Error Estimates for Doubly-Generalized Tikhonov-Phillips Regularization

M. J. CARRIÓ[1*], G. L. MAZZIERI[2] and K. G. TEMPERINI[3]

ABSTRACT. In this work, error estimates are presented for the case in which the regularized solution is obtained by minimizing doubly-generalized Tikhonov-Phillips functionals. The first result is based mainly on an assumption given by a source condition. It is proved that it is possible to replace this assumption by a variational inequality, obtaining analogous result of the error estimate. Finally, relationships are established between the optimality condition associated with the problem, the source condition and the variational inequality. On the other hand, it is known that, in certain cases, the use of two or more penalizing terms is useful. For this reason, generalizations of the results of error estimates are presented for cases in which the regularized solution is a minimizer of doubly-generalized Tikhonov-Phillips functionals with multiple penalizers.

Keywords: inverse problems, Tikhonov-Phillips, error estimate, source condition, variational inequality.

1 INTRODUCTION

In a quite general framework, an inverse problem can be defined as the need of determining x in an equation of the form

$$\mathscr{T}x = y, \tag{1.1}$$

where $\mathscr{T} : \mathscr{X} \to \mathscr{Y}$ is a bounded linear operator between two Banach spaces of infinite dimension (in the classic theory, $\mathscr{X}$ and $\mathscr{Y}$ are Hilbert spaces) and y is the data, supposed to be known, perhaps with a certain degree of error. Frequently, inverse problems are *ill-posed* in the sense of

*Corresponding author: María Josefina Carrió – E-mail: mjcarrio@fbcb.unl.edu.ar

[1]Departamento de Matemática, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Ciudad Universitaria - Paraje El Pozo, C.P. 3000, Santa Fe, Argentina – E-mail: mjcarrio@fbcb.unl.edu.ar https://orcid.org/0000-0002-8676-4135

[2]Instituto de Matemática Aplicada del Litoral, IMAL, CONICET-UNL, Centro Científico Tecnológico CONICET-Santa Fe, Colectora Ruta Nac. 168, km. 472, Paraje "El Pozo", C.P. 3000, Santa Fe, Argentina. Departamento de Matemática, Facultad de Bioquímica y Ciencias Biológicas, Universidad Nacional del Litoral, Ciudad Universitaria - Paraje El Pozo, C.P. 3000, Santa Fe, Argentina. – E-mail: glmazzieri@santafe-conicet.gov.ar https://orcid.org/0000-0002-5368-9601

[3]Instituto de Matemática Aplicada del Litoral, IMAL, CONICET-UNL, Centro Científico Tecnológico CONICET-Santa Fe, Colectora Ruta Nac. 168, km. 472, Paraje "El Pozo", C.P. 3000, Santa Fe, Argentina. Departamento de Matemática, Facultad de Humanidades y Ciencias, Universidad Nacional del Litoral, Ciudad Universitaria - Paraje El Pozo, C.P. 3000, Santa Fe, Argentina – E-mail: ktemperini@santafe-conicet.gov.ar https://orcid.org/0000-0002-0549-7453

Hadamard (i.e., the solution does not exists, it is not unique or it does not depend continuously on the data) and thus arises the need to apply a regularization method. Associated with this method, there is a parameter called the "regularization parameter" and its choice is essential to achieve an adequate approximation of the solution of the inverse problem. There are several parameter choice rules that allow determining its value. The so-called "a-priori" are those that depend only on the noise level of the problem while the "a-posteriori" are those that depend on both the noise level and the data. Finally, there is another type of rules called "heuristics" and they are those that only depend on the noise level through the data. Among them, we can mention the "generalized cross-validation method" introduced in 1979 by G. H. Golub, M. T. Heath and G. Wahba [8], "the L-curve criterion" proposed in 1992 by P. C. Hansen [9] and the rule proposed by K. Ito, B. Jin and J. Zou in 2011 [13], whose construction is based on the stochastic approach to solving an inverse problem. It is important to mention that this type of rule is very useful in those problems where the exact noise level is unknown and only the data is available.

Although there is a wide variety of regularization methods, probably the best known and most commonly and widely used is the Tikhonov–Phillips regularization method, which was originally proposed by D. L. Phillips and A. N. Tikhonov in 1962 and 1963, respectively [16, 19, 20]. Under the classic version of this method (with fidelity term and penalizer, both quadratic) and its generalized version (with quadratic fidelity term and generalized penalizer) it is possible to prove the convergence of the regularized solutions to a minimum penalizing solution, in the case in which the regularization parameter is chosen through an a-priori or a-posteriori rule [3, 5]. It is possible to generalize these convergence results for the doubly-generalized Tikhonov-Phillips regularization method (that is, with generalized fidelity and penalizing terms) for these types of rules [6, 11]. A result due to A. B. Bakushinskii [1] shows that a regularization method cannot be convergent when the associated parameter choice rule depends only on the data of the ill-posed inverse problem. But this does not indicate that the method cannot perform well for small noise levels. For this reason, given that convergence results cannot be obtained, various authors have presented error estimates, that is, bounds for the error between the solution of the problem and the regularized solution when the regularization parameter is chosen using a heuristic rule for a fixed noise level [2, 11, 18].

We propose here to study the doubly-generalized Tikhonov-Phillips regularization method that consists in approximating the solution of the problem (1.1) by a minimizer of the functional, called "TPGG-p functional", given by

$$\mathscr{I}^{p}_{\tilde{\phi},\psi,\eta}(x) \doteq \frac{1}{p}\,\tilde{\phi}(\mathscr{T}x,y)^{p} + \eta\,\psi(x), \tag{1.2}$$

where $p \geq 1$, $\tilde{\phi} : \mathscr{Y} \times \mathscr{Y} \to [0,+\infty)$, $\psi : \mathscr{X} \to [0,+\infty)$ is a convex functional and $\eta > 0$ is the regularization parameter. In particular, if $\tilde{\phi}(\mathscr{T}x,y) = \|\mathscr{T}x - y\|$, the resulting functional is called "TPG-p functional" and is given by

$$\mathscr{I}^{p}_{\psi,\eta}(x) \doteq \frac{1}{p}\,\|\mathscr{T}x - y\|^{p} + \eta\,\psi(x), \tag{1.3}$$

where $p \geq 1$ and $\psi : \mathscr{X} \to [0, +\infty)$ is a convex functional. This work presents error estimates that we have proved for the case in which the regularized solution is obtained by minimizing the TPGG-p functionals given by (1.2). The first estimates are based mainly on an assumption given by a source condition which, as is known, imposes certain smoothness conditions on the minimum penalizing solution. These estimates are particularized for the case in which $\tilde{\phi}(\mathscr{T}x, y) = \|\mathscr{T}x - y\|$ and $p = 2$, leaving evidence that are equivalent to that proposed by K. Ito and B. Jin in [11].

As it is usual, we use the Bregman distance to estimate the error. This distance is a natural way to measure the deviation of elements in a Banach space with respect to a convex functional and it has a term that will be in our interest to be able to bounded. For this purpose, the so-called "variational inequalities" will be used to bound this term of interest and obtain error estimates. We will see here that it is possible to replace the source condition assumption used in the first error estimates by a variational inequality, obtaining similar results from that estimates. Finally, relationships are established between the optimality condition associated with the problem, the source condition and the variational inequality.

It is well known that an adequate choice of the penalizing term, based on the "a-priori" knowledge of certain type of information about the exact solution, will result in regularized solutions which appropriately reflect those characteristics. In recent years, there has been a growing interest in the multi-parameter Tikhonov regularization method that uses multiple constraints as a means of improving the quality of inversion. The multi-parameter regularization adds multiple different penalties which exhibit multi-scale features, while the single-parameter regularization uses a unique penalty which may result in a regularized solution that does not preserve certain features of the original solution. The use of multi-parameter regularization for solving ill-posed problems naturally matches with the multi-resolution analysis framework which has become a standard method to analyze the frequency information of images on different resolutions [12, 14, 15, 21]. For this reason, we present generalizations of the error estimates results obtained for the functionals given in (1.2) for the case in which the regularized solution is a minimizer of doubly-generalized Tikhonov-Phillips functional with multiple penalizers given by

$$\mathscr{J}^{p}_{\tilde{\phi}, \vec{\psi}, \vec{\eta}}(x) \doteq \frac{1}{p} \tilde{\phi}(\mathscr{T}x, y)^p + \vec{\eta} \cdot \vec{\psi}(x), \tag{1.4}$$

where $p \geq 1$, $\tilde{\phi} : \mathscr{Y} \times \mathscr{Y} \to [0, +\infty)$, $\vec{\psi} \doteq (\psi_1, \ldots, \psi_n) : \mathscr{X} \to \mathbb{R}^n$ with $\psi_i : \mathscr{X} \to [0, +\infty)$ convex functionals for all $i = 1, \ldots, n$ and $\vec{\eta} = (\eta_1, \ldots, \eta_n) \in \mathbb{R}^n$ such that $\eta_i > 0$, for all $i = 1, \ldots, n$.

2   PRELIMINARES

As it was previously mentioned, to estimate the a-posteriori error, the Bregman distance will be used and, to present its definition, it is necessary to introduce the concepts of subgradient and subdifferential of a convex functional [18], which are presented below.

Definition 2.1. *Let $\mathscr{X}$ be a Banach space, $\mathscr{X}^*$ the dual space of X and $\psi : \mathscr{X} \to \mathbb{R} \cup \{\infty\}$ a convex functional. Then, $x^* \in \mathscr{X}^*$ is a subgradient of ψ in x if $\psi(z) \geq \psi(x) + [x^*, z - x]$, $\forall z \in \mathscr{X}$,*

where $[x^*, x]$ corresponds to the functional x^* evaluated in x, i.e $[x^*, x] \doteq x^*(x)$. The set $\partial \psi(x)$ of all subgradients of ψ in x is called subdifferential of ψ in x.

It is important to mention here that the subgradient is a generalization of the classical concept of derivative for the case of convex functionals [22].

Definition 2.2. *Let $\mathscr{X}$ be a Banach space, $\psi : \mathscr{X} \to \mathbb{R} \cup \{\infty\}$ a convex functional, $x \in \mathscr{X}$ and $\xi \in \partial \psi(x)$. The Bregman distance in x with respect to ξ and ψ is defined as*

$$d_\xi^\psi(z, x) \doteq \psi(z) - \psi(x) - [\xi, z - x], \ \forall z \in \mathscr{X}.$$

It is immediate to see that $d_\xi^\psi(x, x) = 0$ and the convexity of the functional ψ implies that $d_\xi^\psi(z, x) \geq 0$, for all $z \in \mathscr{X}$. Also, it is easy to prove that if $\mathscr{X}$ is a Hilbert space, the Bregman distance in an element x coincides with $\|x - z\|^2$, for all $z \in \mathscr{X}$, when the functional ψ is chosen as the square of the norm in $\mathscr{X}$. In this way, the Bregman distance extends the concept of norm.

Now, we introduce the concept of duality mapping that will be relevant for its close relationship with the subdifferential of a power of a norm [17, 18]. To define this mapping, it is necessary to introduce the concept of a "gauge function" that consists of a continuous and strictly increasing function $f : [0, +\infty) \to [0, +\infty)$ that verifies $f(0) = 0$ and $\lim_{t \to \infty} f(t) = \infty$. Furthermore, it is necessary to define the set denoted by $2^{\mathscr{X}^*}$ consisting of all subsets of the dual space of $\mathscr{X}$, that is, $2^{\mathscr{X}^*} \doteq \{E : E \subset \mathscr{X}^*\}$.

Definition 2.3. *Let $\mathscr{X}$ be a Banach space. The duality mapping of $\mathscr{X}$ with respect to the gauge function f is the (set-valued) mapping $J : \mathscr{X} \to 2^{\mathscr{X}^*}$ defined by*

$$J(x) = \{x^* \in \mathscr{X}^* : [x^*, x] = \|x^*\| \, \|x\|, \ \|x^*\| = f(\|x\|)\}.$$

The following proposition presents the duality mapping with respect to the gauge function $f(t) = t^{p-1}$ with $p > 1$, for the case of Hilbert spaces.

Proposition 2.1. *Let $\mathscr{X}$ be a Hilbert space and $f(t) = t^{p-1}$ with $p > 1$. The duality mapping J of $\mathscr{X}$ with respect to f is given by $J(x) = \|x\|^{p-2} x$.*

The next result, called *Asplund Theorem*, relates the subdifferential of the primitive of a gauge function (which is convex) and the duality mapping with respect to that function.

Theorem 2.1. *Let $\mathscr{X}$ be a Banach space and f a gauge function. If $F(t) \doteq \int_0^t f(s) \, ds$ then the duality mapping of $\mathscr{X}$ with respect to f is given by*

$$J(x) = \partial F(\|x\|), \ x \in \mathscr{X}.$$

The following corollary is useful to obtain the subdifferential of the convex functional $\dfrac{1}{p}\|x\|^p$ with $p > 1$, that is associated with the fidelity term considered in (1.3).

Corollary 2.1. *If $\mathscr{X}$ is a Banach space and $f(t) = t^{p-1}$ with $p > 1$, then the duality mapping J of $\mathscr{X}$ with respect to f is given by $J(x) = \partial\left(\dfrac{1}{p}\|x\|^p\right)$.*

It is well known that a minimizer of a differentiable and convex function verifies the optimality condition, that is, the derivative of that function vanishes in that minimizer. Similarly, a minimizer z of a convex functional $\mathscr{J}$ defined on a Banach space satisfies the optimality condition given by $0 \in \partial\mathscr{J}(z)$, as it is presented in the following result. Furthermore, the optimality condition implies the existence of a minimizer.

Theorem 2.2. *Let $\mathscr{X}$ be a Banach space, $\mathscr{J} : \mathscr{X} \to \mathbb{R} \cup \{\infty\}$ a convex functional and $z \in \mathscr{D}(\mathscr{J})$ with $\mathscr{J}(z) < \infty$. Then, $\mathscr{J}(z) = \min\limits_{x \in \mathscr{X}} \mathscr{J}(x)$ if and only if $0 \in \partial\mathscr{J}(z)$.*

3 PRINCIPAL RESULTS

3.1 One penalizer

The following result presents the first error estimation that we have obtained for the case in which the regularized solution is a minimizer of a TPGG-p functional. As will be seen, it is possible to particularize this estimate for the case of TPG-p functionals. More precisely, it will be seen that the estimate obtained for the case of TPG-2 functionals is equivalent to that proposed by K. Ito and B. Jin in [11].

Theorem 3.3. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces, $\mathscr{T} \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, $\psi : \mathscr{X} \to [0, +\infty)$, $\tilde{\phi} : \mathscr{Y} \times \mathscr{Y} \to [0, +\infty)$, $\tilde{x}$ the exact solution of (1.1) and $y^\delta \in \mathscr{Y}$ such that $\tilde{\phi}(y, y^\delta) \leq \delta$ and $\tilde{\phi}(y^\delta, y) \leq \delta$. Suppose that there exists $w \in \mathscr{Y}^*$ such that $\xi \doteq \mathscr{T}^\#w \in \partial\psi(\tilde{x})$ and that the functional $\tilde{\phi}$ satisfies the following hypotheses:*

1. $\tilde{\phi}(y_1, y_3) \leq \tilde{\phi}(y_1, y_2) + \tilde{\phi}(y_2, y_3)$, *for all* $y_1, y_2, y_3 \in \mathscr{Y}$;

2. *there exists* $C > 0$ *such that* $\|y_1 - y_2\| \leq C\tilde{\phi}(y_2, y_1)$, $\forall y_1, y_2 \in \mathscr{R}(\mathscr{T})$.

Then, for each minimizer x_η^δ of the functional $\mathscr{J}_{\tilde{\phi},\psi,\eta}^p$ given in (1.2) for data y^δ it follows that

- *if $p > 1$ and q is the conjugate of p then*

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\frac{\delta^p}{p} + \eta\,\delta\,\|w\|\,C + \frac{1}{q}(\eta\,\|w\|\,C)^q}{\eta};$$

- *if $p = 1$ and $\eta\,\|w\|\,C \leq 1$ then*

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta\,\frac{1 + \eta\,\|w\|\,C}{\eta}.$$

Proof. Since x_η^δ is a minimizer of the functional $\mathscr{J}_{\tilde{\phi},\psi,\eta}^p$ for data y^δ, it follows that

$$\frac{1}{p}\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p + \eta\,\psi(x_\eta^\delta) \leq \frac{1}{p}\tilde{\phi}(\mathscr{T}\tilde{x}, y^\delta)^p + \eta\,\psi(\tilde{x}),$$

and, in consequence, $\frac{1}{p}\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p \le \frac{1}{p}\tilde{\phi}(\mathscr{T}\tilde{x}, y^\delta)^p + \eta\left(\psi(\tilde{x}) - \psi(x_\eta^\delta)\right).$

Then,

$$\frac{1}{p}\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p + \eta\, d_\xi^\psi(x_\eta^\delta, \tilde{x}) \le \frac{1}{p}\tilde{\phi}(\mathscr{T}\tilde{x}, y^\delta)^p + \eta\left(\psi(\tilde{x}) - \psi(x_\eta^\delta) + d_\xi^\psi(x_\eta^\delta, \tilde{x})\right)$$

$$= \frac{1}{p}\tilde{\phi}(\mathscr{T}\tilde{x}, y^\delta)^p + \eta\left[\mathscr{T}^\# w, \tilde{x} - x_\eta^\delta\right]$$

$$= \frac{1}{p}\tilde{\phi}(\mathscr{T}\tilde{x}, y^\delta)^p + \eta\left[w, \mathscr{T}\tilde{x} - \mathscr{T}x_\eta^\delta\right], \qquad (3.1)$$

where the last two equalities follow from the source condition and the definition of the dual adjoint operator, respectively. Now, since $\tilde{\phi}(y, y^\delta) \le \delta$ and $\tilde{\phi}(y^\delta, y) \le \delta$, it follows that

$$\frac{1}{p}\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p + \eta\, d_\xi^\psi(x_\eta^\delta, \tilde{x})$$

$$\le \frac{\delta^p}{p} + \eta\,[w, \mathscr{T}\tilde{x} - \mathscr{T}x_\eta^\delta] \qquad\qquad \text{by (3.1) and } \mathscr{T}\tilde{x} = y$$

$$\le \frac{\delta^p}{p} + \eta\,\|w\|\,\|\mathscr{T}\tilde{x} - \mathscr{T}x_\eta^\delta\|$$

$$\le \frac{\delta^p}{p} + \eta\,\|w\|\,C\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) + \eta\,\|w\|\,C\,\tilde{\phi}(y^\delta, y) \qquad \text{by hypotheses 1, 2 and } \mathscr{T}\tilde{x} = y$$

$$\le \frac{\delta^p}{p} + \eta\,\|w\|\,C\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) + \eta\,\|w\|\,C\,\delta.$$

Finally,

$$\frac{1}{p}\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p - \eta\,\|w\|\,C\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) + \eta\, d_\xi^\psi(x_\eta^\delta, \tilde{x}) \le \frac{\delta^p}{p} + \eta\,\|w\|\,C\,\delta. \qquad (3.2)$$

- If $p > 1$, it immediately follows from Young's inequality that

$$\eta\,\|w\|\,C\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) \le \frac{1}{p}\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p + \frac{1}{q}(\eta\,\|w\|\,C)^q,$$

which implies that

$$\eta\, d_\xi^\psi(x_\eta^\delta, \tilde{x}) - \frac{1}{q}(\eta\,\|w\|\,C)^q \le \frac{1}{p}\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p - \eta\,\|w\|\,C\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) + \eta\, d_\xi^\psi(x_\eta^\delta, \tilde{x})$$

$$\le \frac{\delta^p}{p} + \eta\,\|w\|\,C\,\delta,$$

where the last inequality follows from (3.2). Thus,

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \le \frac{\frac{\delta^p}{p} + \eta\,\delta\,\|w\|\,C + \frac{1}{q}(\eta\,\|w\|\,C)^q}{\eta}.$$

- If $p = 1$, by (3.2) it follows that $(1 - \eta\,\|w\|\,C)\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) + \eta\, d_\xi^\psi(x_\eta^\delta, \tilde{x}) \le \delta + \eta\,\|w\|\,C\,\delta$, and by hypothesis $1 - \eta\,\|w\|\,C \ge 0$, finally, it turns out that

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \le \delta\,\frac{1 + \eta\,\|w\|\,C}{\eta}.$$

$\square$

Given that every norm satisfies the triangular inequality and considering $C = 1$, it follows immediately that the hypotheses 1 and 2 of Theorem 3.3 are verified when $\tilde{\phi}$ corresponds to the norm in $\mathscr{Y}$. In this way, if the fidelity term is $\dfrac{1}{p}\|\mathscr{T}x - y\|^p$, with $p \geq 1$, from Theorem 3.3 error estimates are obtained for the case of TPG-p functionals. In particular, for $p = 2$ (corresponding to the classic fidelity term) it turns out that

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\frac{\delta^2}{2} + \eta\delta\|w\| + \frac{1}{2}(\eta\|w\|)^2}{\eta} = \frac{1}{2}\left(\frac{\delta}{\sqrt{\eta}} + \sqrt{\eta}\|w\|\right)^2,$$

which coincides with that of K. Ito and B. Jin in [11]. It is then possible to conclude that the Theorem 3.3 consists of a generalization of the estimation result presented by these authors.

On the other hand, it is appropriate to mention that the inequality (3.1), which constitutes a first error estimate, was obtained by M. Benning and M. Burger in 2011 under slightly different hypotheses about $\tilde{\phi}$ and ψ, using the symmetric Bregman distance [2].

Finally, it is important to observe that the Theorem 3.3 would allow obtaining orders of convergence for the case in which the regularization parameter is chosen with an a-priori or a-posteriori rule. However, if one is interested in obtaining estimates of the error for the case of heuristic rules, one can, from the inequality (3.2), obtain such an estimate as follows

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\delta^p}{p\eta} + \|w\|C\delta - \frac{1}{p}\frac{\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p}{\eta} + \|w\|C\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta). \tag{3.3}$$

From the proof of the Theorem 3.3, it is easy to see that the error estimate given by the inequality (3.3) is less than or equal to the estimate obtain in this theorem.

Next, it is proved that, in the context of Hilbert spaces and under certain hypotheses, the optimality condition for the functional $\mathscr{J}_{\psi,\eta}^2$, with $\psi(x) \doteq \dfrac{1}{2}\|x\|^2$ (given by the existence of minimizers) implies the source condition present in the statement of Theorem 3.3. It should be noted here that, in this context, the dual adjoint operator $\mathscr{T}^\#$ coincides with the adjoint operator $\mathscr{T}^* : \mathscr{Y} \to \mathscr{X}$ of $\mathscr{T}$.

Proposition 3.2. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Hilbert spaces, $\psi(x) \doteq \dfrac{1}{2}\|x\|^2$ and $\tilde{x}$ the exact solution of (1.1). Suppose that $\{x_\eta\}$ is a sequence of minimizers of the functional $\mathscr{J}_{\psi,\eta}^2$ such that $x_\eta \xrightarrow{\omega} \tilde{x}$ when $\eta \to 0^+$ and there exists $w \in \mathscr{Y}$ such that $-\dfrac{\mathscr{T}x_\eta - y}{\eta} \xrightarrow{\omega} w$. Then, $\mathscr{T}^*w \in \partial\psi(\tilde{x})$, where $\mathscr{T}^* : \mathscr{Y} \to \mathscr{X}$ is the adjoint operator of $\mathscr{T}$.*

Proof. Since x_η is a minimizer of the functional $\mathscr{J}_{\psi,\eta}^2$, from Theorem 2.2 it turns out that

$$0 \in \partial\left(\frac{1}{2}\|\mathscr{T}\cdot - y\|^2 + \eta\frac{1}{2}\|\cdot\|^2\right)(x_\eta).$$

Then, by property of the sum, product by a scalar, translation and composition of the subgradient [22] it follows that

$$\partial \left(\frac{1}{2}\|\mathscr{T}\cdot -y\|^2 + \eta\frac{1}{2}\|\cdot\|^2 \right)(x_\eta) = \partial \left(\frac{1}{2}\|\mathscr{T}\cdot -y\|^2 \right)(x_\eta) + \eta\,\partial \left(\frac{1}{2}\|\cdot\|^2 \right)(x_\eta)$$

$$= \mathscr{T}^*\partial \left(\frac{1}{2}\|\cdot -y\|^2 \right)(\mathscr{T}x_\eta) + \eta\,\partial \left(\frac{1}{2}\|\cdot\|^2 \right)(x_\eta)$$

$$= \mathscr{T}^*\partial \left(\frac{1}{2}\|\cdot\|^2 \right)(\mathscr{T}x_\eta - y) + \eta\,\partial \left(\frac{1}{2}\|\cdot\|^2 \right)(x_\eta).$$

Thus, $0 \in \mathscr{T}^*\left(\partial\frac{1}{2}\|\cdot\|^2 \right)(\mathscr{T}x_\eta - y) + \eta\,\partial \left(\frac{1}{2}\|\cdot\|^2 \right)(x_\eta)$ and the functional 0 can be decomposed as $0 = f_\eta + g_\eta$, where $f_\eta \in \mathscr{T}^*\partial \left(\frac{1}{2}\|\cdot\|^2 \right)(\mathscr{T}x_\eta - y)$ and $g_\eta \in \eta\,\partial \left(\frac{1}{2}\|\cdot\|^2 \right)(x_\eta)$. Since $\mathscr{X}$ and $\mathscr{Y}$ are Hilbert spaces, from Proposition 2.1 and Corollary 2.1, it immediately follows that $f_\eta = \mathscr{T}^*(\mathscr{T}x_\eta - y)$ and $g_\eta = \eta\,x_\eta$. Since $g_\eta = -f_\eta$ it turns out that

$$x_\eta = \frac{-\mathscr{T}^*(\mathscr{T}x_\eta - y)}{\eta}.$$

and then, by hypothesis, $x_\eta \xrightarrow{\omega} \mathscr{T}^*w$. By uniqueness of the limit it is concluded that $\tilde{x} = \mathscr{T}^*w$. From Proposition 2.1 and Corollary 2.1 it follows that $\partial\psi(\tilde{x}) = \partial \left(\frac{1}{2}\|\cdot\|^2 \right)(\tilde{x}) = \{\tilde{x}\}$, and thus $\mathscr{T}^*w \in \partial\psi(\tilde{x})$. $\qquad\square$

Finally, the optimality condition associated with the functional TPGG-p with $p > 1$, and convex and weakly lower semicontinuous penalizer imply the source condition present in the statement of Theorem 3.3, as it is proved in the following proposition.

Proposition 3.3. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces, $\psi : \mathscr{X} \to [0,+\infty)$ a convex and weakly lower semicontinuous functional and $\tilde{x}$ the exact solution of (1.1). Suppose that $\{x_\eta\}$ is a sequence of minimizers of the functional $\mathscr{J}^p_{\tilde{\phi},\psi,\eta}$ with $p > 1$ such that $x_\eta \xrightarrow{\omega} \tilde{x}$ when $\eta \to 0^+$ and that the following hypothesis is satisfied:*

for each $\eta > 0$, there exists $w_\eta \in \mathscr{Y}^*$ such that $\partial \left(-\frac{1}{\eta}\frac{\tilde{\phi}(\mathscr{T}\cdot,y)^p}{p} \right)(x_\eta) = \{w_\eta\}$. $\qquad$ (3.4)

If there exists $w \in \mathscr{Y}^*$ such that $w_\eta \xrightarrow{\omega} w$, then $\mathscr{T}^\#w \in \partial\psi(\tilde{x})$, where $\mathscr{T}^\# : \mathscr{Y}^* \to \mathscr{X}^*$ is the dual adjoint operator of $\mathscr{T}$.

Proof. Since x_η is a minimizer of the functional $\mathscr{J}^p_{\tilde{\phi},\psi,\eta}$, from Theorem 2.2 it immediately follows that $0 \in \partial \left(\frac{1}{p}\tilde{\phi}(\mathscr{T}\cdot,y)^p + \eta\,\psi(\cdot) \right)(x_\eta)$. Then, by property of the sum, product by a scalar and composition of the subgradient [22] it follows that

$$0 \in \mathscr{T}^\#\partial \left(\frac{1}{p}\tilde{\phi}(\mathscr{T}\cdot,y)^p \right)(x_\eta) + \eta\,\partial\psi(x_\eta),$$

and then, the functional 0 can be decomposed as $0 = f_\eta + g_\eta$, with

$$f_\eta \in \mathscr{T}^\#\partial \left(\frac{1}{p}\tilde{\phi}(\mathscr{T}\cdot,y)^p \right)(x_\eta) \text{ and } g_\eta \in \eta\,\partial\psi(x_\eta).$$

Now, by hypothesis, it is known that $\partial \left(-\dfrac{1}{\eta} \dfrac{\tilde{\phi}(\mathcal{T}\cdot,y)^p}{p} \right)(x_\eta) = \{w_\eta\}$ and thus $-\dfrac{f_\eta}{\eta} = \mathcal{T}^{\#} w_\eta$. Since $g_\eta = -f_\eta$, it follows that $\mathcal{T}^{\#} w_\eta = \frac{g_\eta}{\eta} \in \partial \psi(x_\eta)$. Analogously to the proof of Proposition 3.2, as a consequence of the weak lower semicontinuity of ψ and since $w_\eta \xrightarrow{\omega} w$ and $x_\eta \xrightarrow{\omega} \tilde{x}$ as $\eta \to 0^+$, it is proved that $\mathcal{T}^{\#} w \in \partial \psi(\tilde{x})$. $\qquad \square$

It should be mentioned here that, under certain additional hypotheses, the assumption (3.4) about the subdifferential of $\dfrac{1}{p}\tilde{\phi}(\mathcal{T}\cdot,y)^p$ with $p > 1$ is verified. For example, if $\tilde{\phi}$ corresponds to the norm in a Hilbert space, from Proposition 2.1 it follows that the subdifferential $\partial \left(\dfrac{1}{p}\tilde{\phi}(\mathcal{T}\cdot,y)^p \right)(x_\eta)$ has a single element. On the other hand, if $\tilde{\phi}$ is Gateaux differentiable then $\partial \tilde{\phi}(x) = \left\{ \nabla \tilde{\phi}(x) \right\}$ [18].

Because the Bregman distance in $\tilde{x} \in \mathcal{X}$ with respect to the convex functional ψ involves the term $[\xi, x - \tilde{x}]$, where $\xi \in \partial \psi(\tilde{x})$ and $x \in \mathcal{X}$, several authors propose the use of inequalities containing this term, called *variational inequalities*, and it is proved that they are a powerful tool to obtain convergence rates [7, 10, 17]. The first results of convergence rates for Tikhonov-Phillips functional minimizers are based on assumptions of the smoothness of the solution with respect to an operator (generally non-linear) defined in a Hilbert space [4] or Banach [3]. These assumptions are expressed in terms of a source condition (generally associated with an equation). However, numerical observations showed that, even when the assumptions of smoothness were not verified, the convergence rate was not necessarily significantly affected. In 2007, B. Hofmann et al. [10] took this observation into account and weakened these assumptions by replacing the source condition with a variational inequality. We propose here to use inequalities of this type to prove a result analogous to that of Theorem 3.3 and thus obtain error estimates under this new approach. First, we will show the relationship between the source condition present in Theorem 3.3 and a variational inequality.

Proposition 3.4. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces, $\mathcal{T} \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\psi : \mathcal{X} \to [0, +\infty)$, $\tilde{\phi} : \mathcal{Y} \times \mathcal{Y} \to [0, +\infty)$, $\tilde{x}$ the exact solution of (1.1) such that there exists $w \in \mathcal{Y}^*$ with $\xi \doteq \mathcal{T}^{\#} w \in \partial \psi(\tilde{x})$. Suppose that the following hypothesis is satisfied:*

$$\text{there exists } C > 0 \text{ such that } \|y_1 - y_2\| \leq C \tilde{\phi}(y_2, y_1), \forall y_1, y_2 \in \mathcal{R}(\mathcal{T}). \qquad (3.5)$$

Then, there exists $\beta \geq 0$ such that the following variational inequality holds:

$$[\xi, \tilde{x} - x] \leq \beta \tilde{\phi}(\mathcal{T}x, \mathcal{T}\tilde{x}), \forall x \in \mathcal{X}. \qquad (3.6)$$

Proof. It is easy to proved that for each $x \in \mathcal{X}$, by hypotheses and by the definition of the dual adjoint operator, it follows that

$$[\xi, \tilde{x} - x] = [\mathcal{T}^{\#} w, \tilde{x} - x] = [w, \mathcal{T}\tilde{x} - \mathcal{T}x] \leq \|w\| \|\mathcal{T}\tilde{x} - \mathcal{T}x\| \leq \|w\| C \tilde{\phi}(\mathcal{T}x, \mathcal{T}\tilde{x}),$$

whence inequality (3.6) is verified considering $\beta \doteq C\|w\|$. $\qquad \square$

In the Proposition 3.5, a result in some "reciprocal" sense to that given in the previous statement, is presented. For this, the following result about the dual adjoint operator $\mathscr{T}^{\#}$ of $\mathscr{T}$ will be used ([17], Lemma 8.21).

Lemma 3.1. *Let $\mathscr{X}$ and $\mathscr{Y}$ be normed spaces, $\mathscr{T} \in \mathscr{L}(\mathscr{X},\mathscr{Y})$ and $x^* \in \mathscr{X}^*$. Then, $x^* \in Ran(\mathscr{T}^{\#})$ if and only if there exists $\tilde{C} > 0$ such that $|[x^*, x]| \leq \tilde{C}\|\mathscr{T}x\|, \forall x \in \mathscr{X}$.*

Proposition 3.5. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces, $\mathscr{T} \in \mathscr{L}(\mathscr{X},\mathscr{Y})$, $\psi : \mathscr{X} \to [0,+\infty)$, $\tilde{\phi} : \mathscr{Y} \times \mathscr{Y} \to [0,+\infty)$, $\tilde{x}$ the exact solution of (1.1). Suppose that there exists $\beta \geq 0$ and $\xi \in \partial\psi(\tilde{x})$ such that $[\xi, \tilde{x} - x] \leq \beta \tilde{\phi}(\mathscr{T}x, \mathscr{T}\tilde{x}), \forall x \in \mathscr{X}$, and that the following hypothesis is satisfied:*

$$\text{there exists } c > 0 \text{ such that } c\,\tilde{\phi}(y_1, y_2) \leq \|y_1 - y_2\|, \forall y_1, y_2 \in \mathscr{R}(\mathscr{T}). \qquad (3.7)$$

Then, there exists $w \in \mathscr{Y}^*$ such that $\xi = \mathscr{T}^{\#}w$.

Proof. By hypotheses, it follows that

$$[\xi, x] = [\xi, \tilde{x} - (\tilde{x} - x)] \leq \beta \tilde{\phi}(\mathscr{T}(\tilde{x} - x), \mathscr{T}\tilde{x}) \leq \beta \frac{1}{c} \|\mathscr{T}x\|,$$

for all $x \in \mathscr{X}$. Then, by Lemma 3.1 we have that $\xi \in Ran(\mathscr{T}^{\#})$, i.e., there exists $w \in \mathscr{Y}^*$ such that $\xi = \mathscr{T}^{\#}w$. $\qquad \square$

In the particular case in which $\tilde{\phi}(\mathscr{T}x, y) \doteq \|\mathscr{T}x - y\|$, the equivalence between the source condition and the variational inequality (3.6) is obtained immediately from the Propositions 3.4 and 3.5. Indeed, the hypotheses (3.5) and (3.7) are verified considering $c = C = 1$.

The following theorem presents an analogous result of the Theorem 3.3 with a modification in its hypotheses since error estimates are established using a variational inequality instead of a source condition. By the Proposition 3.4, it is immediate to see that this new hypothesis is weaker than those used in Theorem 3.3.

Theorem 3.4. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces, $\mathscr{T} \in \mathscr{L}(\mathscr{X},\mathscr{Y})$, $\psi : \mathscr{X} \to [0,+\infty)$, $\tilde{\phi} : \mathscr{Y} \times \mathscr{Y} \to [0,+\infty)$, $\tilde{x}$ the exact solution of (1.1) and $y^\delta \in \mathscr{Y}$ such that $\tilde{\phi}(y, y^\delta) \leq \delta$ and $\tilde{\phi}(y^\delta, y) \leq \delta$. Suppose that the functional $\tilde{\phi}$ satisfies the following hypotheses:*

1. *$\tilde{\phi}(y_1, y_3) \leq \tilde{\phi}(y_1, y_2) + \tilde{\phi}(y_2, y_3)$, for all $y_1, y_2, y_3 \in \mathscr{Y}$;*

2. *there exist $\beta \in [0,\infty)$ and $\xi \in \partial\psi(\tilde{x})$ such that $[\xi, \tilde{x} - x] \leq \beta \tilde{\phi}(\mathscr{T}x, \mathscr{T}\tilde{x})$, for all $x \in \mathscr{X}$.*

Then, for each minimizer x_η^δ of the functional $\mathscr{J}_{\tilde{\phi},\psi,\eta}^p$ given in (1.2) for data y^δ it follows that

- *if $p > 1$ and q is the conjugate of p, then*

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\frac{\delta^p}{p} + \eta\,\delta\beta + \frac{1}{q}(\eta\beta)^q}{\eta};$$

- *if $p = 1$ and $\eta\beta \leq 1$, then*

$$d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \delta\,\frac{1+\eta\beta}{\eta}.$$

Proof. Analogously to the proof of Theorem 3.3, since x_η^δ is a minimizer of the functional $\mathscr{J}_{\tilde{\phi},\psi,\eta}^p$ for data y^δ and $\tilde{\phi}(y, y^\delta) \leq \delta$, it immediately follows that

$$\frac{1}{p}\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p \leq \frac{\delta^p}{p} + \eta\left(\psi(\tilde{x}) - \psi(x_\eta^\delta)\right).$$

Then,

$$
\begin{aligned}
&\frac{1}{p}\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p + \eta\,d_\xi^\psi(x_\eta^\delta, \tilde{x}) \\
&\leq \frac{\delta^p}{p} + \eta\left(\psi(\tilde{x}) - \psi(x_\eta^\delta) + d_\xi^\psi(x_\eta^\delta, \tilde{x})\right) \\
&= \frac{\delta^p}{p} + \eta\,[\xi, \tilde{x} - x_\eta^\delta] \\
&\leq \frac{\delta^p}{p} + \eta\beta\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) + \eta\beta\,\tilde{\phi}(y^\delta, y) \qquad \text{by hypotheses 1, 2 and } \mathscr{T}\tilde{x} = y \\
&\leq \frac{\delta^p}{p} + \eta\beta\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) + \eta\beta\,\delta. \qquad\qquad\qquad \text{since } \tilde{\phi}(y^\delta, y) \leq \delta
\end{aligned}
$$

Finally,

$$\frac{1}{p}\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta)^p - \eta\beta\,\tilde{\phi}(\mathscr{T}x_\eta^\delta, y^\delta) + \eta\,d_\xi^\psi(x_\eta^\delta, \tilde{x}) \leq \frac{\delta^p}{p} + \eta\beta\,\delta,$$

and the proof follows analogously to that of Theorem 3.3 with β instead of $C\|w\|$ in (3.2). $\qquad\square$

From the proof of the Proposition 3.4 it follows that $\beta = C\|w\|$ so that the error estimates obtained in the Theorems 3.3 and 3.4 coincide.

It is appropriate to mention here that if we consider the fidelity term given by $\frac{1}{p}\|\mathscr{T}x - y\|^p$, that is $\tilde{\phi}(\mathscr{T}x, y) = \|\mathscr{T}x - y\|$, the hypothesis 1 of Theorem 3.4 is satisfied as a consequence of the triangular inequality of the norm. In this way, from the Theorem 3.4, it is possible to obtain a result that establishes an estimate of the error between the exact solution of (1.1) and a regularized solution obtained by minimizing TPG-p functionals (in this case, $\beta = \|w\|$ since $C = 1$).

3.2   Multiple penalizers

As previously mentioned, the simultaneous use of two or more penalizers of different nature will, in some way, allow the capturing of different characteristics on the exact solution. This is very usefull, for instance, in image restoration problems in which it is known "a-priori" that the original image is "blocky", i.e. it possesses both regions of high regularity and regions with sharp

discontinuities. Thus, generalizations of the Theorems 3.3 and 3.4 are presented for the case in which the regularized solution is obtained by minimizing the functional given in (1.4).

First, the following theorem presents an error estimate that constitutes a generalization of the Theorem 3.3.

Theorem 3.5. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces, $\mathscr{T} \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, $\tilde{\phi} : \mathscr{Y} \times \mathscr{Y} \to [0, +\infty)$, $\psi_i : \mathscr{X} \to [0, +\infty)$ convex functionals for all $i = 1, \ldots, n$, $\psi_{\vec{\eta}} : \mathscr{X} \to \mathbb{R}$ given by $\psi_{\vec{\eta}}(x) \doteq \vec{\eta} \cdot \vec{\psi}(x)$, where $\vec{\psi}(x) \doteq (\psi_1(x), \ldots, \psi_n(x))$, $\vec{\eta} = (\eta_1, \ldots, \eta_n)$, $\tilde{x}$ the exact solution of (1.1) and $y^\delta \in \mathscr{Y}$ such that $\tilde{\phi}(y, y^\delta) \leq \delta$ and $\tilde{\phi}(y^\delta, y) \leq \delta$. Suppose that, for each $\vec{\eta} \in \mathbb{R}^n$, there exists $w_{\vec{\eta}} \in \mathscr{Y}^*$ such that $\xi_{\vec{\eta}} \doteq \mathscr{T}^\# w_{\vec{\eta}} \in \partial \psi_{\vec{\eta}}(\tilde{x})$ and that the functional $\tilde{\phi}$ satisfies the following hypotheses:*

1. *$\tilde{\phi}(y_1, y_3) \leq \tilde{\phi}(y_1, y_2) + \tilde{\phi}(y_2, y_3)$, for all $y_1, y_2, y_3 \in \mathscr{Y}$;*

2. *there exists $C > 0$ such that $\|y_1 - y_2\| \leq C\tilde{\phi}(y_2, y_1)$, $\forall y_1, y_2 \in \mathscr{R}(\mathscr{T})$.*

Then, if $\vec{\eta}_* \doteq \frac{\vec{\eta}}{\|\vec{\eta}\|_1}$, for each minimizer $x_{\vec{\eta}}^\delta$ of the functional $\mathscr{I}_{\tilde{\phi}, \vec{\psi}, \vec{\eta}}^p$ given in (1.4) for data y^δ, it follows that

- *if $p > 1$ and q is the conjugate of p then*

$$d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^\delta, \tilde{x}) \leq \frac{\frac{1}{p}\delta^p + \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C\delta + \frac{1}{q}\left(\|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C\right)^q}{\|\vec{\eta}\|_1}.$$

- *if $p = 1$ and $\|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \leq 1$ then*

$$d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^\delta, \tilde{x}) \leq \delta \frac{1 + \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C}{\|\vec{\eta}\|_1}.$$

Proof. Since $x_{\vec{\eta}}^\delta$ is a minimizer of the functional $\mathscr{I}_{\tilde{\phi}, \vec{\psi}, \vec{\eta}}^p$ for data y^δ, and $\tilde{\phi}(y, y^\delta) \leq \delta$ and $\mathscr{T}\tilde{x} = y$, it follows that

$$\frac{1}{p}\tilde{\phi}(\mathscr{T}x_{\vec{\eta}}^\delta, y^\delta)^p + \vec{\eta} \cdot \vec{\psi}(x_{\vec{\eta}}^\delta) \leq \frac{1}{p}\tilde{\phi}(\mathscr{T}\tilde{x}, y^\delta)^p + \vec{\eta} \cdot \vec{\psi}(\tilde{x})$$

$$\leq \frac{1}{p}\delta^p + \vec{\eta} \cdot \vec{\psi}(\tilde{x}),$$

and, in consequence,

$$\frac{1}{p}\frac{\tilde{\phi}(\mathscr{T}x_{\vec{\eta}}^\delta, y^\delta)^p}{\|\vec{\eta}\|_1} + \vec{\eta}_* \cdot \vec{\psi}(x_{\vec{\eta}}^\delta) \leq \frac{1}{p}\frac{\delta^p}{\|\vec{\eta}\|_1} + \vec{\eta}_* \cdot \vec{\psi}(\tilde{x})$$

$$\frac{1}{p}\frac{\tilde{\phi}(\mathscr{T}x_{\vec{\eta}}^\delta, y^\delta)^p}{\|\vec{\eta}\|_1} + \psi_{\vec{\eta}_*}(x_{\vec{\eta}}^\delta) \leq \frac{1}{p}\frac{\delta^p}{\|\vec{\eta}\|_1} + \psi_{\vec{\eta}_*}(\tilde{x})$$

$$\frac{1}{p}\frac{\tilde{\phi}(\mathscr{T}x_{\vec{\eta}}^\delta, y^\delta)^p}{\|\vec{\eta}\|_1} \leq \frac{1}{p}\frac{\delta^p}{\|\vec{\eta}\|_1} + \psi_{\vec{\eta}_*}(\tilde{x}) - \psi_{\vec{\eta}_*}(x_{\vec{\eta}}^\delta). \tag{3.8}$$

Now, by hypothesis, there exists $w_{\vec{\eta}_*} \in \mathscr{Y}^*$ such that $\xi_{\vec{\eta}_*} \doteq \mathscr{T}^{\#} w_{\vec{\eta}_*} \in \partial \psi_{\vec{\eta}_*}(\tilde{x})$. Then, since $\psi_{\vec{\eta}_*}$ is a convex functional, from inequality (3.8) it follows that

$$
\begin{aligned}
\frac{1}{p} & \frac{\tilde{\phi}(\mathscr{T} x_{\vec{\eta}}^{\delta}, y^{\delta})^p}{\|\vec{\eta}\|_1} + d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^{\delta}, \tilde{x}) \\
& \leq \frac{1}{p} \frac{\delta^p}{\|\vec{\eta}\|_1} + \psi_{\vec{\eta}_*}(\tilde{x}) - \psi_{\vec{\eta}_*}(x_{\vec{\eta}}^{\delta}) + d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^{\delta}, \tilde{x}) \\
& = \frac{1}{p} \frac{\delta^p}{\|\vec{\eta}\|_1} + \left[\mathscr{T}^{\#} w_{\vec{\eta}_*}, \tilde{x} - x_{\vec{\eta}}^{\delta} \right] \\
& \leq \frac{1}{p} \frac{\delta^p}{\|\vec{\eta}\|_1} + \|w_{\vec{\eta}_*}\| C \tilde{\phi}\left(\mathscr{T} x_{\vec{\eta}}^{\delta}, y^{\delta} \right) + \|w_{\vec{\eta}_*}\| C \tilde{\phi}\left(y^{\delta}, \mathscr{T}\tilde{x} \right) \qquad \text{by hypotheses 1 and 2} \\
& \leq \frac{1}{p} \frac{\delta^p}{\|\vec{\eta}\|_1} + \|w_{\vec{\eta}_*}\| C \tilde{\phi}\left(\mathscr{T} x_{\vec{\eta}}^{\delta}, y^{\delta} \right) + \|w_{\vec{\eta}_*}\| C \delta. \qquad \text{since } \tilde{\phi}(y^{\delta}, y) \leq \delta \text{ and } \mathscr{T}\tilde{x} = y
\end{aligned}
$$

Then,

$$
\frac{1}{p} \tilde{\phi}(\mathscr{T} x_{\vec{\eta}}^{\delta}, y^{\delta})^p - \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \tilde{\phi}\left(\mathscr{T} x_{\vec{\eta}}^{\delta}, y^{\delta} \right) + \|\vec{\eta}\|_1 d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^{\delta}, \tilde{x}) \leq \frac{1}{p} \delta^p + \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \delta.
\tag{3.9}
$$

- If $p > 1$, it immediately follows from Young's inequality that

$$
\|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \tilde{\phi}\left(\mathscr{T} x_{\vec{\eta}}^{\delta}, y^{\delta} \right) \leq \frac{1}{p} \tilde{\phi}(\mathscr{T} x_{\vec{\eta}}^{\delta}, y^{\delta})^p + \frac{1}{q} \left(\|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \right)^q,
$$

which with (3.9) imply that

$$
-\frac{1}{q} \left(\|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \right)^q + \|\vec{\eta}\|_1 d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^{\delta}, \tilde{x}) \leq \frac{1}{p} \delta^p + \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \delta.
$$

Thus,

$$
d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^{\delta}, \tilde{x}) \leq \frac{\frac{1}{p} \delta^p + \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \delta + \frac{1}{q} \left(\|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \right)^q}{\|\vec{\eta}\|_1}.
$$

- If $p = 1$, by (3.9) it turns out that

$$
\left(1 - \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \right) \tilde{\phi}(\mathscr{T} x_{\vec{\eta}}^{\delta}, y^{\delta}) + \|\vec{\eta}\|_1 d_{\xi_{\vec{\eta}_*}}^{\psi}(x_{\vec{\eta}}^{\delta}, \tilde{x}) \leq \delta + \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \delta,
$$

and since $1 - \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C \geq 0$, finally, it follows that

$$
d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^{\delta}, \tilde{x}) \leq \delta \frac{1 + \|\vec{\eta}\|_1 \|w_{\vec{\eta}_*}\| C}{\|\eta\|_1}.
$$

$\square$

Note that if $\eta_i = 0$, for all $i \geq 2$, that is, there is only one penalty term, the error estimate of Theorem 3.5 agrees with that of the Theorem 3.3.

As was done in the case in which there is only one penalty term, a result is presented below that generalizes the Theorem 3.4 for the case in which two or more penalty terms are used and where the hypothesis related to a source condition present in the Theorem 3.5 is replaced by one that considers a variational inequality.

Theorem 3.6. *Let $\mathscr{X}$ and $\mathscr{Y}$ be Banach spaces, $\mathscr{T} \in \mathscr{L}(\mathscr{X}, \mathscr{Y})$, $\tilde{\phi} : \mathscr{Y} \times \mathscr{Y} \to [0, +\infty)$, $\psi_i : \mathscr{X} \to [0, +\infty)$ convex functionals for all $i = 1, \ldots, n$, $\psi_{\vec{\eta}} : \mathscr{X} \to \mathbb{R}$ given by $\psi_{\vec{\eta}}(x) \doteq \vec{\eta} \cdot \vec{\psi}(x)$, where $\vec{\psi}(x) \doteq (\psi_1(x), \ldots, \psi_n(x))$, $\vec{\eta} = (\eta_1, \ldots, \eta_n)$, $\tilde{x}$ the exact solution of (1.1) and $y^\delta \in \mathscr{Y}$ such that $\tilde{\phi}(y, y^\delta) \leq \delta$ and $\tilde{\phi}(y^\delta, y) \leq \delta$. Suppose that the functional $\tilde{\phi}$ satisfies the following hypotheses:*

1. *$\tilde{\phi}(y_1, y_3) \leq \tilde{\phi}(y_1, y_2) + \tilde{\phi}(y_2, y_3)$, for all $y_1, y_2, y_3 \in \mathscr{Y}$;*

2. *there exist $\beta \in [0, \infty)$ and $\xi \in \partial \psi(\tilde{x})$ such that $[\xi, \tilde{x} - x] \leq \beta \, \tilde{\phi}(\mathscr{T}x, \mathscr{T}\tilde{x})$, for all $x \in \mathscr{X}$.*

Then, if $\vec{\eta}_* \doteq \frac{\vec{\eta}}{\|\vec{\eta}\|_1}$, for each minimizer $x_{\vec{\eta}}^\delta$ of the functional $\mathscr{I}_{\tilde{\phi}, \vec{\psi}, \vec{\eta}}^p$ given in (1.4) for data y^δ, it follows that

- *if $p > 1$ and q is the conjugate of p then*

$$d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^\delta, \tilde{x}) \leq \frac{\frac{1}{p} \delta^p + \|\vec{\eta}\|_1 \beta \, \delta + \frac{1}{q}(\|\vec{\eta}\|_1 \beta)^q}{\|\vec{\eta}\|_1}.$$

- *if $p = 1$ and $\|\vec{\eta}\|_1 \beta \leq 1$ then*

$$d_{\xi_{\vec{\eta}_*}}^{\psi_{\vec{\eta}_*}}(x_{\vec{\eta}}^\delta, \tilde{x}) \leq \delta \, \frac{1 + \|\vec{\eta}\|_1 \beta}{\|\vec{\eta}\|_1}.$$

Proof. Analogously to the proof of Theorem 3.5, since $x_{\vec{\eta}}^\delta$ is a minimizer of the functional $\mathscr{I}_{\tilde{\phi}, \vec{\psi}, \vec{\eta}}^p$ for data y^δ and $\tilde{\phi}(y, y^\delta) \leq \delta$, it follows that

$$\frac{1}{p} \frac{\tilde{\phi}(\mathscr{T}x_{\vec{\eta}}^\delta, y^\delta)^p}{\|\vec{\eta}\|_1} \leq \frac{1}{p} \frac{\delta^p}{\|\vec{\eta}\|_1} + \psi_{\vec{\eta}_*}(\tilde{x}) - \psi_{\vec{\eta}_*}(x_{\vec{\eta}}^\delta).$$

Then,

$$\frac{1}{p}\frac{\tilde{\phi}(\mathscr{T}x_{\vec{\eta}}^{\delta},y^{\delta})^{p}}{\|\vec{\eta}\|_{1}}+d_{\xi_{\vec{\eta}*}}^{\psi_{\vec{\eta}*}}(x_{\vec{\eta}}^{\delta},\tilde{x})$$

$$\leq \frac{1}{p}\frac{\delta^{p}}{\|\vec{\eta}\|_{1}}+\psi_{\vec{\eta}*}(\tilde{x})-\psi_{\vec{\eta}*}(x_{\vec{\eta}}^{\delta})+d_{\xi_{\vec{\eta}*}}^{\psi_{\vec{\eta}*}}(x_{\vec{\eta}}^{\delta},\tilde{x})$$

$$= \frac{1}{p}\frac{\delta^{p}}{\|\vec{\eta}\|_{1}}+\left[\xi_{\vec{\eta}*},\tilde{x}-x_{\vec{\eta}}^{\delta}\right]$$

$$\leq \frac{1}{p}\frac{\delta^{p}}{\|\vec{\eta}\|_{1}}+\beta\,\tilde{\phi}\left(\mathscr{T}x_{\vec{\eta}}^{\delta},y^{\delta}\right)+\beta\,\tilde{\phi}\left(y^{\delta},\mathscr{T}\tilde{x}\right) \qquad \text{by hypotheses 1 and 2}$$

$$\leq \frac{1}{p}\frac{\delta^{p}}{\|\vec{\eta}\|_{1}}+\beta\,\tilde{\phi}\left(\mathscr{T}x_{\vec{\eta}}^{\delta},y^{\delta}\right)+\beta\,\delta \qquad \text{since } \tilde{\phi}(y^{\delta},y)\leq\delta \text{ and } \mathscr{T}\tilde{x}=y,$$

and thus,

$$\frac{1}{p}\tilde{\phi}(\mathscr{T}x_{\vec{\eta}}^{\delta},y^{\delta})^{p}-\|\vec{\eta}\|_{1}\beta\,\tilde{\phi}\left(\mathscr{T}x_{\vec{\eta}}^{\delta},y^{\delta}\right)+\|\vec{\eta}\|_{1}d_{\xi_{\vec{\eta}*}}^{\psi_{\vec{\eta}*}}(x_{\vec{\eta}}^{\delta},\tilde{x})\leq\frac{1}{p}\delta^{p}+\|\vec{\eta}\|_{1}\beta\,\delta,$$

and the proof follows analogously to that of Theorem 3.5 with β instead of $\|w_{\vec{\eta}*}\|C$ in (3.9). $\quad\square$

It is appropriate to mention here that for the case in which there is only one penalizing term, that is, if $\eta_{i}=0$, for all $i\geq 2$, the error estimate obtained in Theorem 3.6 is the same as that of Theorem 3.4.

4   CONCLUSIONS

In this work, error estimates were presented for the case in which the regularized solution is obtained by minimizing doubly-generalized Tikhonov-Phillips functionals. In particular, for the case of generalized Tikhonov-Phillips functionals, we have seen that the error estimates obtained coincide with those presented by K. Ito and B. Jin in [11]. The first result obtained was mainly based on an assumption given by a source condition. We have seen that it is possible to replace this assumption by a variational inequality obtaining analogous results of error estimates. Finally, relationships were established between the optimality condition associated with the problem, the source condition and the variational inequality. On the other hand, because it is known that, in certain cases, the use of two or more penalizing terms is useful, generalizations of the results of error estimates are presented for cases in which the regularized solution is obtained by minimizing doubly-generalized Tikhonov-Phillips functionals with multiple penalizers.

Acknowledgments

REFERENCES

[1] A.B. Bakushinskii. Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. *USSR Comp. Math. Math. Phys.*, **24** (1984), 181–182.

[2] M. Benning & M. Burger. Error Estimates for General Fidelities. *Electron. T. Numer. Ana.*, **38** (2011), 44–68.

[3] M. Burger & S. Osher. Convergence rates of convex variational regularization. *Inverse Problems*, **20**(5) (2004), 1411–1421.

[4] H. Engl, K. Kunisch & A. Neubauer. Convergence rates for Tikhonov regularization of non-linear ill-posed problems. *Inverse Problems*, **5**(4) (1989), 523–540.

[5] H.W. Engl, M. Hanke & A. Neubauer. "Regularization of inverse problems", volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht (1996).

[6] J. Flemming. Theory and examples of variational regularization with non-metric fitting functionals. *Journal of Inverse and Ill-posed Problems*, **18**(6) (2010), 677 – 699.

[7] J. Flemming & B. Hofmann. A new Approach to Source Conditions in Regularization with General Residual Term. *Numerical Functional Analysis and Optimization*, **31** (2010), 254–284.

[8] G.H. Golub, M.T. Heath & G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, **21** (1979), 215–223.

[9] P. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM*, **Review 34, no. 4** (1992), 561–580.

[10] B. Hofmann, B. Kaltenbacher, C. Pöschl & O. Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, **23**(3) (2007), 987.

[11] K. Ito & B. Jin. "Inverse problems: Tikhonov theory and algorithms", volume 22 of *Series on applied mathematics*. World Scientific (2015).

[12] K. Ito, B. Jin & T. Takeuchi. Multi-parameter Tikhonov regularization. *Methods and Applications of Analysis*, **18** (2011), 31–46.

[13] K. Ito, B. Jin & J. Zou. A new choice rule for regularization parameters in Tikhonov regularization. *Applicable Analysis*, **90** (2011), 1521–1544.

[14] Y. Lu, L. Shen & Y. Xu. Multi-Parameter Regularization Methods for High-Resolution Image Reconstruction With Displacement Errors. *IEEE Transactions on Circuits and Systems I: Regular Papers*, **54**(8) (2007), 1788–1799.

[15] G.L. Mazzieri, R.D. Spies & K.G. Temperini. Directional convergence of spectral regularization method associated to families of closed operators. *Computational and Applied Mathematics*, **32** (2013), 119–134.

[16] D.L. Phillips. A technique for the numerical solution of certain integral equations of the first kind. *J. Assoc. Comput. Mach.*, **9** (1962), 84–97.

[17] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier & F. Lenzen. "Variational Methods in Imaging", volume 167 of *Applied Mathematical Sciences*. Springer, New York (2009).

[18] T. Schuster, B. Kaltenbacher, B. Hofmann & K. Kazimierski. "Regularization Methods in Banach Spaces". de Gruyter, Berlin, New York (2012).

[19] A.N. Tikhonov. Regularization of incorrectly posed problems. *Soviet Math. Dokl.*, **4** (1963), 1624–1627.

[20] A.N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, **4** (1963), 1035–1038.

[21] Z. Wang. Multi-parameter Tikhonov regularization and model function approach to the damped Morozov principle for choosing regularization parameters. *Journal of Computational and Applied Mathematics*, **236** (2012), 1815–1832.

[22] E. Zeidler. "Nonlinear Functional Analysis and Its Applications, III: Variational methods and optimization". Springer-Verlag, New York (1985). URL `https://books.google.com.ar/books?id=sCNXnAEACAAJ`.