

Comparison of Proximal First-Order Primal and Primal-Dual Algorithms via Performance Estimation

Nizar Bousselmi,
ICTEAM, UCLouvain
Louvain-la-Neuve, Belgium
nizar.bousselmi@uclouvain.be

Nelly Pustelnik,
CNRS, ENS de Lyon
Lyon, France
nelly.pustelnik@ens-lyon.fr

Julien M. Hendrickx,
ICTEAM, UCLouvain
Louvain-la-Neuve, Belgium
julien.hendrickx@uclouvain.be

François Glineur,
CORE/ICTEAM, UCLouvain
Louvain-la-Neuve, Belgium
francois.glineur@uclouvain.be

Abstract—Selecting the fastest algorithm for a specific signal/image processing task is a challenging question. We propose an approach based on the Performance Estimation Problem framework that numerically and automatically computes the worst-case performance of a given optimization method on a class of functions. We first propose a computer-assisted analysis and comparison of several first-order primal optimization methods, namely, the gradient method, the forward-backward, Peaceman-Rachford, and Douglas-Rachford splittings. We tighten the existing convergence results of these algorithms and extend them to new classes of functions. Our analysis is then extended and evaluated in the context of the primal-dual Chambolle-Pock and Condat-Vũ methods.

Index Terms—Complexity analysis, Proximal algorithms, Primal methods, Primal-dual methods

I. INTRODUCTION

Many applications in signal processing and machine learning [1]–[4] require to solve composite problems of the form

$$\min_{x \in \mathcal{H}} f(x) + g(x) \quad (1)$$

or

$$\min_{x \in \mathcal{H}} f(x) + h(Mx) \quad (2)$$

for a finite-dimensional Hilbert space \mathcal{H} and some properties of smoothness and convexity on functions f , g , or h and a linear operator M . A large number of gradient and proximal-based algorithms have been developed to solve such minimization problems [1], [3], [5]–[8]. However, selecting the fastest method according to a target class of functions remains a tedious task, especially when no tight theoretical convergence rate is available.

In this work, we propose to apply the Performance Estimation Problem (PEP) framework [9], [10] to (1) and (2). This analysis allows us to obtain the exact worst-case performance for these optimization schemes, and therefore to identify the fastest strategy for a given class of functions.

State-of-the-art for algorithmic comparisons – The optimization literature is well documented in terms of design of efficient algorithms. However, quantifying the efficiency of the methods is not usually done in a unified way, which makes

comparisons difficult. We often have access to the theoretical worst-case performance of a method in a given setting, which can differ from one analysis to another in terms of function classes and performance criteria. For instance, forward-backward splitting has been studied in terms of convergence of both iterates [11] and function values [12], but usually, the analysis of methods exists only for one criterion.

There are also plenty of numerical analyses and comparisons of methods on specific instances of (1) or (2) that illustrate the behavior of the methods but without any theoretical generalization. For instance, one of the most standard image restoration minimization problems involves f as a quadratic (least squares) term and $h = \lambda \|\cdot\|_1$ where λ is a hyperparameter allowing a tradeoff between the data-fidelity term f and the penalization $\|M \cdot\|_1$. Numerical comparisons are generally performed for a specific instance of f and λ , while the performance on other problems can be affected by the Lipschitz constant of the gradient of f or the value of the hyperparameter λ .

The PEP framework has been used to derive and prove tight rates in different configurations [9], [10], [13], [14]. Currently, PEP can analyze many first-order methods on different function classes and has been recently extended [15] to classes involving a linear operator such as in (2). However, the tool has not been often used to compare different methods in order to select the best one.

Preliminary theoretical studies have been proposed in [16], [17] providing theoretical rate comparisons between different first-order algorithms, thus allowing to establish which method is the fastest. However, this type of analysis faces two limitations: (i) there are no guarantees that these bounds are tight and (ii) such a theoretical analysis is tedious to extend to other algorithmic schemes such as primal-dual methods.

Contributions – First, we complement the theoretical analysis in [16] that provided upper bounds on the worst-case guarantees for problem (1) with f smooth and strongly convex and g smooth and convex. For primal algorithms, we show the tightness of these bounds with PEP and improve them when possible. Then, we extend their results with numerically supported conjectures for additional classes of functions, namely, to the doubly strongly convex case and to problem (2). Finally, we use our tool on primal-dual methods.

N. Bousselmi is a FRIA grantee of the Fonds de la Recherche Scientifique - FNRS. N. Pustelnik is supported by the Fondation Simone et Cino Del Duca - Institut de France.

II. PEP ANALYSIS FOR THE COMPOSITE PROBLEM $f + g$

We consider problem (1) when f is α^{-1} -smooth (i.e. differentiable with α^{-1} -Lipschitz gradient) ρ -strongly convex and g is β^{-1} -smooth μ -strongly convex with $\alpha, \beta \in]0, \infty[$ and $\rho, \mu \in [0, \infty[$, and we seek the largest (i.e. worst case) contraction factor r such that, for every $(x, y) \in \mathcal{H} \times \mathcal{H}$,

$$\|\Phi_{f,g}x - \Phi_{f,g}y\| \leq r\|x - y\| \quad (3)$$

where $\Phi_{f,g} : \mathcal{H} \rightarrow \mathcal{H}$ denotes an algorithmic iteration to solve (1). The corresponding conceptual PEP consists in

$$\begin{aligned} \max_{x,y,f,g} \quad & \|\Phi_{f,g}x - \Phi_{f,g}y\|^2 \\ \text{s.t.} \quad & f \in \mathcal{F}_{\rho,\alpha^{-1}}, g \in \mathcal{F}_{\mu,\beta^{-1}} \\ & \|x - y\|^2 \leq 1, \end{aligned} \quad (4)$$

where $\mathcal{F}_{\rho,\alpha^{-1}}$ denotes the class of α^{-1} -smooth ρ -strongly convex functions. The optimal value of (4) is then equal to the square worst-case contraction factor r^2 .

The seminal work in [9] raised the idea of computing tight bounds on the performance of a method by formulating a maximization problem over a class of functions and initial conditions and then solving it numerically. Later, it has been shown in [10] that the conceptual PEP can often be reformulated as a tractable convex semidefinite program for a large number of methods and function classes. One key idea is to replace the optimization over functions by an optimization over function and gradient values under the constraint that these values are consistent with an actual function in the desired class. Then, one must impose suitable conditions on these discrete variables to guarantee the existence of functions f and g consistent with the discrete points, which are called interpolation conditions. Finally, it is possible to lift the problem as a semidefinite program that can be solved efficiently. In [10], the authors have shown that any fixed-step first-order method (including methods with a proximal operator) on the class of smooth strongly convex functions can be analyzed by PEP. Analyzing a new class of functions with PEP mainly requires deriving an explicit formulation of the corresponding interpolation conditions. Note that PEP allows to consider any performance criterion and initial condition given by a linear combination of the function values and the scalar products between the gradients and the iterates, e.g. $F(x_K) - F^*$, $\|x_K - x^*\|^2$, $\|\nabla F(x_K)\|^2$ where $F := f + g$ or $F := f + h \circ M$ and K is the iteration index.

III. FIRST-ORDER PROXIMAL METHODS TO SOLVE (1)

A. PEP to validate the tightness of theoretical analysis

First, we consider the cases of convex g (i.e. $\mu = 0$). We compare the following five optimization methods and recall for each the upper bound on the worst-case contraction factor r proved in [16]:

- **Gradient method (GM):** Let $x \in \mathcal{H}$, $\tau \in]0, \frac{2}{\alpha^{-1} + \beta^{-1}}[$,

$$\begin{aligned} \Phi_{f,g}x &= x - \tau(\nabla f(x) + \nabla g(x)), & (\text{GM}) \\ r_{\text{GM}}(\tau) &\leq \max\{|1 - \tau\rho|, |1 - \tau(\beta^{-1} + \alpha^{-1})|\}. & (5) \end{aligned}$$

- **Forward-backward splitting (FBS1):** Let $x \in \mathcal{H}$, $\tau \in]0, 2\alpha[$,

$$\begin{aligned} \Phi_{f,g}x &= \text{prox}_{\tau g}(x - \tau\nabla f(x)), & (\text{FBS1}) \\ r_{\text{FBS1}}(\tau) &\leq \max\{|1 - \tau\rho|, |1 - \tau\alpha^{-1}|\}. & (6) \end{aligned}$$

- **Forward-backward splitting (FBS2):** Let $x \in \mathcal{H}$, $\tau \in]0, 2\beta[$,

$$\begin{aligned} \Phi_{f,g}x &= \text{prox}_{\tau f}(x - \tau\nabla g(x)), & (\text{FBS2}) \\ r_{\text{FBS2}}(\tau) &\leq \frac{1}{1 + \tau\rho}. & (7) \end{aligned}$$

- **Peaceman-Rachford splitting (PRS):** Let $x \in \mathcal{H}$, $\tau > 0$,

$$\begin{cases} y &= \text{prox}_{\tau f}(x) \\ \Phi_{f,g}x &= 2\text{prox}_{\tau g}(2y - x) - 2y + x \end{cases} \quad (\text{PRS})$$

$$r_{\text{PRS}}(\tau) = \max_{a \in \{\rho, \alpha^{-1}\}} \frac{|1 - \tau a|}{1 + \tau a}. \quad (8)$$

- **Douglas-Rachford splitting (DRS):** Let $x \in \mathcal{H}$, $\tau > 0$,

$$\begin{cases} y &= \text{prox}_{\tau f}(x) \\ \Phi_{f,g}x &= \text{prox}_{\tau g}(2y - x) - y + x \end{cases} \quad (\text{DRS})$$

$$r_{\text{DRS}}(\tau) \leq \min \left\{ \frac{1 + r_{\text{PRS}}(\tau)}{2}, \frac{1 + \tau^2\rho\beta^{-1}}{(1 + \tau\rho)(1 + \tau\beta^{-1})} \right\}. \quad (9)$$

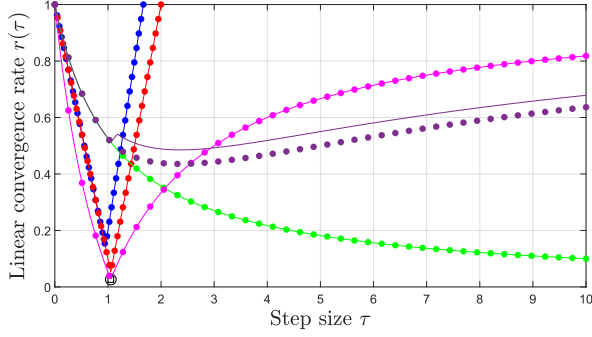
Figure 1 compares the above theoretical rates and those obtained with the PEP analysis described in (4). The results are displayed for three selected configurations of parameters (α, β, ρ) , but are representative of a large number of additional experiments, which are not displayed here.

The numerical results obtained by PEP exactly match the above upper bounds for the first four methods (i.e. GM, FBS1, FBS2, PRS) showing that these upper bounds are tight, i.e. unimprovable. Moreover, inspection of the PEP results suggested that the worst-case functions are univariate and quadratic. Searching for the coefficients of those quadratic functions can then be formulated as a simple optimization problem (see Section IV), whose solution allows us to formally establish tightness of those four bounds.

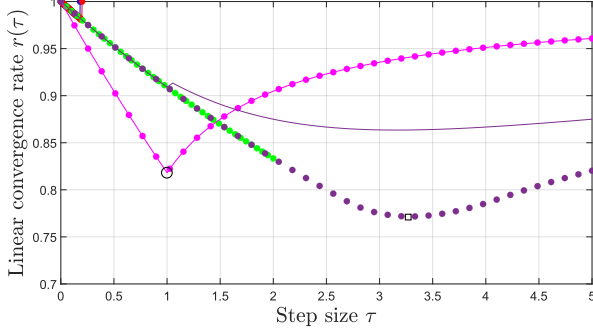
Proposition 1. *Each upper bound (5), (6), (7), and (8) is attained by a quadratic function, and hence describes the exact (unimprovable) worst-case contraction factors of (GM), (FBS1), (FBS2), and (PRS).*

Proof. The quadratic functions are $f(x) = a\frac{x^2}{2}$ and $g(x) = b\frac{x^2}{2}$ with $a \in \{\rho, \alpha^{-1}\}$ and $b \in \{0, \beta^{-1}\}$. \square

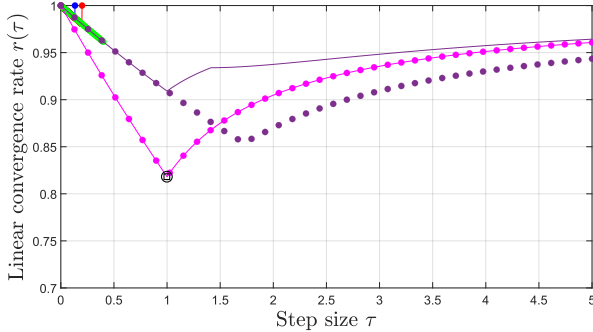
For the fifth method, DRS, Figure 1 shows that the theoretical bound (9) is conservative and can thus be improved. A direct practical consequence of this conservatism is that it may lead us to make sub-optimal choices of methods. For instance, in Figure 1b, the upper bounds obtained by [16] suggested to use PRS with the step-size $\tau = 1$ (black circle). However, DRS behaves much better than predicted by the



(a) $\alpha = 1, \beta = 5, \rho = 0.9, \mu = 0$.



(b) $\alpha = 0.1, \beta = 1, \rho = 0.1, \mu = 0$.



(c) $\alpha = 0.1, \beta = 0.2, \rho = 0.1, \mu = 0$.

Fig. 1: Comparison of the contraction factors from [16] (solid lines) and PEP (dots) of GM (blue), FBS1 (red), FBS2 (green), PRS (magenta), DRS (purple), and the optimal rate predicted by [16] (black circle) and PEP (black square).

existing bound in this case. Indeed, given the exact results provided by PEP among all methods, DRS achieves a better convergence rate when considering a step-size $\tau = 3.3$ (black square in Figure 1b). The configuration displayed in Figure 1b illustrates the drawback of relying on a non-tight analysis of the methods.

B. PEP to evaluate the impact of changing the regularization parameter in the objective function

Knowing the exact worst-case contraction factor of the methods also allows us to understand what is happening when changing the regularizing hyperparameter in $f + \lambda h$, since it can be seen as an instance of our problem with $g = \lambda h$. The

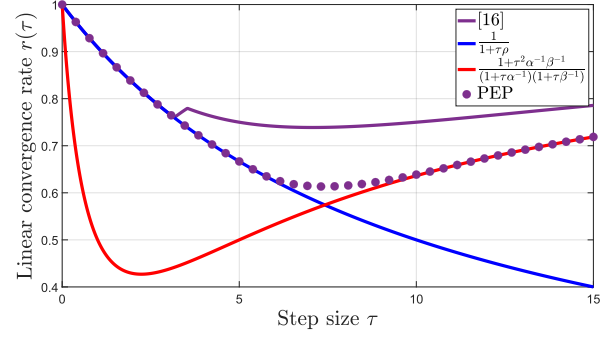


Fig. 2: Worst-case contraction factor of DRS for $\alpha = 1, \beta = 5, \rho = 0.1, \mu = 0$ from [16] (purple solid line) and PEP (purple dots). We identified two regimes (blue and red solid lines).

value of λ directly impacts the smoothness parameter of g . For instance, between Figure 1b and 1c, parameter β is divided by 5, which happens when hyperparameter λ is multiplied by 5. We observe that, as expected, the behavior of FBS2, PRS, and DRS are impacted, and, considering the PEP results, the optimal method is DRS when $\beta = 1$ and PRS when $\beta = 5$.

C. Focus on Douglas-Rachford

Following our observation that the DRS bound is not tight, we now investigate this method in more detail. Figure 2 shows that the rate proved in [16], namely $\frac{1}{1+\tau\rho}$ (displayed in blue), is exact for short step sizes. Looking now at larger step sizes, we find by matching the PEP numerical results and the known analytical performance of the method on simple functions (see Section IV) that the worst-case contraction factor is given by $\frac{1+\tau^2\alpha^{-1}\beta^{-1}}{(1+\tau\alpha^{-1})(1+\tau\beta^{-1})}$ (displayed in red). However, we could not identify the precise behavior of the contraction factor in the middle range (cf. Figure 2 around $\tau = 7$).

Proposition 2. *The exact worst-case contraction factor r_{DRS} of DRS satisfies*

$$r_{DRS}(\tau) \geq \max \left\{ \max_{\substack{a \in \{\rho, \alpha^{-1}\} \\ b \in \{0, \beta^{-1}\}}} \frac{1 + \tau^2 ab}{(1 + \tau a)(1 + \tau b)}, T_0(\tau) \right\} \quad (10)$$

where $T_0(\tau)$ is unidentified.

Proof. The bound is reached by quadratic functions $f(x) = a\frac{x^2}{2}$, $g(x) = b\frac{x^2}{2}$ with $a \in \{\rho, \alpha^{-1}\}$, $b \in \{0, \beta^{-1}\}$ and unidentified functions $f(x) = s_0(x)$ and $g(x) = t_0(x)$ such that $\|\phi_{s_0, t_0} x - \phi_{s_0, t_0} y\| = T_0(\tau)\|x - y\|$ for some x, y . \square

D. Generalization to strongly convex f and g

Based on PEP analysis, we extend the algorithmic comparison proposed in [16] to the case where g is also strongly convex (cf. Figure 3) which was not covered in [16]. Similarly to the convex case $\mu = 0$, we prove in this setting a lower bound on the worst-case contraction factor for the five methods. We summarize the identified analytical expressions in the following proposition.

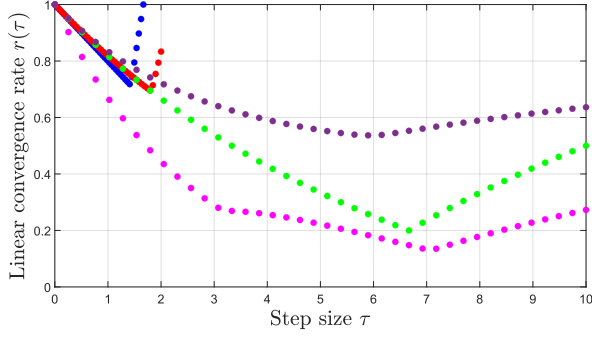


Fig. 3: Contraction factors obtained by PEP of GM (blue), FBS1 (red), FBS2 (green), PRS (magenta), DRS (purple) in the doubly strongly convex case $\alpha = 1$, $\beta = 5$, $\rho = \mu = 0.1$.

Proposition 3. *The exact worst-case contraction factor of GM, FBS1, FBS2, PRS, and DRS respectively satisfy*

$$r_{GM}(\tau) \geq \max\{|1 - \tau(\rho + \mu)|, |1 - \tau(\beta^{-1} + \alpha^{-1})|\}, \quad (11)$$

$$r_{FBS1}(\tau) \geq \frac{1}{1 + \tau\mu} \max\{|1 - \tau\rho|, |1 - \tau\alpha^{-1}|\}, \quad (12)$$

$$r_{FBS2}(\tau) \geq \frac{1}{1 + \tau\rho} \max\{|1 - \tau\mu|, |1 - \tau\beta^{-1}|\}, \quad (13)$$

$$r_{PRS}(\tau) \geq \max_{\substack{a \in \{\rho, \alpha^{-1}\} \\ b \in \{\mu, \beta^{-1}\}}} \frac{|1 - \tau a| |1 - \tau b|}{1 + \tau a |1 + \tau b|}, \quad (14)$$

$$r_{DRS}(\tau) \geq \max \left\{ \max_{\substack{a \in \{\rho, \alpha^{-1}\} \\ b \in \{\mu, \beta^{-1}\}}} \frac{1 + \tau^2 ab}{(1 + \tau a)(1 + \tau b)}, T_\mu(\tau) \right\}, \quad (15)$$

where T_μ is unidentified.

Proof. All the bounds are reached by quadratic functions $f(x) = a\frac{x^2}{2}$, $g(x) = b\frac{x^2}{2}$ with $a \in \{\rho, \alpha^{-1}\}$, $b \in \{\mu, \beta^{-1}\}$ and, for DRS only, unidentified functions $f(x) = s_\mu(x)$, $g(x) = t_\mu(x)$ such that $|\phi_{s_\mu, t_\mu} x - \phi_{s_\mu, t_\mu} y| = T_\mu(\tau) \|x - y\|$ for some x, y . \square

The numerical results of all our experiments were consistent with the contraction factors used as lower bounds in Proposition 3. Therefore, we conjecture that they are exact. As for the case $\mu = 0$, we were not able to identify the regimes T_μ of DRS nor the associated worst-case functions s_μ and t_μ .

IV. LINEAR RATE ON QUADRATIC FUNCTIONS

Interestingly, all worst cases identified so far are attained by univariate quadratic functions. The contraction factor of a method on the class of quadratic functions can be explicitly computed by solving a simple maximization problem.

For example, let us consider PRS applied to the problem $f(x) = a\frac{x^2}{2}$ and $g(x) = b\frac{x^2}{2}$. The proximal step on a quadratic function $f(x) = a\frac{x^2}{2}$ can be explicitly computed as

$$\text{prox}_{\tau f}(x) = \frac{x}{1 + \tau a}. \quad (16)$$

Therefore, after simplifications, PRS becomes

$$\Phi_{f,g} x = \frac{1 - \tau a}{1 + \tau a} \frac{1 - \tau b}{1 + \tau b} x. \quad (17)$$

In this case, problem (4) computing the worst-case contraction factor reduces to

$$r^2(\tau) = \max_{\substack{a \in [\rho, \alpha^{-1}] \\ b \in [\mu, \beta^{-1}]}} \left| \frac{1 - \tau a}{1 + \tau a} \frac{1 - \tau b}{1 + \tau b} \right|^2, \quad (18)$$

namely, a simple optimization problem involving two variables a and b that can often be solved analytically. The same methodology applies to the other methods and also to general quadratic functions (not necessarily univariate).

This methodology allowed us to find the analytical expressions of all the contraction factors gathered in Proposition 3 since the worst-case is achieved with a quadratic function for each of these methods (except for an unidentified non-quadratic worst-case of DRS). Nonetheless, it is important to note that it was not known a priori that the worst-case contraction factors of the methods over general smooth convex functions are attained by quadratic functions. The analysis and numerical results provided by PEP are critical to reach this conclusion.

Conjecture 1. *The worst-case contraction factor of GM, FBS1, FBS2, and PRS on the problem $\min_x f(x) + g(x)$ where $f \in \mathcal{F}_{\rho, \alpha^{-1}}$ and $g \in \mathcal{F}_{\mu, \beta^{-1}}$ are reached by univariate quadratic functions f and g .*

V. CASE $f + h \circ M$

We now move to composite functions of the form (2) involving a linear operator, whose inclusion in the PEP framework was made possible by the very recent work [15]. Specifically, we assume that f is α^{-1} -smooth ρ -strongly convex, h is γ^{-1} -smooth δ -strongly convex, and matrix M such that $\|M\| \triangleq \sigma_{\max}(M) \leq L$ with $\alpha, \rho, L \in]0, \infty[$ and $\gamma, \delta \in [0, \infty[$. This is a particular case of problem (1), therefore, the previous results still apply. However, the composite structure of $g = h \circ M$ can in principle lead to stronger results. We still consider the five same methods and want to quantify the potential gain of performance when applying a method on (2). Note that even if computing the proximal operators of the iterations of the method can be complicated for practical problems, it does not prevent us from easily studying them with PEP.

The only setting where an improvement could be expected is for DRS. Indeed, for the four other methods, the worst-case functions on the class of functions of the form $f + g$ are quadratic. Since, all quadratic functions g can be expressed as a composite function of the form $h \circ M$, the worst-case performance cannot be improved. However, since the worst-case functions of DRS are not identified, this leaves a possibility for an improved rate in the composite case. Still, the PEP numerical results are the same as in Section III, suggesting that there is actually no improvement. It indicates that there exists a worst-case function in the case $f + g$ for

which g is of the form $h \circ M$ (and that is not a quadratic function).

VI. PRIMAL-DUAL METHODS

Unlike the results of [16] for primal methods, no theoretical comparison for proximal primal-dual algorithms seems to be available in the literature so far. We now show how the recent extension of PEP allows us to numerically compute these rates and, therefore, to compare methods. We focus on two primal-dual methods that exploit access to linear operators in their iterations:

- **Chambolle-Pock method (CPM)** [18]: Let x , u , and τ, σ such that $\sigma\tau\|M\|^2 \leq 1$:

$$\begin{cases} x_+ &= \text{prox}_{\tau f}(x - \tau M^T u) \\ u_+ &= \text{prox}_{\sigma h^*}(u + \sigma M(2x_+ - x)). \end{cases} \quad (19)$$

- **Condat-Vũ method (CVM)** [6], [19]: Let x , u and τ, σ such that $\frac{1}{\tau} - \sigma\|M\|^2 \geq \frac{1}{2\alpha}$ (where f is assumed to be α^{-1} -smooth):

$$\begin{cases} x_+ &= x - \tau \nabla f(x) - \tau M^T u \\ u_+ &= \text{prox}_{\sigma h^*}(u + \sigma M(2x_+ - x)). \end{cases} \quad (20)$$

For such primal-dual methods, we are interested in the contraction factor r of the primal-dual iterates (x, u) , defined as

$$\|\phi_{f,g}(x, u) - \phi_{f,g}(x', u')\| \leq r \|(x, u) - (x', u')\|. \quad (21)$$

which cannot be directly compared to the primal contraction factor (3). Similarly to Section V, such worst-case analysis can be formulated and obtained with PEP. Figure 4 below depicts the worst-case contraction factor of CPM (with $\sigma(\tau) = \frac{1}{\tau\|M\|^2}$) and CVM (with $\sigma(\tau) = \frac{1}{\tau\|M\|^2} - \frac{1}{2\alpha\|M\|^2}$) on problem (2) for $\delta = 0$ and $\delta = 0.1$.

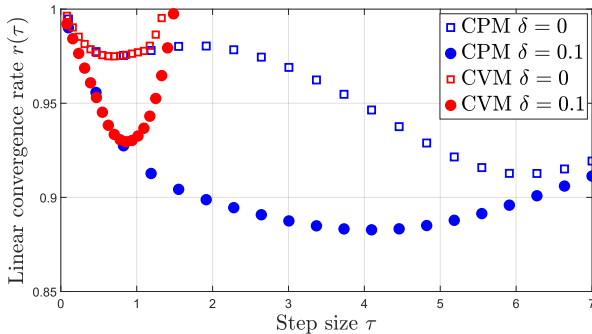


Fig. 4: Contraction factors of CPM (blue) and CVM (red) obtained by PEP for $\alpha = 1$, $\beta = 5$, $\rho = 0.1$, $\delta = 0$ (squares) and $\delta = 0.1$ (dots) and $\|M\| \leq 1$.

We observe the impact of the presence of strong convexity in h . In these two specific settings, CPM performs better.

VII. CONCLUSION

We have shown how the PEP framework allows to easily certify the tightness of (or, in one case, to improve) recent bounds on the contraction factor of the gradient method

and the forward-backward, Peaceman-Rachford, and Douglas-Rachford splitting algorithms, applied to smooth convex functions. We also perform the same analysis in the strongly convex case, obtaining again tight numerical rates. The methodology also highlights interesting and a priori unexpected phenomena, such as the existence of quadratic worst-case functions.

This type of analysis can be extended to any other method, function class, and performance criterion. We demonstrate this possibility on Chambolle-Pock and Condat-Vũ methods opening the way for more extended analysis in further works.

REFERENCES

- [1] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. B. et al., Ed. New York: Springer-Verlag, 2011, pp. 185–212.
- [2] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with sparsity-inducing penalties," *Foundations and Trends® in Machine Learning*, vol. 4, no. 1, pp. 1–106, 2012.
- [3] A. Chambolle and T. Pock, "An introduction to continuous optimization for imaging," *Acta Numerica*, vol. 25, pp. 161–319, 2016.
- [4] N. Pustelnik, A. Benazza-Benhayia, Y. Zheng, and J.-C. Pesquet, "Wavelet-based image deconvolution and reconstruction," *Wiley Encyclopedia of EEE*, 2016.
- [5] A. Nemirovski, "Prox-method with rate of convergence $o(1/t)$ for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems," *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 229–251, 2004.
- [6] L. Condat, "A primal-dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms," *Journal of optimization theory and applications*, vol. 158, no. 2, pp. 460–479, 2013.
- [7] A. Beck, *First-order methods in optimization*. SIAM, 2017.
- [8] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, "Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists," *SIAM Review*, vol. 65, no. 2, pp. 375–435, 2023.
- [9] Y. Drori and M. Teboulle, "Performance of first-order methods for smooth convex minimization: a novel approach," *Mathematical Programming*, vol. 145, no. 1-2, pp. 451–482, 2014.
- [10] A. B. Taylor, J. M. Hendrickx, and F. Glineur, "Smooth strongly convex interpolation and exact worst-case performance of first-order methods," *Mathematical Programming*, vol. 161, pp. 307–345, 2017.
- [11] G. H. Chen and R. T. Rockafellar, "Convergence rates in forward-backward splitting," *SIAM Journal on Optimization*, vol. 7, no. 2, pp. 421–444, 1997.
- [12] A. Beck and M. Teboulle, "Gradient-based algorithms with applications to signal recovery," *Convex optimization in signal processing and communications*, pp. 42–88, 2009.
- [13] A. B. Taylor, J. M. Hendrickx, and F. Glineur, "Exact worst-case performance of first-order methods for composite convex optimization," *SIAM Journal on Optimization*, vol. 27, no. 3, pp. 1283–1313, 2017.
- [14] E. K. Ryu, A. B. Taylor, C. Bergeling, and P. Giselsson, "Operator splitting performance estimation: Tight contraction factors and optimal parameter selection," *SIAM Journal on Optimization*, vol. 30, no. 3, pp. 2251–2271, 2020.
- [15] N. Bousselmi, J. M. Hendrickx, and F. Glineur, "Interpolation conditions for linear operators and applications to performance estimation problems," *arXiv preprint arXiv:2302.08781*, 2023.
- [16] L. M. Briceno-Arias and N. Pustelnik, "Theoretical and numerical comparison of first order algorithms for cocoercive equations and smooth convex optimization," *Signal Processing*, vol. 206, p. 108900, 2023.
- [17] —, "Convergence rate comparison of proximal algorithms for non-smooth convex optimization with an application to texture segmentation," *IEEE Signal Processing Letters*, vol. 29, pp. 1337–1341, 2022.
- [18] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, pp. 120–145, 2011.
- [19] B. C. Vũ, "A splitting algorithm for dual monotone inclusions involving cocoercive operators," *Advances in Computational Mathematics*, vol. 38, no. 3, pp. 667–681, 2013.