Triu_k} : \mathbb{R}^{t(k)} \to \mathbb{R}^{n\times n}$ according to (1.3). We denote the adjoints by triu and $\operatorname{triu_k}$, respectively, and note that

$$
\operatorname{triu}^\dagger = \operatorname{triu}^*, \ \operatorname{Triu}^\dagger = \operatorname{Triu}^*.
$$

Hence,

$$\begin{aligned} D_{+k}(d,\alpha) &= \operatorname{Diag}(d) + \mathrm{Triu_k}(\alpha) \\ &= \begin{bmatrix} \operatorname{Diag} & \mathrm{Triu_k} \end{bmatrix} \begin{pmatrix} d \\ \alpha \end{pmatrix}. \end{aligned}$$

Denote

$$\begin{aligned} \omega_k(d,\alpha) &:= \omega\Big(D_{+k}(d,\alpha)^T W D_{+k}(d,\alpha)\Big) \\ &= \frac{\operatorname{tr}\left(D_{+k}(d,\alpha)^T W D_{+k}(d,\alpha)\right)/n}{\det\left(D_{+k}(d,\alpha)^T W D_{+k}(d,\alpha)\right)^{1/n}} \\ &= \frac{\operatorname{tr}\left(D_{+k}(d,\alpha)^T W D_{+k}(d,\alpha)\right)}{\det(W)^{1/n}\prod_{i=1}^n d_i^{2/n}}. \end{aligned}$$

For the numerator of ω_k we use

$$\begin{aligned} f(d,\alpha) &:= \tfrac{1}{n}\operatorname{tr}\left(D_{+k}(d,\alpha)^T W D_{+k}(d,\alpha)\right) \\ &= \tfrac{1}{n}\Big\langle D_{+k}(d,\alpha), W D_{+k}(d,\alpha)\Big\rangle \\ &= \tfrac{1}{n}\left\langle \begin{pmatrix} d \\ \alpha \end{pmatrix}, D_{+k}^*\Big(W D_{+k}(d,\alpha)\Big)\right\rangle \\ &= \tfrac{1}{n}\begin{pmatrix} d \\ \alpha \end{pmatrix}^T D_{+k}^*\Big(W D_{+k}(d,\alpha)\Big) \\ &= \tfrac{1}{n}\begin{pmatrix} d \\ \alpha \end{pmatrix}^T \begin{bmatrix} \operatorname{diag} \\ \mathrm{triu_k} \end{bmatrix}\Big(W D_{+k}(d,\alpha)\Big) \\ &= \tfrac{1}{n}\begin{pmatrix} d \\ \alpha \end{pmatrix}^T \begin{bmatrix} \operatorname{diag} W\Big(\operatorname{Diag}(d) + \mathrm{Triu_k}(\alpha)\Big) \\ \mathrm{triu_k}\, W\Big(\operatorname{Diag}(d) + \mathrm{Triu_k}(\alpha)\Big) \end{bmatrix} \\ &= \tfrac{1}{n}\begin{pmatrix} d \\ \alpha \end{pmatrix}^T \begin{bmatrix} \operatorname{diag} W \operatorname{Diag} & \operatorname{diag} W\, \mathrm{Triu_k} \\ \mathrm{triu_k}\, W \operatorname{Diag} & \mathrm{triu_k}\, W\, \mathrm{Triu_k} \end{bmatrix} \begin{pmatrix} d \\ \alpha \end{pmatrix}. \end{aligned}$$

and the gradient is therefore

$$\nabla f(d,\alpha) = \frac{2}{n}\begin{bmatrix} \operatorname{diag} W \operatorname{Diag} & \operatorname{diag} W\, \mathrm{Triu_k} \\ \mathrm{triu_k}\, W \operatorname{Diag} & \mathrm{triu_k}\, W\, \mathrm{Triu_k} \end{bmatrix} \begin{pmatrix} d \\ \alpha \end{pmatrix}.$$

The denominator of ω_k is

$$g(d,\alpha) := \det(W)^{1/n}\prod_{i=1}^n d_i^{2/n}$$

and thus

$$\nabla g(d,\alpha) = \frac{2}{n} g(d,\alpha) \begin{pmatrix} 1/d_1 \\ 1/d_2 \\ \vdots \\ 1/d_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

For simplicity, denote $\bar{d}^{-1} := (1/\bar{d}_1, 1/\bar{d}_2, \ldots, 1/\bar{d}_n)^T \in \mathbb{R}^n$. Then,

$$\begin{aligned}
\nabla\omega_k(d,\alpha) &= \tfrac{1}{g(d,\alpha)^2}\big(g(d,\alpha)\nabla f(d,\alpha) - f(d,\alpha)\nabla g(d,\alpha)\big) \\
&= \tfrac{1}{g(d,\alpha)}\left(\nabla f(d,\alpha) - \tfrac{2}{n} f(d,\alpha)\begin{pmatrix} d^{-1} \\ 0_{t(k)} \end{pmatrix}\right).
\end{aligned}$$

Finally, the proof follows from noticing that

$$\begin{aligned}
(\bar{d},\bar{\alpha}) \text{ satisfies } (1.5) \text{ and } (1.6) \quad &\iff \quad \tfrac{n}{2}\nabla f(\bar{d},\bar{\alpha}) = \begin{pmatrix} \bar{d}^{-1} \\ 0_{t(k)} \end{pmatrix} \\
&\implies \quad f(\bar{d},\bar{\alpha}) = 1.
\end{aligned}$$

Hence, (1.5) and (1.6) implies $\nabla\omega_k(\bar{d},\bar{\alpha}) = 0$, i.e., $(\bar{d},\bar{\alpha})$ is optimal. ∎

The following Example 1.3 and Example 1.4 solve (1.6) for $k = 1$ and $k = 2$.

Example 1.3 ($k = 1$)**.** *Let $W \in \mathbb{S}^n_{++}$ be given. Set*

$$\bar{d}_i = \begin{cases} W_{i,i}^{-1/2}, & \text{if } i \in [n-1] \\ \left(\frac{W_{1,1}W_{n,n} - W_{1,n}^2}{W_{1,1}}\right)^{-1/2}, & \text{if } i = n. \end{cases}$$

and

$$\bar{\alpha} = -\frac{W_{1n}}{W_{11}}\bar{d}_n.$$

Then the optimal D_{+1}-diagonal upper triangular scaling is given by

$$(\bar{d},\bar{\alpha}) = \operatorname*{argmin}_{d\in\mathbb{R}^n_{++},\alpha\in\mathbb{R}} \omega\big(D_{+1}(d,\alpha)^T W D_{+1}(d,\alpha)\big).$$

Example 1.4 ($k = 2$)**.** *Let $W \in \mathbb{S}^n_{++}$ be given. Set*

$$\bar{d}_i = \begin{cases} W_{i,i}^{-1/2}, & \text{if } i \in [n-2] \\ \left(\frac{W_{1,1}W_{n-1,n-1} - W_{1,n-1}^2}{W_{1,1}}\right)^{-1/2}, & \text{if } i = n-1 \\ \left(W_{n,n} + \frac{W_{1,n}^2 W_{2,2} - 2W_{1,n}W_{2,n}W_{1,2} + W_{2,n}^2 W_{1,1}}{W_{1,2}^2 - W_{1,1}W_{2,2}}\right)^{-1/2}, & \text{if } i = n. \end{cases}$$

$$\begin{aligned}
\bar{\alpha}_{1,n} &= \left(\tfrac{W_{1,n}W_{2,2} - W_{1,2}W_{2,n}}{W_{1,2}^2 - W_{1,1}W_{2,2}}\right)\bar{d}_n, \\
\bar{\alpha}_{1,n-1} &= -\tfrac{W_{1,n-1}}{W_{1,1}}\bar{d}_{n-1}, \\
\bar{\alpha}_{2,n} &= \left(\tfrac{W_{1,1}W_{2,n} - W_{1,2}W_{1,n}}{W_{1,2}^2 - W_{1,1}W_{2,2}}\right)\bar{d}_n.
\end{aligned}$$

Then the optimal D_{+2}-diagonal upper triangular scaling is given by

$$(\bar{d},\bar{\alpha}) = \operatorname*{argmin}_{d\in\mathbb{R}^n_{++},\alpha\in\mathbb{R}^3} \omega\big(D_{+2}(d,\alpha)^T W D_{+2}(d,\alpha)\big).$$

2 Further empirical results

2.1 Clustering of eigenvalues for sparse test matrices

In this section, we provide further numerical results to illustrate the advantage of the ω-optimal diagonal preconditioner over the κ-optimal diagonal preconditioner. We use 5 distinct test matrices from the SuiteSparse Matrix Collection of sizes between 100 and 150. We do not use larger matrices as the computation of the κ-optimal diagonal preconditioner is very inefficient. For this matrices we compare the efficacy of the optimal ω and κ diagonal preconditioners for (1) cluster the eigenvalues and (2) reducing the number of iterations of `Matlab`'s preconditioner conjugate gradient. For (1) we perform the same experiment than in Section 2.2 of the main paper and for (2) we replicate the empirics in Section 4.1 of the main paper. The results for (1) are illustrated in Figures 2.1 to 2.5 while the conclusions from (2) are presented in Table 2.1.

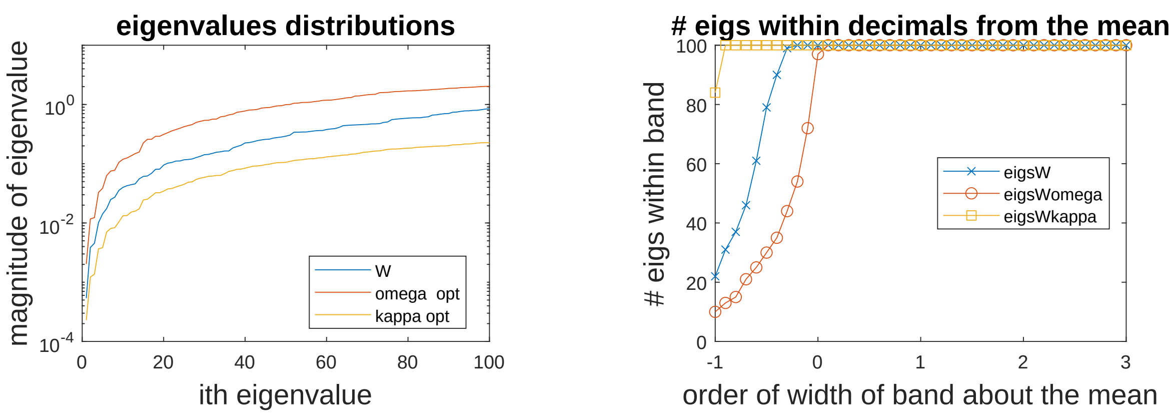

Figure 2.1: diagonal preconditioners: distribution of eigenvalues for matrix `nos4`

name	n	$nnz(W)$	None	Omega	Kappa
`nos4`	100	594	76	68	70
`bcsstk03`	112	604	569	132	141
`bcsstk04`	132	3,648	550	67	86
`lund_a`	147	2,449	343	85	93
`Trefethen_150`	150	2,040	95	9	14

Table 2.1: diagonal preconditioners: number of iterations

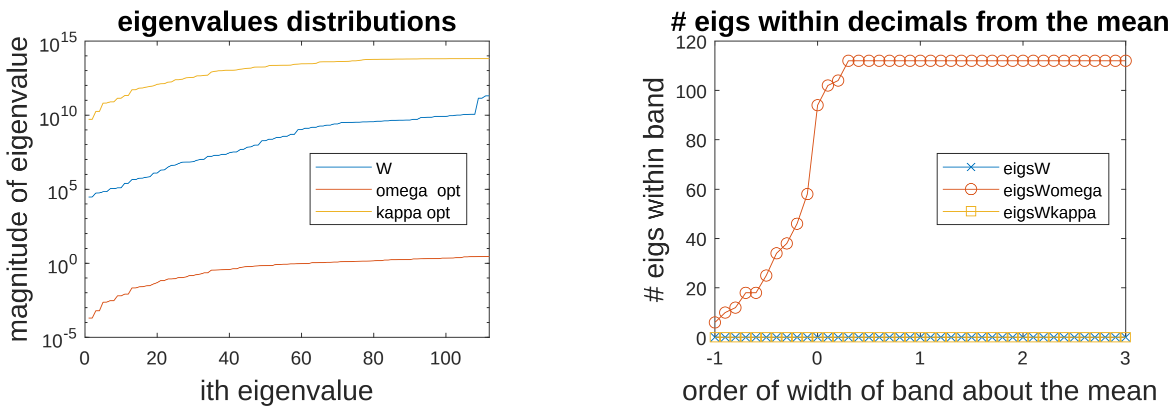

Figure 2.2: diagonal preconditioners: distribution of eigenvalues for matrix `bcsstk03`

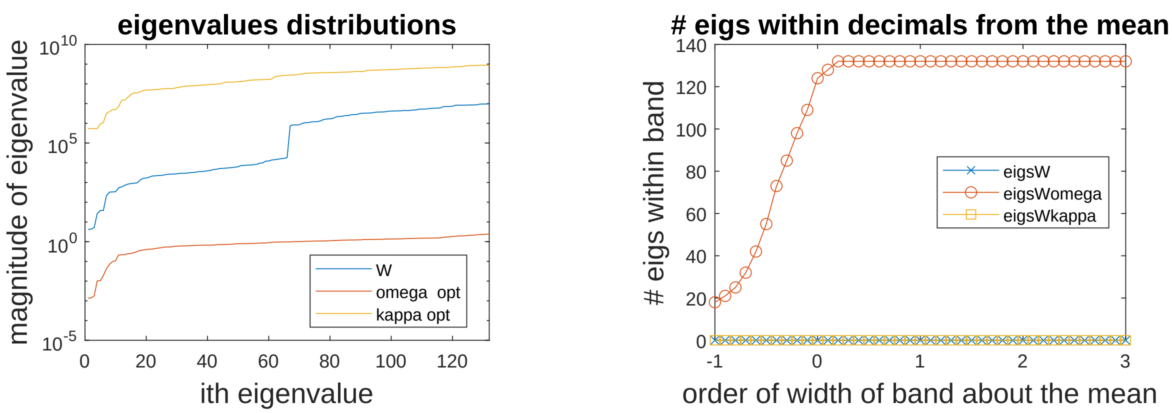

Figure 2.3: diagonal preconditioners: distribution of eigenvalues for matrix `bcsstk04`

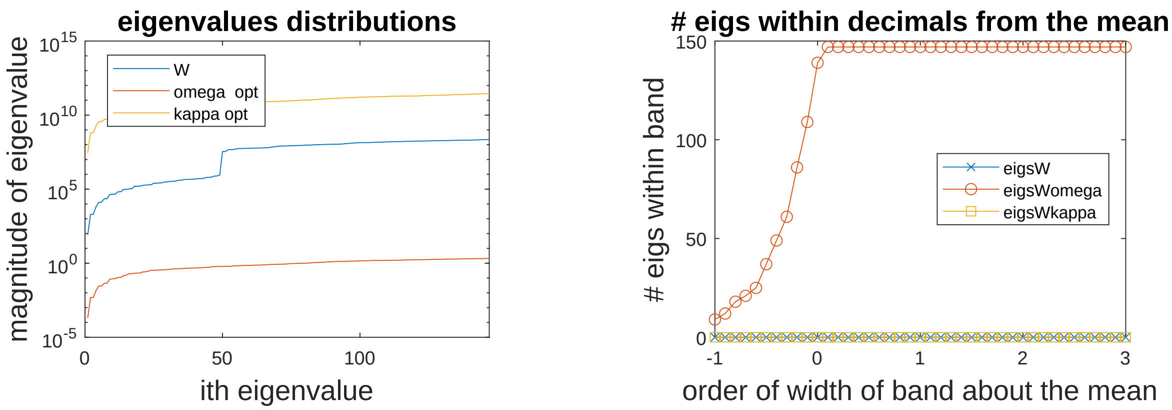

Figure 2.4: diagonal preconditioners: distribution of eigenvalues for matrix `lund_a`

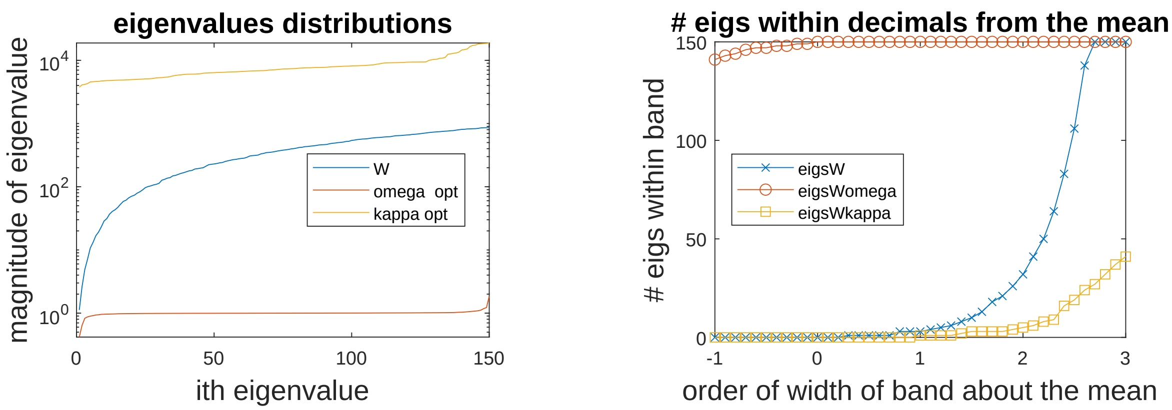

Figure 2.5: diagonal preconditioners: distribution of eigenvalues for matrix `Trefethen_150`

Index