- Chapter 1 -

Periodic Forward-Backward Solver

We consider the following optimization problem,

$$\underset{\boldsymbol{x}\in\mathbb{R}^p}{\text{minimize}}\ \ f(\boldsymbol{x}) + \frac{1}{2}\|\boldsymbol{H}\boldsymbol{x}-\boldsymbol{y}\|^2, \tag{1.1}$$

where $f \in \Gamma_0(\mathbb{R}^p)$, $\boldsymbol{y} \in \mathbb{R}^q$ and $\boldsymbol{H} \in \mathbb{R}^{q\times p}$ is not the zero matrix. Algorithm 1 corresponds to an m-periodic forward-backward algorithm. We want to derive conditions on the steps $(\gamma_i)_{1\leq i\leq m}$ that guarantee the convergence of this algorithm.

Algorithm 1: m-periodic forward-backward algorithm

Initialization: Let $\boldsymbol{x}_1 \in \mathbb{R}^p$.
for $n = 1, 2, \dots$ **do**
  $\boldsymbol{x}_{0,n} = \boldsymbol{x}_n$;
  for $i \in \{1, \dots, m\}$ **do**
    $\boldsymbol{x}_{i,n} = \text{prox}_{\gamma_i f}\left(\boldsymbol{x}_{i-1,n} - \gamma_i \boldsymbol{H}^\top\left(\boldsymbol{H}\boldsymbol{x}_{i-1,n} - \boldsymbol{y}\right)\right)$;
  end
  $\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n\left(\boldsymbol{x}_{m,n} - \boldsymbol{x}_n\right)$;
end

§ 1.1 General notation

For every $i \in \{1, \dots, m\}$, we define the following objects,

$$\mathcal{R}_i = \text{prox}_{\gamma_i f},\ \ \boldsymbol{W}_i = \boldsymbol{I}_p - \gamma_i \boldsymbol{H}^\top\boldsymbol{H},\ \ b_i = \gamma_i\boldsymbol{H}^\top\boldsymbol{y} \text{ and } \mathcal{T}_i = \mathcal{R}_i(\boldsymbol{W}_i\cdot + b_i).$$

We also introduce $\mathcal{F}$, the set of fixed points of $\mathcal{T}_m \circ\cdots\circ \mathcal{T}_1$ as follows,

$$\mathcal{F} = \{\boldsymbol{x} \in \mathbb{R}^p \mid (\mathcal{T}_m\circ\cdots\circ\mathcal{T}_1)(\boldsymbol{x}) = \boldsymbol{x}\}.$$

We define the following variational inequality problem which will be used in the convergence analysis of the proposed algorithm.

Problem 1.1.1 Find $(\overline{\boldsymbol{x}}_1, \ldots, \overline{\boldsymbol{x}}_m) \in (\mathbb{R}^p)^m$ such that the following system of variational inequalities is satisfied.

$$\begin{cases} b_1 \in \overline{\boldsymbol{x}}_1 - \boldsymbol{W}_1 \overline{\boldsymbol{x}}_m + \gamma_1 \partial f(\overline{\boldsymbol{x}}_1) \\ b_2 \in \overline{\boldsymbol{x}}_2 - \boldsymbol{W}_2 \overline{\boldsymbol{x}}_1 + \gamma_2 \partial f(\overline{\boldsymbol{x}}_2) \\ \vdots \\ b_m \in \overline{\boldsymbol{x}}_m - \boldsymbol{W}_m \overline{\boldsymbol{x}}_{m-1} + \gamma_m \partial f(\overline{\boldsymbol{x}}_m) \end{cases} \tag{1.2}$$

Furthermore, we denote by $\boldsymbol{W}$ the following operator $\boldsymbol{W} = \boldsymbol{W}_m \circ \cdots \circ \boldsymbol{W}_1$. It is worth noting that all matrices $(\boldsymbol{W}_i)_{1 \leq i \leq m}$ are diagonalizable in the same basis, hence they commute. Therefore, $\boldsymbol{W}$ is diagonalizable and its eigen values are the products of the eigen values of $(\boldsymbol{W}_i)_{1 \leq i \leq m}$ in the aforementioned basis. The maximal and minimal eigen values of $\boldsymbol{H}^\top \boldsymbol{H}$ are noted β_+ and β_-, respectively. If $\beta_- > 0$, then the its condition number is denoted by $\kappa = \frac{\beta_+}{\beta_-}$. We also refer to the upperbound for the step in the classic forward-backward algorithm as $\overline{\gamma}$, which is defined as follows,

$$\overline{\gamma} = \frac{2}{\beta_+}.$$

Similarly, we denote by η_+ and η_- the maximal and minimal eigen values of $\boldsymbol{W}$.

In addition, we set $\theta_0 = 1$, and for every $i \in \{1, \ldots, m\}$ we define the following quantity,

$$\begin{aligned} \theta_i = \|\boldsymbol{W}_i \circ \cdots \circ \boldsymbol{W}_1\| + \sum_{k=1}^{i-1} \sum_{1 \leq j_1 < \ldots < j_k \leq i-1} \|\boldsymbol{W}_i \circ \cdots \circ \boldsymbol{W}_{j_k+1}\| \\ \times \|\boldsymbol{W}_{j_k} \circ \cdots \circ \boldsymbol{W}_{j_{k-1}+1}\| \cdots \|\boldsymbol{W}_{j_1} \circ \cdots \circ \boldsymbol{W}_1\|. \end{aligned} \tag{1.3}$$

For each setting of the hyperparameters of Algorithm 1, we will study the requirements under which the condition below is satisfied. Our interest in (1.4) is justified by Remark 1.1.3.

Condition 1.1.2 Let $\theta_0 = 1$ and $(\forall i \in \{1, \ldots, m\})$ θ_i be defined as in (1.3). Then, there exists $\alpha \in \left[\frac{1}{2}, 1\right[$ such that the following inequality is satisfied.

$$\|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| - \|\boldsymbol{W}\| + 2\theta_m \leq 2^m \alpha. \tag{1.4}$$

Remark 1.1.3 It is worth noting that, from [Combettes and Pesquet, 2018, Proposition 3.6], if Condition 1.1.2 holds, then [Combettes and Pesquet, 2018, Condition 3.1], which is required in [Combettes and Pesquet, 2018, Theorem 4.7], is satisfied.

§ 1.2 Preliminary results

As recalled below, the sequence $(\theta_i)_{1\leq \leq m}$ can be computed recursively.

Proposition 1.2.1 *[Combettes and Pesquet, 2018, Lemma 3.3(i)] For every $i \in \{1,\ldots,m\}$, let $\boldsymbol{W}_i \in \mathbb{R}^{p\times p}$ and let θ_i be defined as in (1.3). Then, $(\forall i \in \{1,\ldots,m\})$ $\theta_i = \sum_{k=0}^{i-1} \theta_k \|\boldsymbol{W}_i \circ \cdots \circ \boldsymbol{W}_{k+1}\|$.*

We also give explicit expressions for the norm of some matrices, which will be useful in the rest of our study.

Proposition 1.2.2 *Let $\alpha \in \left[\frac{1}{2}, 1\right[$. Using the notation of Section 1.1, we have*

$$(\forall i \in \{1,\ldots,m\}) \quad \|\boldsymbol{W}_i\| = \max\left\{1-\gamma_i\beta_-, \gamma_i\beta_+ - 1\right\},$$

and

$$\|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| = \max\left\{\eta_+ - 2^m(1-\alpha), 2^m(1-\alpha) - \eta_-\right\}.$$

Proof. Results simply follow from the fact that $(\forall i \in \{1,\ldots,m\})$ $\boldsymbol{W}_i$ and $\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p$ are symmetric, so their norms can be expressed in terms of their extremal eigen values. □

§ 1.3 Constant stepsize

In this section we study Algorithm 1 under the following assumption.

Assumption 1.3.1 *There exists $\gamma \in \mathbb{R}_+^*$ such that $(\forall i \in \{1,\ldots,m\})$ $\gamma_i = \gamma$.*

When Assumption 1.3.1 is satisfied, we denote by $\boldsymbol{W}_0$ the operator $\boldsymbol{I}_p - \gamma\boldsymbol{H}^\top\boldsymbol{H}$, i.e. $(\forall i \in \{1,\ldots,m\})$ $\boldsymbol{W}_i = \boldsymbol{W}_0$. Similarly, we let $\mathcal{R}_0 = \mathrm{prox}_{\gamma f}$, $b_0 = \gamma\boldsymbol{H}^\top\boldsymbol{y}$ and $\mathcal{T}_0 = \mathcal{R}_0(\boldsymbol{W}_0 \cdot + b_0)$. First, we derive preliminary results in Propositions 1.3.2 and 1.3.3 below, that are related to the computation of θ_m and of $\|\boldsymbol{W}\|$.

Proposition 1.3.2 *Under Assumption 1.3.1 and using notation from Section 1.1, the following holds,*

(i) $\|\boldsymbol{W}\| = \|\boldsymbol{W}_0\|^m$,

(ii) $(\forall i \in \{1,\ldots,m\})$ $\theta_i = 2^{i-1}\|\boldsymbol{W}_0\|^i$.

Proof. Result (i) simply follows from the definition of $\boldsymbol{W}$ and from the fact that $\boldsymbol{W}_0$ is symmetric. We show result (ii) by induction. From Proposition 1.2.1, the result is true for $i = 1$: $\theta_1 = \|\boldsymbol{W}_0\| = 2^{1-1}\|\boldsymbol{W}_0\|^1$. If $m \geq 2$, take $i \in \{2, \ldots, m\}$ and assume that the result holds for every $j \in \{1, \ldots, i-1\}$. Then,

$$\theta_i = \|\boldsymbol{W}_0\|^i + \sum_{k=1}^{i-1} 2^{k-1} \|\boldsymbol{W}_0\|^k \|\boldsymbol{W}_0^{i-k}\|. \tag{1.5}$$

Since $\boldsymbol{W}_0$ is symmetric, (1.5) becomes

$$\begin{aligned}\theta_i &= \|\boldsymbol{W}_0\|^i + \textstyle\sum_{k=1}^{i-1} 2^{k-1}\|\boldsymbol{W}_0\|^i \\ &= \|\boldsymbol{W}_0\|^i + \|\boldsymbol{W}_0\|^i \left(\tfrac{1-2^{i-1}}{1-2}\right) \\ &= 2^{i-1}\|\boldsymbol{W}_0\|^i,\end{aligned}$$

which completes the proof. □

Proposition 1.3.3 *Under Assumption 1.3.1 and using notation from Section 1.1, the following statements are true.*

(i) If m is odd, then $\eta_- = (1-\gamma\beta_+)^m$ and $\eta_+ = (1-\gamma\beta_-)^m$.

(ii) If m is even, then $\eta_+ = (\max\{1-\gamma\beta_-, \gamma\beta_+ - 1\})^m$.

Proof. Results simply follow from the variations of $\eta \mapsto \eta^m$, which is increasing on $\mathbb{R}$ if m is odd; and which is decreasing on $]-\infty, 0]$ and increasing on $[0, +\infty[$ if m is even. □

We can show now that Condition 1.1.2 is satisfied if and only if the constant step is below an certain bound.

Proposition 1.3.4 *We assume that Assumption 1.3.1 holds and we use notation from Section 1.1. Then, Condition 1.1.2 holds if and only if one of the following conditions is satisfied,*

(i) m is odd and $\gamma < \overline{\gamma}$,

(ii) m is even and $\gamma \leq \overline{\gamma}$.

Proof. First, we show by contraposition that (i) and (ii) are necessary conditions for Condition 1.1.2. The triangular inequality leads to

$$(\forall \alpha \in [1/2, 1[) \quad \|\boldsymbol{W}\| \leq \|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| + 2^m(1-\alpha). \tag{1.6}$$

From (1.6) and Proposition 1.3.2, we deduce that

$$\|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| - \|\boldsymbol{W}\| + 2\theta_m \geq 2^m(\|\boldsymbol{W}_0\|^m - 1) + 2^m\alpha \tag{1.7}$$

If $\gamma > \overline{\gamma}$, then $\gamma > \frac{2}{\beta_+ + \beta_-}$ and $\|\boldsymbol{W}_0\| = \gamma\beta_+ - 1 > 1$. Hence, from (1.7), we deduce that Condition 1.1.2 cannot hold. By contraposition, Condition 1.1.2 implies that $\gamma \leq \overline{\gamma}$, which leads to the necessary condition (ii).
Now, let us show that the inequality is strict when m is odd. Assume that $\gamma = \overline{\gamma}$, then $\|\boldsymbol{W}_0\| = 1$. Thus, if Condition 1.1.2 holds, then we have an equality in (1.6) and

$$\|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| = 1 - 2^m(1-\alpha). \tag{1.8}$$

If m is odd, then, from Proposition 1.3.3 (i), $1 = -(1-\gamma\beta_+)^m = -\eta_-$, and (1.8) can be rewritten as

$$\|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| = -\eta_- - 2^m(1-\alpha).$$

Thus, from Proposition 1.2.2, we deduce that

$$\begin{aligned} -\eta_- - 2^m(1-\alpha) &\geq 2^m(1-\alpha) - \eta_- && (1.9)\\ \alpha &\geq 1. && (1.10) \end{aligned}$$

Hence, Condition 1.1.2 cannot hold since it requires $\alpha \in \left[\frac{1}{2}, 1\right[$, which leads to the necessary condition (i).
Second, we show that (i) and (ii) are sufficient conditions for Condition 1.1.2.

(i) Assume that $\gamma < \overline{\gamma}$ and that $\beta_- > 0$. Then, from the definition of $\overline{\gamma}$ and from Proposition 1.2.2, we deduce that $\|\boldsymbol{W}_0\| < 1$. Hence, $\underline{\alpha} < 1$, where $\underline{\alpha} = \frac{1+\|\boldsymbol{W}_0\|^m}{2}$. On the other hand, the triangular inequality leads to

$$\|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| - \|\boldsymbol{W}\| \leq 2^m(1-\alpha). \tag{1.11}$$

Plugging Proposition 1.3.2 into (1.11) leads to

$$\begin{aligned} \|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| - \|\boldsymbol{W}\| + 2\theta_m &\leq 2^m(1-\alpha+\|\boldsymbol{W}_0\|^m)\\ &\leq 2^m(2\underline{\alpha} - \alpha). \end{aligned}$$

Hence, for every $\alpha \in \left[\frac{1}{2}, 1\right[$ such that $\underline{\alpha} \leq \alpha$, (1.4) holds and Condition 1.1.2 is satisfied.

(ii) Now, we assume that $\gamma < \overline{\gamma}$ and that $\beta_- = 0$. From the definition of $\overline{\gamma}$ we have $\gamma\beta_+ - 1 < 1 = 1 - \gamma\beta_-$. Thus, Proposition 1.2.2 leads to $\|\boldsymbol{W}_0\| = 1$ and Proposition 1.3.3 gives $\eta_+ = 1$. If m is odd, then $-(1-\gamma\beta_+)^m < 1$, which leads to $-\eta_- < \eta_+$ (see Proposition 1.3.3 (i)). If m is even, then, by definition of $\boldsymbol{W}$, all of its eigen values are positive and, in particular, $\eta_- \geq 0$. In both cases, we have $\eta_- + \eta_+ > 0$. Hence, there exists $\alpha \in \left[\frac{1}{2}, 1\right[$ such that $\eta_- + \eta_+ > 2^{m+1}(1-\alpha)$. With this α, from Proposition 1.2.2, we have $\|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| = \eta_+ - 2^m(1-\alpha)$. Thus,

$$
\begin{aligned}
\|\boldsymbol{W} - 2^m(1-\alpha)\boldsymbol{I}_p\| - \|\boldsymbol{W}\| + 2\theta_m &= 1 - 2^m(1-\alpha) - 1 + 2^m \\
&= 2^m\alpha,
\end{aligned}
$$

and Condition 1.1.2 is satisfied.

(iii) Assume that m is even, then $\eta_- \geq 0$. In addition, let $\gamma = \overline{\gamma}$. Since $\beta_- \geq 0$, by definition of $\overline{\gamma}$ we have $\gamma\beta_+ - 1 = 1 \geq 1 - \gamma\beta_-$. Hence, $\|\boldsymbol{W}_0\| = 1$ and $\eta_+ = 1$. Therefore, like in the previous setting, we have $\eta_+ + \eta_- > 0$. Similarly, we find α in $\left[\frac{1}{2}, 1\right[$ such that Condition 1.1.2 holds.

The proof is complete. □

Remark 1.3.5 Unlike for the classic forward-backward algorithm, when m is even, the upper bound $\overline{\gamma}$ is included. It is worth noting that, for this configuration ($\gamma = \overline{\gamma}$ and m is even), the lower bound for α is $1 - \frac{\eta_- + 1}{2^{m+1}}$, which is always strictly lower than 1, and which converges to 1 as $m \to +\infty$.

We state our main theorem below.

Theorem 1.3.6 *Under Assumption 1.3.1 and using notation from Section 1.1, if one of the following condition holds,*

(i) m is odd and $\gamma < \overline{\gamma}$,

(ii) m is even and $\gamma \leq \overline{\gamma}$,

then $(\mathcal{T}_0)^m$ is α-averaged for some $\alpha \in \left[\frac{1}{2}, 1\right[$ and, if in addition the following conditions are satisfied,

(iii) there exists a solution to problem (1.1),

(iv) $(\forall n \in \mathbb{N})$ $\lambda_n \in \left]0, \frac{1}{\alpha}\right[$ and $\sum_{n\in\mathbb{N}} \lambda_n(1 - \alpha\lambda_n) = +\infty$,

then $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ converges to a point $\overline{\boldsymbol{x}}_m \in \mathcal{F}$, $(\forall i \in \{1, \ldots, m-1\})$ $(\boldsymbol{x}_{i,n})_{n\in\mathbb{N}}$ converges to $\overline{\boldsymbol{x}}_i = (\mathcal{T}_0)^i(\overline{\boldsymbol{x}}_m)$, and $(\overline{x}_1, \ldots, \overline{x}_m)$ is a solution to Problem 1.1.1.

Proof. If condition (i) or (ii) holds, then we deduce from Proposition 1.3.4 that Condition 1.1.2 is satisfied. Hence, we can apply [Combettes and Pesquet, 2018, Proposition 3.6] which states that [Combettes and Pesquet, 2018, Condition 3.1] holds for $\alpha \in \left[\frac{1}{2}, 1\right[$. Then, it follows from [Combettes and Pesquet, 2018, Theorem 3.8] that $(\mathcal{T}_0)^m$ is α-averaged. If, in addition to (i) or (ii), conditions (iii) and (iv) are satisfied, then we can apply [Combettes and Pesquet, 2018, Theorem 4.7] which leads to the result since $\mathbb{R}^p$ is a finite-dimensional space. □

From Theorem 1.3.6 we deduce Corollary 1.3.7 below, which provides a stronger result related to the solutions to Problem 1.1.

Corollary 1.3.7 *Under Assumption 1.3.1 and using notation from Section 1.1, if the following conditions are satisfied,*

(i) $\gamma < \overline{\gamma}$,

(ii) $\boldsymbol{H}^\top \boldsymbol{H}$ *is invertible, i.e.* $\beta_- > 0$,

(iii) there exists a solution to problem (1.1),

(iv) $(\forall n \in \mathbb{N})\ \lambda_n \in \left]0, \frac{1}{\alpha}\right[$ *and* $\sum_{n\in\mathbb{N}} \lambda_n(1-\alpha\lambda_n) = +\infty$,

then $\mathcal{T}_0$ *is strictly quasinonexpansive (see [Bauschke and Combettes, 2017, Definition 4.1(vi)]), and* $(\boldsymbol{x}_n)_{n\in\mathbb{N}}$ *converges to a point* $\overline{\boldsymbol{x}}_m$ *which is a solution to problem 1.1. Furthermore,* $(\forall i \in \{1,\dots,m-1\})$ $(\boldsymbol{x}_{i,n})_{n\in\mathbb{N}}$ *also converges to* $\overline{\boldsymbol{x}}_m$.

Proof. If $\gamma < \overline{\gamma}$ and $\beta_- > 0$, then $\|\boldsymbol{W}_0\| < 1$. Since $\mathcal{R}_0$ is nonexpansive, it follows that $\mathcal{T}_0$ is strictly nonexpansive, hence it is also strictly quasinonexpansive. Then, we deduce from [Bauschke and Combettes, 2017, Corollary 4.50] that $\mathcal{F}$ is the set of fixed point of $\mathcal{T}_0$, which is also the set of solutions to problem 1.1 (see [Combettes and Wajs, 2005, Proposition 3.1]). The result then directly follows from Theorem 1.3.6. □

§ 1.4 2-periodic forward-backward with varying stepsize

In this section, we investigate Algorithm 2 below, which is a special case of Algorithm 1, where $m = 2$. Note that, since $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$ commute, from now on, without loss of generality, we consider $\gamma_1 \le \gamma_2$. We are going to show that, by allowing $\gamma_1 \neq \gamma_2$, one can obtain looser sufficient conditions compared to Section 1.3 for satisfying Condition 1.1.2.

Algorithm 2: 2-periodic forward-backward algorithm

Initialization: Let $\boldsymbol{x}_1 \in \mathbb{R}^p$.

for $n = 1, 2, \ldots$ **do**

$\quad \boldsymbol{x}_{1,n} = \operatorname{prox}_{\gamma_1 f}\left(\boldsymbol{x}_n - \gamma_1 \boldsymbol{H}^\top \left(\boldsymbol{H}\boldsymbol{x}_n - \boldsymbol{y}\right)\right);$

$\quad \boldsymbol{x}_{2,n} = \operatorname{prox}_{\gamma_2 f}\left(\boldsymbol{x}_{1,n} - \gamma_2 \boldsymbol{H}^\top \left(\boldsymbol{H}\boldsymbol{x}_{1,n} - \boldsymbol{y}\right)\right);$

$\quad \boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \lambda_n \left(\boldsymbol{x}_{2,n} - \boldsymbol{x}_n\right);$

end

1.4.1 Notation

In this section, we introduce notation specific to the study of Algorithm 1.4. For every $i \in \{1, \ldots, 5\}$ let γ_i be defined as follows,

$$\gamma^{(1)} = \frac{\sqrt{2}(\sqrt{2}-1)}{\beta_+ - (\sqrt{2}-1)\beta_-}, \qquad \gamma^{(2)} = \frac{1}{2\beta_-},$$

$$\gamma^{(3)} = \frac{\sqrt{2}}{\beta_+ + (\sqrt{2}-1)\beta_-}, \qquad \gamma^{(4)} = \frac{\beta_- + \beta_+}{\beta_-^2 + \beta_+^2}, \quad \text{and} \quad \gamma^{(5)} = \frac{2}{\beta_+ + \beta_-}.$$

We also introduce the functions ψ, φ and ζ as follows,

$$(\forall \gamma \in \mathbb{R} \setminus \{\gamma^{(4)}\}) \quad \psi(\gamma) = \frac{\gamma^{(5)} - \gamma}{\gamma^{(4)} - \gamma}\gamma^{(4)},$$

$$(\forall \gamma \in \mathbb{R} \setminus \{\gamma^{(2)}\}) \quad \varphi(\gamma) = \frac{\gamma\gamma^{(2)}}{\gamma - \gamma^{(2)}},$$

and

$$(\forall \gamma \in \mathbb{R} \setminus \{\overline{\gamma}/4\}) \quad \zeta(\gamma) = \frac{\gamma\overline{\gamma}}{4\gamma - \overline{\gamma}}.$$

Finally, $\{\xi^{(i)}\}_{1 \le i \le 6}$ will refer to the following set of functions.

$$\begin{aligned}
(\forall \gamma \in \mathbb{R} \setminus \{2\gamma^{(2)}\}) \qquad & \xi^{(1)}(\gamma) = \overline{\gamma}\left(1 + \frac{\gamma}{2(2\gamma^{(2)} - \gamma)}\right) \\
(\forall \gamma \in \mathbb{R} \setminus \{\overline{\gamma}/2\}) \qquad & \xi^{(2)}(\gamma) = \left(1 + \frac{\overline{\gamma} - \gamma}{2\gamma - \overline{\gamma}}\right)\overline{\gamma} \\
(\forall \gamma \in \mathbb{R} \setminus \{2\gamma^{(2)}\}) \qquad & \xi^{(3)}(\gamma) = \frac{2\gamma^{(2)}\overline{\gamma}}{(1 - \kappa^{-1})(2\gamma^{(2)} - \gamma)} \\
(\forall \gamma \in \mathbb{R} \setminus \{\gamma^{(5)}\}) \qquad & \xi^{(4)}(\gamma) = \left(1 + \frac{\gamma^{(5)}}{\gamma^{(5)} - \gamma}\right)\frac{\overline{\gamma}}{2} \\
(\forall \gamma \in \mathbb{R}^*) \qquad & \xi^{(5)}(\gamma) = \left(1 + \frac{\overline{\gamma}}{\gamma(1 - \kappa^{-1})}\right)\frac{\overline{\gamma}}{2} \\
(\forall \gamma \in \mathbb{R} \setminus \{2\gamma^{(2)}\}) \qquad & \xi^{(6)}(\gamma) = \left(1 + \frac{2\gamma^{(2)}}{2\gamma^{(2)} - \gamma}\right)\gamma^{(5)}
\end{aligned}$$

1.4.2 Main results

We present in this section our main results regarding conditions that ensure the convergence of Algorithm 2. First, we give Proposition 1.4.1 below, which links Condition 1.1.2 to a simple condition on the norms of operators $\boldsymbol{W}$, $\boldsymbol{W}_1$ and $\boldsymbol{W}_2$.

Proposition 1.4.1 *Consider Algorithm 2 and notation from Section 1.1. The following statements hold.*

(i) If Condition 1.1.2 is satisfied, then $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| \leq 2$.

(ii) If $\|\boldsymbol{W}\|+\|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$*, then* $\beta_- > 0$ *and Condition 1.1.2 is satisfied.*

Proof. From Proposition 1.2.1 we have $\theta_0 = 1$, $\theta_1 = \|\boldsymbol{W}_1\|$ and $\theta_2 = \|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\|$. From the triangular inequality, for every $\alpha \in \left[\frac{1}{2}, 1\right[$, we have

$$\|\boldsymbol{W} - 2^2(1-\alpha)\boldsymbol{I}_p\| - \|\boldsymbol{W}\| + 2\theta_2 \geq 4\alpha + 2(\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| - 2).$$

Hence, if Condition 1.1.2 holds, then $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| \leq 2$, which leads to (i). On the other side, from the triangular inequality we also have

$$\|\boldsymbol{W} - 2^2(1-\alpha)\boldsymbol{I}_p\| - \|\boldsymbol{W}\| + 2\theta_2 \leq 2^2(1-\alpha) + 2(\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\|). \quad (1.12)$$

If $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$, then there exists $\alpha \in \left[\frac{1}{2}, 1\right[$ such that $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| \leq 2(2\alpha - 1)$. Plugging this into (1.12) leads to

$$\|\boldsymbol{W} - 2^2(1-\alpha)\boldsymbol{I}_p\| - \|\boldsymbol{W}\| + 2\theta_2 \leq 2^2(1-\alpha) + 2^2(2\alpha - 1) = 4\alpha. \quad (1.13)$$

Thus, Condition 1.1.2 holds. If $\beta_- = 0$, then $\|\boldsymbol{W}_1\| \geq 1$, $\|\boldsymbol{W}_2\| \geq 1$ and $\|\boldsymbol{W}\| \geq 1$. Hence, by contraposition, if $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$, then $\beta_- > 0$, which completes the proof. □

Remark 1.4.2 It is worth noting that, by definition of $\boldsymbol{W}$, a sufficient condition for $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$ is $\|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 1$.

Next, we provide a result regarding the equivalence between the sufficient condition in Proposition 1.4.1 (ii) and some constraints on the steps γ_1 and γ_2. We first address the case $\beta_- < \beta_+$, the situtation when $\beta_- = \beta_+$ being dealt with in a separate proposition.

Proposition 1.4.3 *Assume that* $0 < \beta_- < \beta_+$ *and use notation from Sections 1.1 and 1.4.1. Under the assumption that* $0 \leq \gamma_1 \leq \gamma_2$*, the following statements hold.*

(i) If $0 \leq \gamma_1 \leq \gamma_2 \leq \overline{\gamma}$, $(\gamma_1, \gamma_2) \notin \{(0,0), (0,\overline{\gamma}), (\overline{\gamma}, \overline{\gamma})\}$*, then* $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$.

(ii) *If* $0 \leq \gamma_1 \leq \gamma^{(5)}$ *and* $\overline{\gamma} < \gamma_2$*, then* $\|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 1$ *if and only if* $\gamma_1 \neq 0$ *and* $\gamma_2 < \xi^{(1)}(\gamma_1)$*.*

(iii) *If* $\gamma^{(5)} < \gamma_1$ *and* $\overline{\gamma} < \gamma_2$*, then* $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$ *if and only if* $\gamma_1 < \overline{\gamma}$ *and* $\gamma_2 < \xi^{(2)}(\gamma_1)$*.*

(iv) *If* $0 \leq \gamma_1 \leq \gamma^{(5)}$ *and* $\overline{\gamma} < \gamma_2$*, then* $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$ *and* $\beta_+ \leq \beta^*$ *hold if and only if one of the conditions below is satisfied,*

(a) $0 < \gamma_1 \leq \frac{\overline{\gamma}}{4}$ *and,* $\gamma_2 < \xi^{(4)}(\gamma_1)$ *if* $\psi(\gamma_1) < \xi^{(3)}(\gamma_1)$*, or* $\gamma_2 < \xi^{(3)}(\gamma_1)$ *else.*

(b) $\kappa \leq \sqrt{2}(\sqrt{2}+1)$*,* $\frac{\overline{\gamma}}{4} < \gamma_1 < \frac{\overline{\gamma}}{3}$ *and,* $\gamma_2 < \xi^{(4)}(\gamma_1)$ *and* $\gamma_2 \leq \zeta(\gamma_1)$ *if* $\psi(\gamma_1) < \xi^{(3)}(\gamma_1)$*, or* $\gamma_2 < \xi^{(3)}(\gamma_1)$ *else.*

(c) $\kappa > \sqrt{2}(\sqrt{2}+1)$*,* $\frac{\overline{\gamma}}{4} < \gamma_1 \leq \gamma^{(1)}$ *and,* $\gamma_2 < \xi^{(4)}(\gamma_1)$ *and* $\gamma_2 \leq \zeta(\gamma_1)$ *if* $\psi(\gamma_1) < \xi^{(3)}(\gamma_1)$*, or* $\gamma_2 < \xi^{(3)}(\gamma_1)$ *else.*

(d) $\kappa > \sqrt{2}(\sqrt{2}+1)$*,* $\gamma^{(1)} < \gamma_1 < \frac{\overline{\gamma}}{3}$*,* $\gamma_2 \leq \zeta(\gamma_1)$ *and* $\gamma_2 < \xi^{(3)}(\gamma_1)$*.*

(v) *If* $0 \leq \gamma_1 \leq \gamma^{(5)}$ *and* $\overline{\gamma} < \gamma_2$*, then* $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$ *and* $\beta_- \geq \beta^*$ *hold if and only if* $\kappa < 2$*,* $\varphi(\gamma_1) \leq \gamma_2$ *and one of the conditions below is satisfied,*

(a) $\gamma^{(2)} < \gamma_1 \leq \gamma^{(3)}$ *and* $\gamma_2 < \xi^{(6)}(\gamma_1)$*.*

(b) $\gamma^{(3)} < \gamma_1 < \gamma^{(4)}$ *and* $\gamma_2 < \xi^{(6)}(\gamma_1)$ *if* $\xi^{(5)}(\gamma_1) > \psi(\gamma_1)$*, or* $\gamma_2 < \xi^{(5)}(\gamma_1)$ *else.*

(c) $\gamma^{(4)} \leq \gamma_1 \leq \gamma^{(5)}$ *and* $\gamma_2 < \xi^{(5)}(\gamma_1)$*.*

The proof of Proposition 1.4.3 is given in Section 1.4.3. It is worth noting that, in view of Proposition 1.4.10, the above result provides upperbounds that are greater than $\overline{\gamma}$, as opposed to Proposition 1.3.4. However, it should be noted that, at least one step must be lower than $\overline{\gamma}$. We illustrate Proposition 1.4.3 in Figure 1.1 for different values of the condition number κ.

Finally, we address the case $\beta_+ = \beta_-$ in the following proposition, whose proof is given in Section 1.4.3.

Proposition 1.4.4 *Assume that* $\kappa = 1$ *and use notation from Sections 1.1 and 1.4.1. Under the assumption that* $0 \leq \gamma_1 \leq \gamma_2$*,* $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$ *is satisfied if and only if one of the conditions below holds.*

(i) $0 \leq \gamma_1 \leq \gamma_2 \leq \overline{\gamma}$ *and* $(\gamma_1, \gamma_2) \notin \{(0,0), (0,\overline{\gamma}), (\overline{\gamma},\overline{\gamma})\}$

(ii) $0 < \gamma_1 < \gamma^{(5)}$ *and* $\overline{\gamma} < \gamma_2 < \xi^{(1)}(\gamma_1)$

(iii) $\gamma_1 = \gamma^{(5)}$

(iv) $\gamma^{(5)} < \gamma_1 < \overline{\gamma} < \gamma_2 < \xi^{(2)}(\gamma_1)$

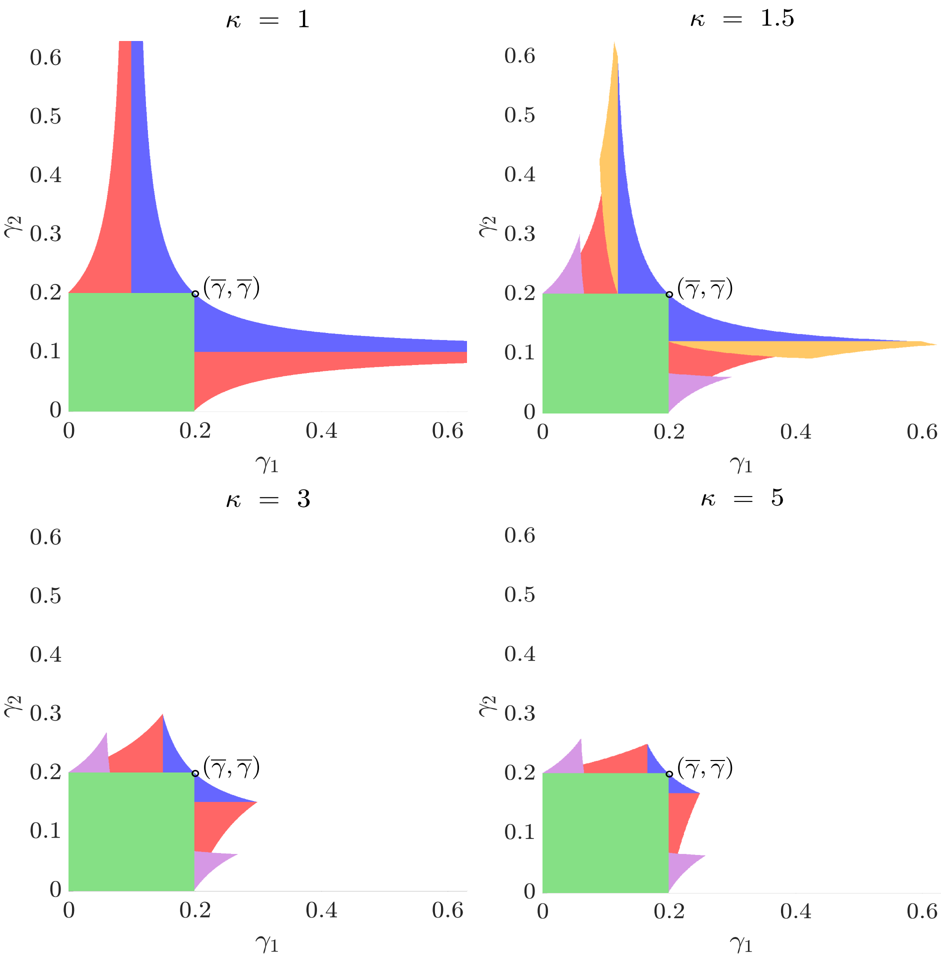

Figure 1.1: *Illustration of Propositions 1.4.3 and 1.4.4 and of the influence of the condition number of* $\boldsymbol{H}^\top\boldsymbol{H}$. *As mentioned earlier,* γ_1 *and* γ_2 *are interchangable, so we do not assume that* $\gamma_1 \leq \gamma_2$ *to plot these figures. In this picture,* $\beta_+ = 10$. *When* $\kappa > 1$, *colors have the following meaning: green, red, blue, purple and orange represent the conditions given by Proposition 1.4.3* (i)-(v), *respectively. For* $\kappa = 1$, *green, red and blue represent the conditions given by Proposition 1.4.4* (i), (ii) *and* (iv), *respectively.*

1.4.3 Proofs

Preliminary results

First, we derive the result below about the ordering of the reference steps.

Proposition 1.4.5 *Using notation from Sections 1.1 and 1.4.1, if* $0 < \beta_- < \beta_+$, *then*

(i) $\frac{\overline{\gamma}}{4} < \gamma^{(1)} < \frac{\overline{\gamma}}{2} < \gamma^{(4)} < \gamma^{(5)}$.

(ii) $\gamma^{(1)} < \frac{\overline{\gamma}}{3}$ *if and only if* $\kappa > \sqrt{2}(\sqrt{2}+1)$.

(iii) $\gamma^{(2)} < \gamma^{(4)}$ *and* $\gamma^{(2)} < \gamma^{(3)} < \gamma^{(4)}$ *are both equivalent to* $\kappa < \sqrt{2}+1$.

(iv) $\gamma^{(2)} < \gamma^{(5)}$ *if and only if* $\kappa < 3$.

Proof. We assume that $0 < \beta_- < \beta_+$. Hence, we have $\beta_+ - (\sqrt{2}-1)\beta_- > 0$ and

$$\begin{aligned}\frac{\overline{\gamma}}{4} < \gamma^{(1)} \quad &\Longleftrightarrow \quad \beta_+ - (\sqrt{2}-1)\beta_- < (4-2\sqrt{2})\beta_+ \\ &\Longleftrightarrow \quad (1-\sqrt{2})\beta_- < (3-2\sqrt{2})\beta_+,\end{aligned}$$

which is satisfied since $1-\sqrt{2} < 0$ and $3-2\sqrt{2} > 0$. Similarly,

$$\gamma^{(1)} < \frac{\overline{\gamma}}{2} \iff \sqrt{2}(\sqrt{2}-1)\beta_+ < \beta_+ - (\sqrt{2}-1)\beta_- \iff \beta_- < \beta_+$$

and

$$\frac{\overline{\gamma}}{2} < \gamma^{(4)} \iff \beta_+^2 + \beta_-^2 < \beta_+^2 + \beta_+\beta_- \iff \beta_- < \beta_+.$$

Finally,

$$\gamma^{(4)} < \gamma^{(5)} \iff (\beta_+ + \beta_-)^2 < 2(\beta_+^2 + \beta_-^2) \iff 0 < (\beta_+ - \beta_-)^2,$$

which completes the proof of (i). Then, we have

$$\begin{aligned}\gamma^{(1)} < \frac{\overline{\gamma}}{3} \quad &\Longleftrightarrow \quad 3\sqrt{2}(\sqrt{2}-1)\beta_+ < 2(\beta_+ - (\sqrt{2}-1)\beta_-) \\ &\Longleftrightarrow \quad (2\sqrt{2}-3)\beta_+ < -\sqrt{2}(\sqrt{2}-1)\beta_- \\ &\Longleftrightarrow \quad \sqrt{2}(\sqrt{2}+1) < \kappa,\end{aligned}$$

where we used $(\sqrt{2}-1)(\sqrt{2}+1) = 1$. This proves (ii). Furthermore,

$$\gamma^{(2)} < \gamma^{(4)} \iff \beta_+^2 + \beta_-^2 < 2\beta_+\beta_- + 2\beta_-^2 \iff 0 < (\sqrt{2}+1-\kappa)(\sqrt{2}-1+\kappa),$$

which is equivalent to $\kappa < \sqrt{2}+1$, since $\kappa > 1$. In addition,

$$\gamma^{(2)} < \gamma^{(3)} \iff \beta_+ + (\sqrt{2}-1)\beta_- < 2\sqrt{2}\beta_- \iff \kappa < \sqrt{2}+1,$$

and

$$\begin{aligned}\gamma^{(3)} < \gamma^{(4)} \quad &\Longleftrightarrow \quad \sqrt{2}(\beta_-^2 + \beta_+^2) < (\beta_+ + \beta_-)(\beta_+ + (\sqrt{2}-1)\beta_-) \\ &\Longleftrightarrow \quad (\sqrt{2}-1)\beta_+^2 + \beta_-^2 - \sqrt{2}\beta_+\beta_- < 0 \\ &\Longleftrightarrow \quad (\sqrt{2}-1)(\kappa-1)(\kappa-(\sqrt{2}+1)) < 0, \end{aligned} \tag{1.14}$$

which is also equivalent to $\kappa < \sqrt{2}+1$. This shows (iii). Eventually,

$$\gamma^{(2)} < \gamma^{(5)} \iff \beta_+ + \beta_- < 4\beta_- \iff \kappa < 3,$$

which completes the proof of (iv). □

Second, we study the variations of functions ψ, φ, ζ and $\{\xi^{(i)}\}_{1\leq i\leq 6}$. These technical results are not interesting per se but they play a role in the derivation of the conditions on the steps γ_1 and γ_2.

Proposition 1.4.6 *Assume that $0<\beta_-<\beta_+$. Using notation from Section 1.1, the following properties are satisfied.*

(i) ψ is strictly increasing on $]0,\gamma^{(4)}[$ and for every $\gamma\in]\gamma^{(4)},+\infty[$ we have $\psi(\gamma)<\gamma^{(4)}$.

(ii) ζ is strictly decreasing on $]\frac{\overline{\gamma}}{4},+\infty[$, $\zeta\left(\frac{\overline{\gamma}}{3}\right)=\overline{\gamma}$ and $\zeta(\gamma^{(1)})=\psi(\gamma^{(1)})$.

(iii) If $\kappa<\sqrt{2}+1$, then, for every $\gamma\in]\gamma^{(2)},\gamma^{(4)}[$, we have $\psi(\gamma)\leq\varphi(\gamma)$ if and only if $\gamma\leq\gamma^{(3)}$.

Proof. Assume that $0<\beta_-<\beta_+$. Then, $(\forall\gamma\neq\gamma^{(4)})\ \psi\prime(\gamma)=\frac{\gamma^{(5)}/\gamma^{(4)}-1}{(1-\gamma/\gamma^{(4)})^2}$. Since we have $\gamma^{(5)}>\gamma^{(4)}$ from Proposition 1.4.5 (i), it follows that $\psi\prime$ is strictly positive on $[0,\gamma^{(4)}[$ and on $]\gamma^{(4)},+\infty[$. Thus, ψ is strictly increasing on these intervals. The fact that $\lim_{\gamma\to+\infty}\psi(\gamma)=\gamma^{(4)}$ completes the proof for (i). For every $\gamma\in]\overline{\gamma}/4,+\infty[$, $\zeta\prime(\gamma)=-(4\gamma/\overline{\gamma}-1)^{-2}<0$. Hence, ζ is strictly decreasing on this interval. Furthermore, $\zeta\left(\frac{\overline{\gamma}}{3}\right)=\frac{\overline{\gamma}}{3(4/3-1)}=\overline{\gamma}$. From Proposition 1.4.5 (i) we have $\frac{\overline{\gamma}}{4}<\gamma^{(1)}$ and

$$\zeta(\gamma^{(1)})=\frac{\sqrt{2}(\sqrt{2}-1)}{(3-2\sqrt{2})\beta_++(\sqrt{2}-1)\beta_-}=\frac{\sqrt{2}}{(\sqrt{2}-1)\beta_++\beta_-}.$$

On the other hand,

$$\begin{aligned}\psi(\gamma^{(1)}) &= \frac{2(\beta_+-(\sqrt{2}-1)\beta_-)-(2-\sqrt{2})(\beta_++\beta_-)}{(\beta_++\beta_-)(\beta_+-(\sqrt{2}-1)\beta_-)-(2-\sqrt{2})(\beta_+^2+\beta_-^2)}\\ &= \frac{\sqrt{2}(\beta_+-\beta_-)}{(\sqrt{2}-1)\beta_+^2-\beta_-^2+\sqrt{2}(\sqrt{2}-1)\beta_+\beta_-}\\ &= \frac{\sqrt{2}(\beta_+-\beta_-)}{(\beta_+-\beta_-)((\sqrt{2}-1)\beta_++\beta_-)}=\zeta(\gamma_1).\end{aligned}$$

This completes the proof of (ii). Assume that $\kappa<\sqrt{2}+1$, then $\gamma^{(2)}<\gamma^{(3)}<\gamma^{(4)}$ (Proposition 1.4.5 (iii)). On the one hand,

$$\varphi(\gamma^{(3)})=\frac{\sqrt{2}}{2\sqrt{2}\beta_--\beta_+-(\sqrt{2}-1)\beta_-}=\frac{\sqrt{2}}{(\sqrt{2}+1)\beta_--\beta_+}.$$

On the other hand,

$$\begin{aligned}\psi(\gamma^{(3)}) &= \frac{2(\beta_+ + (\sqrt{2}-1)\beta_-) - \sqrt{2}(\beta_+ + \beta_-)}{(\beta_+ + \beta_-)(\beta_+ + (\sqrt{2}-1)\beta_-) - \sqrt{2}(\beta_+^2 + \beta_-^2)} \\ &= \frac{\sqrt{2}(\sqrt{2}-1)(\beta_+ - \beta_-)}{(1-\sqrt{2})\beta_+^2 - \beta_-^2 + \sqrt{2}\beta_+\beta_-} \\ &= \frac{\sqrt{2}(\sqrt{2}-1)(\beta_+ - \beta_-)}{(\sqrt{2}-1)(\beta_+ - \beta_-)((\sqrt{2}+1)\beta_- - \beta_+)} = \varphi(\gamma^{(3)}).\end{aligned}$$

In addition, ψ and φ are strictly increasing and decreasing on $]\gamma^{(2)}, \gamma^{(4)}[$, respectively. This completes the proof of (iii). □

Proposition 1.4.7 *Assume that $0 < \beta_- < \beta_+$ and use notation from Sections 1.1 and 1.4.1. Then, the following statements hold.*

(i) For every $\gamma \in [0, \gamma^{(5)}[$, $\psi(\gamma) < \xi^{(3)}(\gamma)$ if and only if $\psi(\gamma) < \xi^{(4)}(\gamma)$.

(ii) For every $\gamma \in]0, \gamma^{(5)}[$, $\psi(\gamma) < \xi^{(5)}(\gamma)$ if and only if $\psi(\gamma) < \xi^{(6)}(\gamma)$.

(iii) If $2 \le \kappa < \sqrt{2}+1$, then for every $\gamma \in]\gamma^{(2)}, \gamma^{(3)}]$ we have $\xi^{(6)}(\gamma) \le \varphi(\gamma)$.

(iv) If $2 \le \kappa < 3$, then for every $\gamma \in]\gamma^{(2)}, \gamma^{(5)}]$ we have $\xi^{(5)}(\gamma) \le \varphi(\gamma)$.

Proof. It is worth noting that $2\gamma^{(2)} = \frac{1}{\beta_-} > \frac{2}{\beta_+ + \beta_-} = \gamma^{(5)}$. Let $\gamma \in [0, \gamma^{(5)}[$,

$$\begin{aligned}\psi(\gamma) < \xi^{(3)}(\gamma) &\iff \frac{2 - (\beta_+ + \beta_-)\gamma}{\beta_+ + \beta_- - (\beta_+^2 + \beta_-^2)\gamma} < \frac{2}{(\beta_+ - \beta_-)(1 - \beta_-\gamma)} \\ &\iff \gamma^2\beta_-(\beta_+^2 - \beta_-^2) + \gamma(\beta_+^2 + 5\beta_-^2 - 2\beta_-\beta_+) - 4\beta_- < 0.\end{aligned}$$

Similarly, we have

$$\begin{aligned}\psi(\gamma) < \xi^{(4)}(\gamma) &\iff \frac{2 - (\beta_+ + \beta_-)\gamma}{\beta_+ + \beta_- - (\beta_+^2 + \beta_-^2)\gamma} < \left(1 + \frac{2}{2 - (\beta_+ + \beta_-)\gamma}\right)\frac{1}{\beta_+} \\ &\iff \gamma^2\beta_-(\beta_+^2 - \beta_-^2) + \gamma(\beta_+^2 + 5\beta_-^2 - 2\beta_+\beta_-) - 4\beta_- < 0.\end{aligned}$$

The above equivalences lead to (i). We proceed the same way to prove (ii): $\psi(\gamma) < \xi^{(5)}(\gamma)$ and $\psi(\gamma) < \xi^{(6)}(\gamma)$ are both equivalent to

$$-\gamma^2\beta_-(\beta_+ - \beta_-)^2 + \gamma(3\beta_+^2 + 3\beta_-^2 - 2\beta_+\beta_-) - 2(\beta_+ + \beta_-) < 0,$$

which completes the proof for (i) and (ii).
Now, assume that $2 \le \kappa < \sqrt{2}+1$. Let $\gamma \in]\gamma^{(2)}, \gamma^{(3)}]$. Then,

$$\begin{aligned}\xi^{(6)}(\gamma) \le \varphi(\gamma) &\iff \frac{2}{\beta_+ + \beta_-}\left(1 + \frac{1}{1 - \beta_-\gamma}\right) \le \frac{\gamma}{2\beta_-\gamma - 1} \\ &\iff 2(2 - \beta_-\gamma)(2\beta_-\gamma - 1) \le \gamma(\beta_+ + \beta_-)(1 - \beta_-\gamma) \\ &\iff (\gamma\beta_-)^2(\kappa - 3) + \gamma\beta_-(9 - \kappa) - 4 \le 0.\end{aligned}$$

In addition, the above second degree polynomial is maximal at $(9-\kappa)/(2(3-\kappa))$, and we have $\beta_-\gamma^{(3)} < 1 < (9-\kappa)/(2(3-\kappa))$ and

$$\begin{aligned}
\xi^{(6)}(\gamma^{(3)}) \leq \varphi(\gamma^{(3)}) \quad &\iff \quad \frac{2}{\beta_+ + \beta_-}\left(2 + \frac{\sqrt{2}\beta_-}{\beta_+ - \beta_-}\right) \leq \frac{\sqrt{2}}{(\sqrt{2}+1)\beta_- - \beta_+} \\
&\iff \quad 2\sqrt{2} \leq (2\sqrt{2}+1)(\kappa-1)^2,
\end{aligned}$$

which is satisfied since $2 \leq \kappa$ by assumption. Hence, $\xi^{(6)}(\gamma) \leq \varphi(\gamma)$, which proves (iii).
Now, assume that $2 \leq \kappa < 3$. Let $\gamma \in]\gamma^{(2)}, \gamma^{(5)}]$. Then,

$$\begin{aligned}
\xi^{(5)}(\gamma) \leq \varphi(\gamma) \quad &\iff \quad \frac{1}{\beta_+}\left(1 + \frac{2}{\gamma(\beta_+ - \beta_-)}\right) \leq \frac{\gamma}{2\beta_-\gamma - 1} \\
&\iff \quad (2 + \gamma(\beta_+ - \beta_-))(2\beta_-\gamma - 1) \leq \beta_+\gamma^2(\beta_+ - \beta_-) \\
&\iff \quad (\gamma\beta_-)^2(\kappa-1)(2-\kappa) + \gamma\beta_-(5-\kappa) - 2 \leq 0.
\end{aligned}$$

In addition, the above second degree polynomial is maximal at $(5-\kappa)/(2(\kappa-1)(\kappa-2))$. It can be shown that, under the assumption $2 \leq \kappa < 3$, we have $\beta_-\gamma^{(5)} < (5-\kappa)/(2(\kappa-1)(\kappa-2))$ and

$$\xi^{(5)}(\gamma^{(5)}) \leq \varphi(\gamma^{(5)}) \iff \frac{2}{\beta_+ - \beta_-} \leq \frac{2}{3\beta_- - \beta_+} \iff 2 \leq \kappa,$$

which is satisfied by assumption. Hence, $\xi^{(5)}(\gamma) \leq \varphi(\gamma)$, which completes the proof of (iv). □

Third, we give a result which is related to the computation of $\|\boldsymbol{W}\|$.

Proposition 1.4.8 *Use notation from Sections 1.1 and 1.4.1 and assume that $0 < \beta_- < \beta_+$. Let $0 \leq \gamma_1 \leq \gamma_2$ and let $h : \beta \in \mathbb{R} \mapsto (1-\gamma_1\beta)(1-\gamma_2\beta)$. Then, $h(\beta_+) \leq -h(\beta_-)$ if and only if, $\gamma_1 < \gamma^{(4)}$ and $\gamma_2 \geq \psi(\gamma_1)$.*

Proof. Assume that $0 < \beta_- < \beta_+$ and let $0 \leq \gamma_1 \leq \gamma_2$. Then, the following equivalences hold,

$$\begin{aligned}
h(\beta_+) \leq -h(\beta_-) \quad &\iff \gamma_2\left(\left(\beta_+^2 + \beta_-^2\right)\gamma_1 - (\beta_+ + \beta_-)\right) \leq \gamma_1(\beta_+ + \beta_-) - 2 \\
&\iff \gamma_2(\gamma_1 - \gamma^{(4)}) \leq \gamma^{(4)}(\gamma_1 - \gamma^{(5)}). \qquad (1.15)
\end{aligned}$$

In addition, if $\gamma_1 > \gamma^{(5)}$ then, from Proposition 1.4.5 (i), we deduce that $\gamma_1 > \gamma^{(4)}$ and, under this condition, the right-hand side in (1.15) is equivalent to $\gamma_2 \leq \psi(\gamma_1)$. However, from Proposition 1.4.6 (i) we have $\psi(\gamma_1) < \gamma^{(4)} < \gamma_1 \leq \gamma_2$. Hence, if $\gamma_1 > \gamma^{(5)}$ then $h(\beta_+) > -h(\beta_-)$. If $\gamma_1 = \gamma^{(5)}$, then $\gamma_2(\gamma_1 - \gamma^{(4)}) > 0$ and (1.15) is not satisfied. Finally, if $\gamma_1 < \gamma^{(5)}$, then $\gamma_1 - \gamma^{(5)} < 0$, and, since $\gamma_2 \geq 0$, (1.15) is equivalent to $\gamma_1 < \gamma^{(4)}$ and

$\gamma_2 \geq \psi(\gamma_1)$, which completes the proof. $\square$

It is worth noting that the norm of $\boldsymbol{W}$ might depend on eigen values of $\boldsymbol{H}^\top \boldsymbol{H}$ other than β_+ and β_-. In order to provide simple characterizations that only depends on the extremal eigen values, we give two configurations where $\|\boldsymbol{W}\|$ can be expressed using only β_+ and β_-.

Proposition 1.4.9 *Assume that $0 < \beta_- < \beta_+$, $0 \leq \gamma_1 \leq \gamma_2$, and use notation from Sections 1.1 and 1.4.1. Let $\beta^* = \frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2}$ and $h : \beta \mapsto (1-\gamma_1\beta)(1-\gamma_2\beta)$. Under these assumptions,*

(i) if $\beta_+ \leq \beta^*$, then $\|\boldsymbol{W}\| = \max\{h(\beta_-), -h(\beta_+)\}$. In addition, $\beta_+ \leq \beta^*$ if and only if $\gamma_1 \leq \frac{\overline{\gamma}}{4}$, or, $\frac{\overline{\gamma}}{4} < \gamma_1 \leq \frac{\overline{\gamma}}{2}$ and $\gamma_2 \leq \zeta(\gamma_1)$.

(ii) if $\beta_- \geq \beta^*$ then $\|\boldsymbol{W}\| = \max\{-h(\beta_-), h(\beta_+)\}$. In addition, $\beta_- \geq \beta^*$ if and only if $\gamma_1 > \gamma^{(2)}$ and $\gamma_2 \geq \varphi(\gamma_1)$.

Proof. Assume that $0 < \beta_- < \beta_+$ and $0 \leq \gamma_1 \leq \gamma_2$. Then,

$$\beta_+ \leq \beta^* \iff 2\gamma_1\gamma_2\beta_+ \leq \gamma_1 + \gamma_2 \iff \gamma_2 \left(4\frac{\gamma_1}{\overline{\gamma}} - 1\right) \leq \gamma_1.$$

The above inequality is equivalent to $\gamma_1 \leq \frac{\overline{\gamma}}{4}$, or, $\gamma_1 > \frac{\overline{\gamma}}{4}$ and $\gamma_1 \leq \gamma_2 \leq \zeta(\gamma_1)$. In addition, if $\gamma_1 > \frac{\overline{\gamma}}{4}$, then $\gamma_1 \leq \zeta(\gamma_1)$ if and only if $\gamma_1 \leq \frac{\overline{\gamma}}{2}$. Similarly,

$$\beta_- \geq \beta^* \iff 2\gamma_1\gamma_2\beta_- \geq \gamma_1 + \gamma_2 \iff \gamma_2 \left(\frac{\gamma_1}{\gamma^{(2)}} - 1\right) \geq \gamma_1.$$

Since $\gamma_1 > 0$ and $\gamma^{(2)} > 0$, the above inequality directly is equivalent to $\gamma_1 > \gamma^{(2)}$ and $\gamma_2 \geq \varphi(\gamma_1)$. Finally, the results regarding the expression of $\|\boldsymbol{W}\|$ simply follows from its definition and from the variations of h, which is decreasing on $[0, \beta^*]$ and increasing on $[\beta^*, +\infty[$. This completes the proof. $\square$

Finally, we show that the set of functions $\{\xi^{(i)}\}_{1\leq i\leq 6}$ can provide upper-bounds on the steps that are greater than the standard bound $\overline{\gamma}$.

Proposition 1.4.10 *Assume that $0 < \beta_- < \beta_+$ and use notation from Sections 1.1 and 1.4.1. Let $\gamma \in \mathbb{R}$. If $0 < \gamma < 2\gamma^{(2)}$, then $\xi^{(1)}(\gamma) > \overline{\gamma}$, $\xi^{(3)}(\gamma) > \overline{\gamma}$ and $\xi^{(6)}(\gamma) > \overline{\gamma}$. If $0 < \gamma < \gamma^{(5)}$, then $\xi^{(4)}(\gamma) > \overline{\gamma}$. If $0 < \gamma \leq \gamma^{(5)}$, then $\xi^{(5)}(\gamma) > \overline{\gamma}$. And, if $\overline{\gamma}/2 < \gamma < \overline{\gamma}$, then $\xi^{(2)}(\gamma) > \overline{\gamma}$.*

Proof. Assume that $0 < \beta_- < \beta_+$. It is worth noting that $\kappa > 1$. If $0 < \gamma < 2\gamma^{(2)}$, then $\gamma/(2\gamma^{(2)} - \gamma) > 0$ and $\xi^{(1)}(\gamma) > \overline{\gamma}$. Similarly, if $\overline{\gamma}/2 < \gamma < \overline{\gamma}$, then $\overline{\gamma}(\overline{\gamma} - \gamma)/(2\gamma - \overline{\gamma}) > 0$, and $\xi^{(2)} > \overline{\gamma}$. If $0 < \gamma < 2\gamma^{(2)}$, then $\xi^{(3)}(\gamma) \geq \overline{\gamma}/(1-\kappa^{-1}) > \overline{\gamma}$ by assumption on κ. If $0 < \gamma < \gamma^{(5)}$, then

$\gamma^{(5)}/(\gamma^{(5)}-\gamma) > 1$, and $\xi^{(4)}(\gamma) > \overline{\gamma}$. If $0 < \gamma \leq \gamma^{(5)}$, then, by definition of $\gamma^{(5)}$ and since $\kappa > 1$, we have $\gamma < 2/(\beta_+ - \beta_-) = \overline{\gamma}/(1-\kappa^{-1})$. Hence, $\xi^{(5)}(\gamma) > \overline{\gamma}$. Finally, if $0 < \gamma < 2\gamma^{(2)}$, then $\xi^{(6)}(\gamma) \geq 2\gamma^{(5)} > \overline{\gamma}$. This completes the proof. □

Proof of Proposition 1.4.3

First, we derive some necessary and sufficient conditions on the norm of the linear operators such that Condition 1.1.2 is satisfied.

Proof. Assume that $0 < \beta_- < \beta_+$ and that $0 \leq \gamma_1 \leq \gamma_2$.

First, let $0 \leq \gamma_1 \leq \gamma_2 \leq \overline{\gamma}$. Then, for every $i \in \{1,2\}$, $\gamma_i\beta_+ - 1 \leq 1$ and $1 - \gamma_i\beta_- \leq 1$. If $(\gamma_1, \gamma_2) \notin \{(0,0),(0,\overline{\gamma}),(\overline{\gamma},\overline{\gamma})\}$ then these inequalities are strict for at least one index. Hence, we have $\|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 1$. Combining this with $\|\boldsymbol{W}\| \leq \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\|$ leads to (i).

Now, assume that $0 \leq \gamma_1 \leq \gamma^{(5)}$ and $\overline{\gamma} < \gamma_2$. Then, $\|\boldsymbol{W}_1\| = 1 - \gamma_1\beta_-$ and $\|\boldsymbol{W}_2\| = \gamma_2\beta_+ - 1$. Hence,

$$\|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 1 \iff (1-\gamma_1\beta_-)(\gamma_2\beta_+ - 1) < 1 \iff \gamma_2 < \xi^{(1)}(\gamma_1).$$

Note that, $\xi^{(1)}$ is strictly increasing on $[0, 2\gamma^{(2)}[$ and that $\xi^{(1)}(0) = \overline{\gamma}$. Thus, $\overline{\gamma} < \xi^{(1)}(\gamma_1) \iff \gamma_1 \neq 0$. This completes the proof for (ii).

Assume that $\gamma^{(5)} < \gamma_1$ and that $\overline{\gamma} < \gamma_2$. Then, $\|\boldsymbol{W}_1\| = \gamma_1\beta_+ - 1$ and $\|\boldsymbol{W}_2\| = \gamma_2\beta_+ - 1$. In addition, for every $i \in \{1,2\}$ and every $\beta \in [\beta_-, \beta_+]$, $\gamma_i\beta_+ - 1 > 1 - \gamma_i\beta_- \geq 1 - \gamma_i\beta$ and $\gamma_i\beta_+ - 1 \geq \gamma_i\beta - 1$, thus $\gamma_i\beta_+ - 1 \geq |1-\gamma_i\beta|$. Therefore, $\|\boldsymbol{W}\| = \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\|$ and

$$\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff (\gamma_1\beta_+ - 1)(\gamma_2\beta_+ - 1) < 1 \iff \gamma_2 < \xi^{(2)}(\gamma_1).$$

In addition, $\xi^{(2)}$ is strictly decreasing on $]\gamma^{(5)}, +\infty[$ and $\xi^{(2)}(\overline{\gamma}) = \overline{\gamma}$, which completes the proof for (iii).

Let $h : \beta \mapsto (1-\gamma_1\beta)(1-\gamma_2\beta)$. Until the end of the proof we assume that $0 \leq \gamma_1 \leq \gamma^{(5)}$ and $\overline{\gamma} < \gamma_2$. Then, $\|\boldsymbol{W}_1\| = 1 - \gamma_1\beta_-$ and $\|\boldsymbol{W}_2\| = \gamma_2\beta_+ - 1$. If $\gamma_1 = 0$, then $(\forall \beta \in [\beta_-, \beta_+])\ -h(\beta_+) = \gamma_2\beta_+ - 1 \geq |1-\gamma_2\beta| = |h(\beta)|$. Hence, $\|\boldsymbol{W}\| = \gamma_2\beta_+ - 1$ and $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff \gamma_2 < \overline{\gamma}$, which is in contradiction with our assumption on γ_2. From now on we assume that $\gamma_1 > 0$.

It follows from Proposition 1.4.9 (i) that $\beta_+ \leq \beta^*$ if and only if $\gamma_1 \leq \frac{\overline{\gamma}}{4}$, or $\frac{\overline{\gamma}}{4} < \gamma_1 \leq \frac{\overline{\gamma}}{2}$ and $\gamma_2 \leq \zeta(\gamma_1)$. We deal with these two cases separately.

Assume that $0 < \gamma_1 \leq \frac{\overline{\gamma}}{4}$ and $\overline{\gamma} < \gamma_2$. From Proposition 1.4.9 (i) we deduce that $\|\boldsymbol{W}\| = \max\{h(\beta_-), -h(\beta_+)\}$. In addition, from Proposition 1.4.5 (i) we deduce that $\gamma_1 < \gamma^{(4)}$. Hence, from Proposition 1.4.8 there are two case,

(i) if $\gamma_2 < \psi(\gamma_1)$, then $h(\beta_+) > -h(\beta_-)$ and $\|\boldsymbol{W}\| = h(\beta_-)$. In this configuration,

$$\begin{aligned}
&\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \\
&\qquad \iff \quad (1-\gamma_1\beta_-)(\gamma_2\beta_+ - 1) + (1-\gamma_1\beta_-)(1-\gamma_2\beta_-) < 2 \\
&\qquad \iff \quad \gamma_2 < \xi^{(3)}(\gamma_1).
\end{aligned} \tag{1.16}$$

(ii) if $\gamma_2 \geq \psi(\gamma_1)$, then $h(\beta_+) \leq -h(\beta_-)$ and $\|\boldsymbol{W}\| = -h(\beta_+)$. In this configuration,

$$\begin{aligned}
&\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \\
&\qquad \iff \quad (1-\gamma_1\beta_-)(\gamma_2\beta_+ - 1) + (1-\gamma_1\beta_+)(\gamma_2\beta_+ - 1) < 2 \\
&\qquad \iff \quad \gamma_2 < \xi^{(4)}(\gamma_1).
\end{aligned} \tag{1.17}$$

In addition, if $\psi(\gamma_1) < \xi^{(3)}(\gamma_1)$ then $\gamma_2 < \psi(\gamma_1) \implies \gamma_2 < \xi^{(3)}(\gamma_1) \implies \|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$ and it follows from Proposition 1.4.7 (i) that $\psi(\gamma_1) < \xi^{(4)}(\gamma_1)$. If $\psi(\gamma_1) \geq \xi^{(3)}(\gamma_1)$ then it follows from Proposition 1.4.7 (i) that $\psi(\gamma_1) \geq \xi^{(4)}(\gamma_1)$ and $\gamma_2 \geq \psi(\gamma_1) \implies \|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| \geq 2$. This completes the proof for (iva).
Assume that $\frac{\overline{\gamma}}{4} < \gamma_1 \leq \frac{\overline{\gamma}}{2}$ and $\overline{\gamma} < \gamma_2 \leq \zeta(\gamma_1)$. From Proposition 1.4.9 (i) we deduce that $\|\boldsymbol{W}\| = \max\{h(\beta_-), -h(\beta_+)\}$. Note that, from Proposition 1.4.6 (ii), $\overline{\gamma} < \gamma_2 \leq \zeta(\gamma_1)$ holds if and only if $\gamma_1 < \frac{\overline{\gamma}}{3}$. Now, there are two cases according to the value of κ.

(i) If $\kappa \leq \sqrt{2}(\sqrt{2}+1)$, then it follows from Proposition 1.4.5 (ii) that $\frac{\overline{\gamma}}{3} \leq \gamma^{(1)}$. Hence, Proposition 1.4.6 (i) and (ii) leads to $\psi(\gamma_1) < \zeta(\gamma_1)$. Thus there are two cases according to how γ_2 compares with $\psi(\gamma_1)$.

   (a) If $\gamma_2 < \psi(\gamma_1)$, then Proposition 1.4.8 leads to $h(\beta_+) > -h(\beta_-)$ and $\|\boldsymbol{W}\| = h(\beta_-)$. Therefore, in this configuration, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$ if and only if $\gamma_2 < \xi^{(3)}(\gamma_1)$.

   (b) If $\psi(\gamma_1) \leq \gamma_2 \leq \zeta(\gamma_1)$, then Proposition 1.4.8 leads to $h(\beta_+) \leq -h(\beta_-)$ and $\|\boldsymbol{W}\| = -h(\beta_+)$. Therefore, in this configuration, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2$ if and only if $\gamma_2 < \xi^{(4)}(\gamma_1)$.

   We use Proposition 1.4.7 (i) to complete the proof for (ivb) (see the proof for (iva) for more details).

(ii) If $\kappa > \sqrt{2}(\sqrt{2}+1)$, then it follows from Proposition 1.4.5 (iii) that $\frac{\overline{\gamma}}{3} > \gamma^{(1)}$. If, in addition,

   (a) $\frac{\overline{\gamma}}{4} < \gamma_1 \leq \gamma^{(1)}$, then $\psi(\gamma_1) \leq \zeta(\gamma_1)$ and the proof for (ivc) is the same as for (ivb).

   (b) $\gamma^{(1)} < \gamma_1 < \frac{\overline{\gamma}}{3}$, then $\gamma_2 \leq \zeta(\gamma_1) < \psi(\gamma_1)$. Thus, $\|\boldsymbol{W}\| = h(\beta_-)$ and $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff \gamma_2 < \xi^{(3)}(\gamma_1)$, which completes the proof for (ivd).

We still work under the assumption that $0 \leq \gamma_1 \leq \gamma^{(5)}$ and $\overline{\gamma} < \gamma_2$ and we are now going to show (va)-(vc). From Proposition 1.4.9 (ii) we have $\beta_- \geq \beta^*$ if and only if $\gamma_1 > \gamma^{(2)}$ and $\gamma_2 \geq \varphi(\gamma_1)$. Hence, if $\beta_- \geq \beta^*$, then we have necessarily $\gamma^{(2)} < \gamma^{(5)}$ and, from Proposition 1.4.5 (iv), we must have $\kappa < 3$. First, we assume that $\kappa < \sqrt{2}+1$, $\gamma_1 > \gamma^{(2)}$ and $\gamma_2 \geq \varphi(\gamma_1)$. Then, $\beta_- \geq \beta^*$ and Proposition 1.4.9 (ii) leads to $\|\boldsymbol{W}\| = \max\{-h(\beta_-), h(\beta_+)\}$. In addition, Proposition 1.4.5 (iii) leads to $\gamma^{(2)} < \gamma^{(3)} < \gamma^{(4)}$. There are three cases depending on the value of γ_1.

(i) If $\gamma^{(2)} < \gamma_1 \leq \gamma^{(3)}$, then it follows from Proposition 1.4.6 (iii) that $\psi(\gamma_1) \leq \varphi(\gamma_1) \leq \gamma_2$. Therefore, Proposition 1.4.8 leads to $h(\beta_+) \leq -h(\beta_-)$ and $\|\boldsymbol{W}\| = -h(\beta_-)$. Under these conditions,

$$\begin{aligned}
&\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \\
&\qquad\iff (1-\gamma_1\beta_-)(\gamma_2\beta_+ - 1) + (1-\gamma_1\beta_-)(\gamma_2\beta_- - 1) < 2 \\
&\qquad\iff (1-\gamma_1\beta_-)(\gamma_2(\beta_+ + \beta_-) - 2) < 2 \\
&\qquad\iff \gamma_2 < \xi^{(6)}(\gamma_1).
\end{aligned}$$

However, from Proposition 1.4.10 (iii), if $\xi^{(6)}(\gamma_1) > \varphi(\gamma_1)$ then we must have $\kappa < 2$. This completes the proof for (va).

(ii) If $\gamma^{(3)} < \gamma_1 < \gamma^{(4)}$, then Proposition 1.4.6 (iii) leads to $\psi(\gamma_1) > \varphi(\gamma_1)$. If, in addition,

(a) $\varphi(\gamma_1) \leq \gamma_2 < \psi(\gamma_1)$, then Proposition 1.4.8 leads to $h(\beta_+) > -h(\beta_-)$ and $\|\boldsymbol{W}\| = h(\beta_+)$. Therefore,

$$\begin{aligned}
&\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \\
&\qquad\iff (1-\gamma_-\beta_-)(\gamma_2\beta_+ - 1) + (1-\gamma_1\beta_+)(1-\gamma_2\beta_+) < 2 \\
&\qquad\iff (\gamma_2\beta_+ - 1)\gamma_1(\beta_+ - \beta_-) < 2 \\
&\qquad\iff \gamma^{(2)} < \xi^{(5)}(\gamma_1).
\end{aligned}$$

(b) $\psi(\gamma_1) \leq \gamma_2$, then Proposition 1.4.8 leads to $h(\beta_+) \leq -h(\beta_-)$ and $\|\boldsymbol{W}\| = -h(\beta_-)$ and $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff \gamma_2 < \xi^{(6)}(\gamma_1)$.

If $\kappa \geq 2$, then Propositions 1.4.10 (iv) and 1.4.7 (ii) lead to $\xi^{(5)}(\gamma_1) \leq \varphi(\gamma_1) < \psi(\gamma_1)$ and $\xi^{(6)}(\gamma_1) \leq \psi(\gamma_1)$. Thus, if $\kappa \geq 2$ then $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| \geq 2$. Proposition 1.4.7 (ii) completes the proof for (vb).

(iii) If $\gamma^{(4)} \leq \gamma_1 \leq \gamma^{(5)}$, then Proposition 1.4.8 leads to $h(\beta_+) > -h(\beta_-)$ and $\|\boldsymbol{W}\| = h(\beta_+)$. Therefore, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff \gamma_2 < \xi^{(5)}(\gamma_1)$. If $\kappa \geq 2$, then Proposition 1.4.10 (iv) leads to $\xi^{(5)}(\gamma_1) \leq \varphi(\gamma_1)$, so $\gamma_2 \geq \xi^{(5)}(\gamma_1)$ since we assumed that $\gamma_2 \geq \varphi(\gamma_1)$. This completes the proof for (vc).

Second, we assume that $\sqrt{2}+1 \le \kappa < 3$, $\gamma_1 > \gamma^{(2)}$ and $\gamma_2 \ge \varphi(\gamma_1)$. From Proposition 1.4.5 (iii) we have $\gamma^{(2)} \ge \gamma^{(4)}$. Hence, $\gamma_1 > \gamma^{(4)}$ and Proposition 1.4.8 leads to $h(\beta_+) > -h(\beta_-)$ and $\|\boldsymbol{W}\| = h(\beta_+)$. Under these conditions, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff \gamma_2 < \xi^{(5)}(\gamma_1)$. However, from Proposition 1.4.7 (iv) we have $\xi^{(5)}(\gamma_1) \le \varphi(\gamma_1)$. Therefore, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| \ge 2$, which completes the proof. □

Proof of Proposition 1.4.4

Proof. Assume that $\kappa = 1$. Let β denote $\beta_- = \beta_+$. Then, $\|\boldsymbol{W}\| = \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| = |1-\gamma_1\beta||1-\gamma_2\beta|$. In order to show the equivalence, we are going to deal with all possible cases regarding the valeus of γ_1 and γ_2.

If $0 \le \gamma_1 \le \gamma_2 \le \overline{\gamma}$ then $(\forall i \in \{1,2\})$ $|1-\gamma_i\beta| \le 1$ and, if $(\gamma_1,\gamma_2) \notin \{(0,0),(0,\overline{\gamma}),(\overline{\gamma},\overline{\gamma})\}$, then this inequality is strict for at least one index. Hence, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| = 2|1-\gamma_1\beta||1-\gamma_2\beta| < 2$.

If $\gamma_1 = 0$ and $\overline{\gamma} < \gamma_2$, then $\|\boldsymbol{W}\|_1 = 1$ and $\|\boldsymbol{W}_2\| = \gamma_2\beta - 1$. Hence, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff 2(\gamma_2\beta - 1) < 2 \iff \gamma_2 < \overline{\gamma}$, which is a contradiction.

If $0 < \gamma_1 < \gamma^{(5)}$ and $\overline{\gamma} < \gamma_2$, then $\|\boldsymbol{W}_1\| = 1-\gamma_1\beta$ and $\|\boldsymbol{W}_2\| = \gamma_2\beta-1$. Thus, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff 2(1-\gamma_1\beta)(\gamma_2\beta-1) < 2 \iff \gamma_2 < \xi^{(1)}(\gamma_1)$.

If $\gamma_1 = \gamma^{(5)} = 1/\beta$, then $\|\boldsymbol{W}_1\| = 0$ and $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| = 0 < 2$ is satisfied for every γ_2.

If $\gamma^{(5)} < \gamma_1$ and $\overline{\gamma} < \gamma_2$, then $\|\boldsymbol{W}_1\| = \gamma_1\beta-1$ and $\|\boldsymbol{W}_2\| = \gamma_2\beta-1$. Therefore, $\|\boldsymbol{W}\| + \|\boldsymbol{W}_1\|\|\boldsymbol{W}_2\| < 2 \iff 2(\gamma_1\beta-1)(\gamma_2\beta-1) < 2 \iff \gamma_2 < \xi^{(2)}(\gamma_1)$. Since we are working under the assumption that $\gamma_1 \le \gamma_2$, $\gamma_2 < \xi^{(2)}(\gamma_1)$ implies that $\gamma_1 < \xi^{(2)}(\gamma_1)$, which leads to $\gamma_1 < \overline{\gamma}$. The proof is complete. □

Bibliography

H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces.* Springer, 2017. doi: 10.1007/978-3-319-48311-5. 7

P. L. Combettes and J.-C. Pesquet. Deep neural network structures solving variational inequalities. *arXiv preprint arXiv:1808.07526*, 2018. 2, 3, 6, 7

P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling and Simulation*, 4(4):1168–1200, 2005. 7