AN INEXACT MODIFIED QUASI-NEWTON METHOD FOR NONSMOOTH REGULARIZED OPTIMIZATION

NATHAN ALLAIRE*, SÉBASTIEN LE DIGABEL†, AND DOMINIQUE ORBAN‡

Abstract. We introduce method iR2N, a modified proximal quasi-Newton method for minimizing the sum of a $\mathcal{C}^1$ function f and a lower semi-continuous prox-bounded h that permits inexact evaluations of f, ∇f and of the relevant proximal operators. Both f and h may be nonconvex. In applications where the proximal operator of h is not known analytically but can be evaluated via an iterative procedure that can be stopped early, or where the accuracy on f and ∇f can be controlled, iR2N can save significant computational effort and time. At each iteration, iR2N computes a step by approximately minimizing the sum of a quadratic model of f, a model of h, and an adaptive quadratic regularization term that drives global convergence. In our implementation, the step is computed using a variant of the proximal-gradient method that also allows inexact evaluations of the smooth model, its gradient, and proximal operators. We assume that it is possible to interrupt the iterative process used to evaluate proximal operators when the norm of the current iterate is larger than a fraction of that of the minimum-norm optimal step, a weaker condition than others in the literature. Under standard assumptions on the accuracy of f and ∇f, we establish global convergence in the sense that a first-order stationarity measure converges to zero and a worst-case evaluation complexity in $O(\epsilon^{-2})$ to bring said measure below $\epsilon > 0$. Thus, inexact evaluations and proximal operators do not deteriorate asymptotic complexity compared to methods that use exact evaluations. We illustrate the performance of our implementation on problems with ℓ_p-norm, ℓ_p total-variation and the indicator of the nonconvex pseudo p-norm ball as regularizers. On each example, we show how to construct an effective stopping condition for the iterative method used to evaluate the proximal operator that ensures satisfaction of our inexactness assumption. Our results show that iR2N offers great flexibility when exact evaluations are costly or unavailable, and highlight how controlled inexactness can reduce computational effort effectively and significantly.

Key words. Nonsmooth optimization, nonconvex optimization, modified quasi-Newton method, proximal quasi-Newton method, regularized optimization, composite optimization, proximal gradient method, inexact proximal operator, inexact evaluations

AMS subject classifications. 65F22, 90C30, 90C53

1. Introduction.

We consider the problem class

$$\underset{x\in\mathbb{R}^n}{\text{minimize}}\; f(x) + h(x), \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is proper, lower semi-continuous (lsc), and both may be nonconvex. In practice, h, called the *regularizer*, is designed to promote desirable properties in solutions, such as sparsity. We develop method iR2N, a variant of the modified proximal quasi-Newton algorithm R2N of Diouane et al. [25] that allows for inexact evaluations of f and ∇f, as well as of the relevant proximal operators. Among other applications, evaluations of f and ∇f are inexact when they result from the discretization of a differential or integral operator [8], from the sampling of a sum of a large number of terms, as in machine learning applications [34], or from using multiple floating-point systems [31]. Like

*GERAD and Department of Mathematics and Industrial Engineering, Polytechnique Montréal. E-mail: nathan.allaire@etud.polymtl.ca.

†GERAD and Department of Mathematics and Industrial Engineering, Polytechnique Montréal. E-mail: sebastien.le.digabel@gerad.ca. Research partially supported by NSERC Discovery Grant RGPIN–2024–05086.

‡GERAD and Department of Mathematics and Industrial Engineering, Polytechnique Montréal. E-mail: dominique.orban@gerad.ca. Research partially supported by NSERC Discovery Grant RGPIN–2020–06535.

R2N, iR2N computes a step at each iteration by approximately minimizing the sum of a quadratic model of f, a model of h, and an adaptive quadratic regularization term. The subproblem is solved with method iR2, which is to method R2 of Aravkin et al. [3] as iR2N is to R2N, i.e., proximal operators are evaluated inexactly. Method R2 may be viewed as a variant of the standard proximal-gradient method with adaptive step length, and is a special case of R2N. We consider settings where proximal operators do not have a closed-form expression, and one must thus rely on inexact evaluations. Specifically, we focus on scenarios where proximal operators can be evaluated by running a convergent algorithm that can be terminated early with appropriate guarantees detailed below. Special cases that fit our assumptions include choices of convex and nonconvex h, including the ℓ_p-norm total variation (TV), ℓ_p-norm regularizer and the indicator of the nonconvex ℓ_p-pseudo norm ball with $0 < p < 1$. Method iR2N reduces to R2N when f, ∇f and proximal operators are evaluated exactly. We establish global convergence of iR2N under standard assumptions on the inexactness of f and ∇f, and provided the inexact proximal operator yields a step whose norm is at least a fraction of the norm of an optimal step. We also establish that worst-case evaluation complexity of iR2N is of the same order as that of R2N. Thus, inexact evaluations do not degrade worst-case complexity. Our remaining assumptions are standard. To emphasize our assumptions on inexact evaluations, we simplify those assumptions of [25] that would complicate the analysis. In particular, we assume that ∇f is Lipschitz continuous, but its Lipschitz constant need not be known nor approximated. However, it should be clear that iR2N remains convergent under the more general assumptions of [25] with its worst-case complexity affected accordingly. It should also be clear that minor alterations of our approach would establish that the proximal quasi-Newton trust-region algorithm of Aravkin et al. [4, 5] remains convergent under inexact evaluations and its asymptotic worst-case complexity is unchanged. Such minor alterations would also establish convergence and complexity of Levenberg-Marquardt variants in the vein of [6] that are useful when f is a least-squares residual.

We report computational experience with the proximal operator of the ℓ_p norm, the total variation in ℓ_p norm, and the indicator of the nonconvex ℓ_p-pseudo norm ball. Each of those proximal operators must be evaluated via an iterative procedure. For each, we devise a stopping condition that ensures satisfaction of our assumption on inexact proximal operator evaluations. Our results show that iR2N offers great flexibility in settings where exact evaluations are costly or unavailable, and highlight how controlled inexactness can be exploited to reduce computational effort effectively and significantly. We provide an efficient Julia implementation of iR2N as part of the open-source package `RegularizedOptimization.jl` [7].

Related Research. Most numerical methods for (1.1) require the evaluation of one or more proximal operators [32] at each iteration. The proximal operator of h with step size $\nu > 0$ at $q \in \mathbb{R}^n$ is

$$\operatorname*{prox}_{\nu h}(q) := \operatorname*{argmin}_{u \in \mathbb{R}^n} \tfrac{1}{2}\|u - q\|^2 + \nu h(u) \subseteq \mathbb{R}^n. \tag{1.2}$$

For given h and q, (1.2) can be empty, a singleton or contain multiple elements, one of which must be identified. Beck [11] and Chierchia et al. [19] summarize the closed-form of (1.2) for a large number of choices of h relevant in applications. The standard proximal-gradient method [26] along with most proximal methods in the literature assume that obtaining an element of (1.2) *exactly* is possible.

For certain choices of h, it is necessary to apply an iterative method to appoximate an element of (1.2), e.g., the total variation (TV) with ℓ_p regularization $h(x) = \|Dx\|_p$,

where $p \geq 1$ and D is the upper bidiagonal finite-difference operator with a diagonal of negative ones and a superdiagonal of ones. Finding an element in (1.2) for the TV-ℓ_p can be achieved via the taut-string method [9] or the fast TV denoising method [20]. As in other methods in the literature for various choices of convex h [10, 24, 27], the latter monitor the duality gap between a convex problem and its dual. Those algorithms have guaranteed convergence properties and can be terminated early, i.e., short of optimality. In the above, the evaluation of (1.2) is inexact in the sense that a convergent process to identify a global minimizer is applied and can be stopped short of optimality according to an optimality criterion.

A somewhat more complicated scenario is the algorithm described by Yang et al. [43] for the case where h is the indicator of the "ball" in pseudo-norm ℓ_p with $p \in (0, 1)$. The evaluation of the proximal operator requires solving a nonconvex problem to global optimality in that case, and their algorithm is not guaranteed to always succeed. We return to this problem in Section 4.

Other concepts of inexactness of the proximal operator appear in the literature. For convex h, Rockafellar [36] requires that an approximate solution of (1.2) be a certain distance from the optimal set. Still for convex h, Barré et al. [10] unveil multiple ways to define inexactness by finding a primal-dual point in a certain relaxed subdifferential. Salzo and Villa [38] define three approximations: they compute z such that either *(i)* $\|z - \operatorname{prox}_{\nu h}(q)\| \leq \epsilon$, *(ii)* $\nu^{-1}(q - z)$ lies in a relaxation of the subdifferential of h at z, or *(iii)* $z \in \operatorname{prox}_{\nu h}(q + e)$ with $\|e\| \leq \epsilon$ for some $\epsilon > 0$. Chen et al. [18] extend proximal inexactness by introducing the concept of $(\gamma, \delta, \varepsilon)$-proximal-gradient stationary point (PGSP) for convex h based on the Goldstein subdifferential. The PGSP generalizes the three concepts of [38] by jointly relaxing spatial and functional exactness and directly quantifying the first-order residual, thus also encompassing Rockafellar's [36] and relaxed subgradient formulations within a unified framework. For nonconvex h, Gu et al. [27] say that an element is an inexact solution of (1.2) if its objective value is within ϵ of its optimal value.

To cope with inexact evaluations of the proximal operator, classical schemes must be revised to preserve convergence guarantees. The seminal inexact proximal-point algorithm (iPPA) of Rockafellar [36] allows summably controlled errors in the resolvant computation of a maximal monotone operator and still ensures global convergence with linear/superlinear behavior under suitable parameter growth. Building on the accelerated estimate-sequence framework, Salzo and Villa [38] establish that the accelerated iPPA retains $O(1/k)$ decay under inexactness of type (i) above, and optimal $O(1/k^2)$ decay under inexactness of type (ii). Schmidt et al. [39] establish an $O(1/k)$ rate for proximal-gradient and an $O(1/k^2)$ rate for an accelerated variant under inexactness similar to (iii) above. Extensions include inertial, variable-metric forward–backward schemes with relative inner accuracy and uniform symmetric positive definite metrics [16]; nonconvex inexact (accelerate) proximal gradient with guarantees matching the exact counterparts under calibrated error schedules [27]; adaptive, implementable stopping rules that preserve $O(\epsilon^{-2})$ iteration complexity and enable support identification [24]; and accelerated proximal gradient under relative error criteria that maintain an $O(1/k^2)$ rate [13]. For nonconvex problems, the sequence generated by an inexact proximal-gradient (or splitting) method can still be shown to converge to a first-order critical point under an assumption of type (iii) above on the approximation errors [41]. Finally, for weakly convex functions, recent results establish global convergence for inexact proximal algorithms under inexactness of type (i) above, allowing controlled inexactness in the proximal steps while maintaining convergence [28].

Notation. The Euclidean norm is $\|\cdot\|$. When required, other norms are denoted with different symbols. We use f, h, m, ϕ, φ, ξ and ψ for functions. Other lowercase Latin letters denote vectors in $\mathbb{R}^n$. Exceptions are p and q, which are standard to denote a pair of dual ℓ_p and ℓ_q norms, and r, which denotes a radius. Uppercase A and B are matrices, L is a Lipschitz constant, and O is used for the Landau notation. Lowercase Greek letters denote scalars. Calligraphic letters denote sets.

2. Background.

2.1. Variational Analysis Concepts. We say that $h : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is proper if $h(x) < +\infty$ for at least one $x \in \mathbb{R}^n$ and lower semi-continuous (lsc) at $\bar{x}$ if $\liminf_{x\to\bar{x}} h(x) = h(\bar{x})$. It is lsc if it is lsc at all $\bar{x} \in \mathbb{R}^n$. We say that h is prox-bounded at x if there is $\lambda > 0$ such that $w \mapsto h(w) + \frac{1}{2}\lambda^{-1}\|w - x\|^2$ is bounded below [37, Definition 1.23]. The threshold of prox-boundedness of h at x is the supremum of all such λ at x, and is denoted λ_x. We say that h is *uniformly prox-bounded* if there is $\lambda \in \mathbb{R}_+ \cup \{+\infty\}$ such that $\lambda_x \geq \lambda$ for all $x \in \mathbb{R}^n$.

For $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ and $\bar{x} \in \operatorname{dom}(\phi)$, the Fréchet subdifferential of ϕ at $\bar{x}$ is

$$\widehat{\partial}\phi(\bar{x}) := \left\{ v \in \mathbb{R}^n \;\middle|\; \liminf_{x\to\bar{x}} \frac{\phi(x) - \phi(\bar{x}) - v^T(x - \bar{x})}{\|x - \bar{x}\|} \geq 0 \right\}.$$

The limiting subdifferential $\partial\phi(\bar{x})$ of ϕ at $\bar{x}$ is the set of elements $v \in \mathbb{R}^n$ such that there exists a sequence $\{x_k\} \to \bar{x}$ with $\{\phi(x_k)\} \to \phi(\bar{x})$, and there exists $v_k \in \widehat{\partial}\phi(x_k)$ for all k such that $\{v_k\} \to v$. It always holds that $\widehat{\partial}\phi(\bar{x}) \subseteq \partial\phi(\bar{x})$.

If ϕ is proper, we say that $\bar{x}$ is stationary for ϕ, or for the problem of minimizing ϕ, if $0 \in \widehat{\partial}\phi(\bar{x})$. If ϕ is proper and has a local minimum at $\bar{x}$, then $\bar{x}$ is stationary for ϕ. In the special case where $\phi = f + h$ with f continuously differentiable and h proper, then $\partial\phi(x) = \nabla f(x) + \partial h(x)$ [37, Theorem 10.1]. We say that $f : \mathbb{R}^n \to \mathbb{R}$ has Lipschitz-continuous gradient with Lipschitz constant $L \geq 0$ if for all x and $s \in \mathbb{R}^n$,

$$|f(x + s) - f(x) - \nabla f(x)^T s| \leq \tfrac{1}{2}L\|s\|^2. \tag{2.1}$$

2.2. Models. In this work, we focus on three sources of inexactness: the objective, its gradient and the proximal operator evaluations. We denote $\widehat{f}$ and $\widehat{\nabla} f$ the inexact counterparts of f and ∇f. At each iteration, R2N computes a step s_{cp} defined below that serves to define a stationarity measure and that results from a proximal operator evaluation. Accordingly, in iR2N, we denote its inexact counterpart $\widehat{s}_{\text{cp}}$. We follow [3, 6, 25] and structure the iterations of an algorithm around two sets of models, but, since the only information we have access to is inexact, those are based on $\widehat{f}$ and $\widehat{\nabla} f$.

For $\nu > 0$ and $x \in \mathbb{R}^n$, the first-order models

$$\varphi_{\text{cp}}(s; x) := \widehat{f}(x) + \widehat{\nabla} f(x)^T s \tag{2.2}$$

$$\psi(s; x) \;\approx h(x + s) \tag{2.3}$$

$$m_{\text{cp}}(s; x, \nu^{-1}) := \varphi_{\text{cp}}(s; x) + \tfrac{1}{2}\nu^{-1}\|s\|^2 + \psi(s; x) \tag{2.4}$$

serve to generalize the concept of Cauchy point, hence the subscript "cp", where we use the symbol "$\approx$" to mean that the left-hand side is an approximation of the right-hand side. We will be more specific in Assumption 3.3 below. The dual role of models (2.2)–(2.4) is to define a threshold for sufficient decrease at each iteration, and to define a measure of approximate stationarity.

For $\sigma > 0$, $x \in \mathbb{R}^n$ and $B(x) = B(x)^T \in \mathbb{R}^{n\times n}$, the second-order models

$$\varphi(s;x) \coloneqq \widehat{f}(x) + \widehat{\nabla} f(x)^T s + \tfrac{1}{2} s^T B(x) s \tag{2.5}$$

$$m(s;x,\sigma) \coloneqq \varphi(s;x) + \tfrac{1}{2}\sigma\|s\|^2 + \psi(s;x), \tag{2.6}$$

are used to compute a step. Because $\varphi_{\text{cp}}(\cdot;x)$ is linear and $\varphi(\cdot;x)$ is quadratic for fixed x, both have globally Lipschitz-continuous gradient.

We follow [3, 6, 25] and require that all models that we consider satisfy the following assumption.

ASSUMPTION 2.1. *For all $x \in \mathbb{R}^n$, $\psi(\cdot;x)$ is proper, lsc and uniformly prox-bounded. In addition, $\psi(0;x) = h(x)$ and $\partial\psi(0;x) \subseteq \partial h(x)$.*

2.3. The Proximal-Gradient Method. The direct generalization of the gradient method to nonsmooth regularized optimization is the proximal-gradient method [26]. For (1.1), the proximal-gradient iteration can be written

$$\begin{aligned} x_{k+1} &= x_k + s_{k,\text{cp}} \\ s_{k,\text{cp}} &\in \operatorname*{argmin}_s \tfrac{1}{2}\nu_k^{-1}\|s + \nu_k\widehat{\nabla} f(x_k)\|^2 + \psi(s;x_k) \end{aligned} \tag{2.7}$$

$$\begin{aligned} &= \operatorname*{argmin}_s \widehat{\nabla} f(x_k)^T s + \tfrac{1}{2}\nu_k^{-1}\|s\|^2 + \psi(s;x_k) \\ &= \operatorname*{argmin}_s m_{\text{cp}}(s;x_k,\nu_k^{-1}), \end{aligned} \tag{2.8}$$

where $\nu_k > 0$ is an appropriate step length, though it is typically used with $\psi(s;x_k) \coloneqq h(x_k + s)$. We call $s_{k,\text{cp}}$ a Cauchy step. It turns out that $s_{k,\text{cp}}$ exists provided ν_k is sufficiently small.

PROPOSITION 2.1 (37, Theorem 1.25). *Let $\varphi_{\text{cp}}(\cdot;x)$ be as in (2.2), and $\psi(\cdot;x)$ be proper, lsc, prox-bounded with threshold $\lambda_x > 0$ and such that $\psi(0;x) = h(x)$. For any $0 < \nu < \lambda_x$, the set $\operatorname{argmin}_s m_{\text{cp}}(s;x,\nu^{-1})$ is nonempty and compact.*

We denote s_{cp} an element of $\operatorname{argmin}_s m_{\text{cp}}(s;x,\nu^{-1})$ when one exists. When s_{cp} is well defined, the quantity

$$\begin{aligned} \xi_{\text{cp}}(s_{\text{cp}},x,\nu^{-1}) &\coloneqq (\varphi_{\text{cp}} + \psi)(0;x) - (\varphi_{\text{cp}} + \psi)(s_{\text{cp}};x) \\ &= (\widehat{f} + h)(x) - (\varphi_{\text{cp}} + \psi)(s_{\text{cp}};x) \end{aligned} \tag{2.9}$$

is central to the algorithm and the analysis, as it is in [3, 6, 25], where it plays the dual role of defining Cauchy decrease and serving as stationarity measure. Indeed, under standard assumptions, x is stationary for (1.1) if $\xi_{\text{cp}}(s_{\text{cp}};x,\nu^{-1}) = 0$ [25, Lemma 3.5]. We diverge slightly from those references and, for reasons that become clear later, note that $\nu^{-1}\|s_{\text{cp}}\|$ can equally be used as stationarity measure.

PROPOSITION 2.2. *Let $x \in \mathbb{R}^n$ and $\psi(\cdot;x)$ be proper, lsc, prox-bounded with threshold $\lambda_x > 0$ and such that $\partial\psi(0;x) \subseteq \partial h(x)$. Let $0 < \nu < \lambda_x$ and $s_{\text{cp}} \in \operatorname{argmin}_s m_{\text{cp}}(s;x,\nu^{-1})$. If $s_{\text{cp}} = 0$, then $0 \in \widehat{\nabla} f(x) + \partial h(x)$. If, in addition, $\widehat{\nabla} f(x) = \nabla f(x)$, then x is stationary for (1.1).*

Proof. If $s_{\text{cp}} = 0$, then $\xi_{\text{cp}}(s_{\text{cp}};x,\nu^{-1}) = 0$ by (2.9). The rest of the proof is identical to that of [25, Lemma 3.5]. ∎

In the special case $h = 0$, i.e., smooth optimization, $s_{\text{cp}} = -\nu\nabla f(x)$. Thus, we normalize and use $\nu^{-1}\|s_{\text{cp}}\|$ as stationarity measure.

The identification of an s_{cp}, when one exists, coincides with the identification of an element in the image of a proximal operator (1.2), i.e., $s_{\text{cp}} \in \text{prox}_{\nu\psi(\cdot;x)}(-\nu\widehat{\nabla} f(x))$. It is the computation of an element in such a set that represents the main computational challenge in problems for which the set is not known analytically, so that one must resort to an iterative numerical method. In that case, the s_{cp} computed is inexact, and we refer to this situation as an inexact evaluation of the proximal operator.

The following result is hidden inside the proof of [15, Lemma 2].

PROPOSITION 2.3. *Let f have Lipschitz-continuous gradient with Lipschitz constant $L \geq 0$ and let h be proper, lsc and prox-bounded at $x \in \mathbb{R}^n$ with threshold $\lambda_x > 0$. Let $0 < \nu < \min(1/L, \lambda_x)$, and let $s \in \mathbb{R}^n$ be such that*

$$f(x) + \nabla f(x)^T s + \tfrac{1}{2}\nu^{-1}\|s\|^2 + h(x+s) \leq (f+h)(x). \tag{2.10}$$

Then,

$$(f+h)(x) - (f+h)(x+s) \geq \tfrac{1}{2}(\nu^{-1} - L)\|s\|^2. \tag{2.11}$$

Proof. We inject $f(x) + \nabla f(x)^T s \geq f(x+s) - \frac{1}{2}L\|s\|^2$, which follows from (2.1), into (2.10) and obtain (2.11). ∎

Proposition 2.3 applied to $\varphi_{\text{cp}}(\cdot;x)$, $\psi(\cdot;x)$ and $s_{\text{cp}} \in \text{prox}_{\nu\psi(\cdot;x)}(-\nu\widehat{\nabla} f(x))$, yields

$$\xi_{\text{cp}}(s_{\text{cp}}; x, \nu^{-1}) \geq \tfrac{1}{2}\nu^{-1}\|s_{\text{cp}}\|^2, \tag{2.12}$$

because the Lipschitz constant of $\nabla\varphi_{\text{cp}}(\cdot;x)$ is zero.

By contrast, we denote an approximate Cauchy step resulting from an *inexact* minimization of (2.4) as $\widehat{s}_{\text{cp}}$. We will be more specific about the meaning of inexactness in that context in Assumption 3.5. Accordingly, we define

$$\widehat{\xi}_{\text{cp}}(\widehat{s}_{\text{cp}}; x, \nu^{-1}) \coloneqq (\varphi_{\text{cp}} + \psi)(0;x) - (\varphi_{\text{cp}} + \psi)(\widehat{s}_{\text{cp}}; x). \tag{2.13}$$

Proposition 2.3 states that (2.11) also holds for any s that produces simple decrease in (2.4); s need not be an exact minimizer. Thus, if we apply a descent procedure to minimize (2.4) starting from $s = 0$, any iterate, denoted generically as $\widehat{s}_{\text{cp}}$, generated by that procedure will satisfy (2.11), i.e.,

$$(\varphi_{\text{cp}} + \psi)(0;x) - (\varphi_{\text{cp}} + \psi)(\widehat{s}_{\text{cp}}; x) \geq \tfrac{1}{2}\nu^{-1}\|\widehat{s}_{\text{cp}}\|^2. \tag{2.14}$$

Thus, an exact minimizer in (2.8) would produce a Cauchy step $s_{k,\text{cp}}$ that satisfies (2.12). For brevity, we write $\xi_{k,\text{cp}} \coloneqq \xi_{\text{cp}}(s_{k,\text{cp}}; x_k, \nu_k^{-1})$ and $\widehat{\xi}_{k,\text{cp}}$ instead of $\widehat{\xi}_{\text{cp}}(\widehat{s}_{k,\text{cp}}; x_k, \nu_k^{-1})$. The above shows that $\xi_{k,\text{cp}} \geq \frac{1}{2}\nu_k^{-1}\|s_{k,\text{cp}}\|^2$ and $\widehat{\xi}_{k,\text{cp}} \geq \frac{1}{2}\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\|^2$ provided $\widehat{s}_{k,\text{cp}}$ results in simple decrease in (2.4) from $s = 0$.

Proposition 2.2 indicates that one role of the first-order models (2.2)–(2.4), and hence of $\widehat{s}_{k,\text{cp}}$ and $\widehat{\xi}_{k,\text{cp}}$ is to determine approximate stationarity. The role of the second-order models (2.5)–(2.6) is to allow us to compute a step that improves upon the (inexact) Cauchy step. Minimizing the second-order model is a well-defined problem for all sufficiently large σ_k.

PROPOSITION 2.4 (25, Proposition 3.3). *Let $\varphi(\cdot;x)$ be defined as in (2.5), and let $\psi(\cdot;x)$ be proper, lsc and prox-bounded with threshold $\lambda_x > 0$ and such that $\psi(0;x) = h(x)$. For any $\sigma > \lambda_x^{-1} - \lambda_{\min}(B(x))$, the set $\operatorname{argmin}_s m(s;x,\sigma)$ is nonempty and compact, where $\lambda_{\min}$ represents the smallest eigenvalue.*

3. Algorithm and Convergence Analysis. Our algorithm is a modification of method R2N of Diouane et al. [25]. At a general iteration k, an approximate Cauchy step $\widehat{s}_{k,\text{cp}}$ is computed together with the corresponding value of $\widehat{\xi}_{k,\text{cp}}$ by minimizing (2.4) inexactly. If x_k is not approximately stationary, a step s_k is computed by approximately minimizing (2.6). Because only $\widehat{f}$, and not f, is available, we compute the ratio of achieved versus predicted decrease

$$\widehat{\rho}_k := \frac{\widehat{f}(x_k) + h(x_k) - (\widehat{f}(x_k + s_k) + h(x_k + s_k))}{\varphi(0; x_k) + \psi(0; x_k) - (\varphi(s_k; x_k) + \psi(s_k; x_k))} \tag{3.1}$$

to accept or reject s_k. Acceptance of s_k occurs when $\widehat{\rho}_k \geq \widehat{\eta}_1 > 0$, which indicates that sufficient decrease occurs in $\widehat{f} + h$. The parameters of the algorithm, specifically $\sigma_{\min}$, together with assumptions on the accuracy of $\widehat{f}$, are chosen so that acceptance of s_k also implies that sufficient decrease occurs in $f + h$. We then update σ_k accordingly, as in R2N. All that is required of s_k is that it satisfy a sufficient decrease condition—see Assumption 3.4 below. That can be achieved, for instance, by computing $\widehat{s}_{k,\text{cp}}$ from a single (inexact) proximal-gradient iteration on (2.6) with a well-chosen step length ν_k starting from $s = 0$, and computing s_k by continuing the (inexact) proximal-gradient iterations from $\widehat{s}_{k,\text{cp}}$. Should $\|s_k\|$ be much larger than $\|\widehat{s}_{k,\text{cp}}\|$, we reset s_k to $\widehat{s}_{k,\text{cp}}$ as in R2N. The procedure is formally stated as Algorithm 3.1. We refer the reader to [25] for more background.

3.1. Assumptions. Intentionally, our assumptions are not the most general under which convergence of Algorithm 3.1 can be shown to occur. We have done so in order to highlight the influence of our assumptions on the inexactness of the objective, gradient and proximal operators evaluations on the analysis. We refer the interested reader to [25] for the current most general assumptions. Nonetheless, we expect that our convergence guarantees remain valid under the weaker assumptions, at the cost of a more intricate analysis.

Our first assumption concerns Lipschitz-continuity of the gradient. Technically, this assumption is only necessary for the complexity analysis; convergence can be guaranteed under continuous differentiability only.

ASSUMPTION 3.1. *∇f is Lipschitz-continuous with constant $L \geq 0$—see (2.1).*

We assume that $\{B_k\}$ is bounded; a common assumption in the literature. Under appropriate growth conditions, convergence is preserved even if $\{B_k\}$ is allowed to grow unbounded [25].

ASSUMPTION 3.2. *There exists $\kappa_B > 0$ such that $\|B_k\| \leq \kappa_B$ for all k.*

Assumption 3.2 is trivially satisfied when $B_k = 0$, as in [3, Algorithm 6.1]. It is also satisfied in [12] where the objective is strongly convex and the model Hessian is defined by a positive definite limited-memory quasi-Newton update. Under standard assumptions, the LBFGS and LSR1 updates satisfy Assumption 3.2 [5, 17].

Our next assumption bounds the discrepancy between h and its model ψ.

ASSUMPTION 3.3. *There exists $\kappa_h > 0$ such that $|\psi(x, s) - h(x + s)| \leq \kappa_h \|s\|^2$ for all x and $s \in \mathbb{R}^n$.*

The bound $\|s\|^2$ in Assumption 3.3 can be relaxed to $o(\|s\|)$ [25]. Assumption 3.3 is satisfied when $\psi(s; x) = h(x + s)$, and when $h(x) = g(c(x))$ where c is twice continuously differentiable with bounded second derivatives and g is globally Lipschitz continuous if we select $\psi(s; x) = g(c(x) + \nabla c(x)^T s)$.

Algorithm 3.1 iR2N

Given $\kappa_f > 0$, $\kappa_\nabla > 0$, choose constants $0 < \gamma_3 \leq 1 < \gamma_1 \leq \gamma_2$, $0 < \widehat{\eta}_1 \leq \widehat{\eta}_2 < 1$.
Choose $0 < \theta_1 < 1 < \theta_2$.
Choose $\sigma_{\min} > 4\kappa_f\theta_1\theta_2^2/(\widehat{\eta}_1(1-\theta_1))$ and $\sigma_0 \geq \sigma_{\min}$.
for $k = 0, 1, \ldots$ **do**
    Choose $B_k := B(x_k) \in \mathbb{R}^{n\times n}$ such that $B_k = B_k^T$.
    Set $\nu_k := \theta_1/(\|B_k\| + \sigma_k)$.
    repeat
        Compute $\widehat{s}_{k,\mathrm{cp}}$ an approximate solution of $\min_s m_{\mathrm{cp}}(s; x_k, \nu_k^{-1})$ and $\widehat{\xi}_{k,\mathrm{cp}}$.
        Compute a step s_k such that $m(s_k; x_k, \sigma_k) \leq m(\widehat{s}_{k,\mathrm{cp}}; x_k, \sigma_k)$.
        if $\|s_k\| > \theta_2\|\widehat{s}_{k,\mathrm{cp}}\|$ **then**
            Reset $s_k = \widehat{s}_{k,\mathrm{cp}}$.
        end if
    until $\widehat{f}$ and $\widehat{\nabla} f$ satisfy Assumption 3.6.
    Compute the ratio $\widehat{\rho}_k$ as in (3.1).
    if $\widehat{\rho}_k \geq \widehat{\eta}_1$ **then**
        Set $x_{k+1} = x_k + s_k$.
    else
        Set $x_{k+1} = x_k$.
    end if
    Update the regularization parameter according to

$$\sigma_{k+1} \in \begin{cases} [\gamma_3\sigma_k,\ \sigma_k] & \text{if } \widehat{\rho}_k \geq \widehat{\eta}_2, \quad \text{very successful iteration} \\ [\sigma_k,\ \gamma_1\sigma_k] & \text{if } \widehat{\eta}_1 \leq \widehat{\rho}_k < \widehat{\eta}_2, \quad \text{successful iteration} \\ [\gamma_1\sigma_k,\ \gamma_2\sigma_k] & \text{if } \widehat{\rho}_k < \widehat{\eta}_1 \quad \text{unsuccessful iteration} \end{cases}$$

    Reset $\sigma_{k+1} = \max(\sigma_{k+1},\ \sigma_{\min})$.
end for

The next assumption drives the convergence analysis and states that the step s_k computed at iteration k should result in a decrease at least comparable to that induced by the approximate Cauchy step in the first-order model.

ASSUMPTION 3.4. *There is $\theta_1 \in (0,1)$ such that $\varphi(0;x) + \psi(0;x) - (\varphi(s_k;x) + \psi(s_k;x)) \geq (1-\theta_1)\widehat{\xi}_{k,\mathrm{cp}}$ for all k.*

As we now show, Assumption 3.4 holds for s_k computed as stated in Algorithm 3.1.

LEMMA 3.1. *For $\theta_1 \in (0,1)$ and s_k as in Algorithm 3.1, Assumption 3.4 holds.*

Proof. The proof of [25, Proposition 3] applies with $s = s_k$ and $\widehat{s}_{k,\mathrm{cp}}$ in place of s_{cp}. Indeed, it remains valid for any $s \in \mathbb{R}^n$ and $s_{\mathrm{cp}} \in \mathbb{R}^n$ as long as $m(s;x,\sigma) \leq m(s_{\mathrm{cp}};x,\sigma)$, which is guaranteed by step 7 of Algorithm 3.1. ∎

We ensure that Step 7 in Algorithm 3.1 holds because the inexact Cauchy step $\widehat{s}_{k,\mathrm{cp}}$ coincides with the first (inexact) step of the proximal gradient method applied to $m(s;x_k,\sigma_k)$ from $s = 0$ with an appropriate step length ν_k. Therefore, computing s_k by continuing the proximal iterations from $\widehat{s}_{k,\mathrm{cp}}$ leads to further decrease in $m(s;x_k,\sigma_k)$.

The next assumption requires the norm of the computed step $\widehat{s}_{k,\mathrm{cp}}$ to be at least a fraction of that of an exact step $s_{k,\mathrm{cp}}$.

ASSUMPTION 3.5. *There exists* $\kappa_s \in (0,1]$ *such that, for all* k,

$$\|\widehat{s}_{k,\text{cp}}\| \geq \kappa_s \min\{\|s_{k,\text{cp}}\| \mid s_{k,\text{cp}} \in \underset{\nu\psi(\cdot;x_k)}{\text{prox}}\ (-\nu_k \widehat{\nabla} f(x_k))\}.$$

In the experiments of Section 4, $\psi(\cdot;x_k)$ satisfies the assumptions of Proposition 2.1 and, therefore, the minimum in Assumption 3.5 is well defined.

Assumption 3.5 holds when $s_{k,\text{cp}}$ is computed exactly, i.e., $\widehat{s}_{k,\text{cp}} = s_{k,\text{cp}}$. Indeed, let $\|s_{k,\min}\|$ be the smallest norm across all possible choices of $s_{k,\text{cp}}$. Several cases can occur. Firstly, if $\|s_{k,\min}\| > 0$, then $\|s_{k,\text{cp}}\| > 0$ necessarily, and Assumption 3.5 holds with $\kappa_s := \min(1, \|s_{k,\text{cp}}\|/\|s_{k,\min}\|)$. If, on the other hand, $\|s_{k,\min}\| = 0$, the same holds if we compute $s_{k,\text{cp}} \neq 0$ but, should we compute $s_{k,\text{cp}} = 0$, Proposition 2.2 would imply that x_k is stationary and the iterations would stop. This case will be clarified in Lemma 3.5.

Details on how we satisfy Assumption 3.5 when $\widehat{s}_{k,\text{cp}} \neq s_{k,\text{cp}}$ in certain situations relevant in practice can be found in Section 4. We further comment on Assumption 3.5 in Section 6.

In the same fashion as [31], we bound evaluation errors in terms of the step. Similar assumptions are made in [22] in a trust-region context.

ASSUMPTION 3.6. *There exist* $\kappa_f > 0$ *and* $\kappa_\nabla > 0$ *such that, for all* $k \in \mathbb{N}$,

$$|f(x_k) - \widehat{f}(x_k)| \leq \kappa_f \|s_k\|^2, \tag{3.2}$$

$$|f(x_k + s_k) - \widehat{f}(x_k + s_k)| \leq \kappa_f \|s_k\|^2, \tag{3.3}$$

$$\|\nabla f(x_k) - \widehat{\nabla} f(x_k)\| \leq \kappa_\nabla \|s_k\|. \tag{3.4}$$

Finally, we assume that the objective is bounded below, which is only required in the complexity analysis.

ASSUMPTION 3.7. *There exists* $(f+h)_{\text{low}} \in \mathbb{R}$ *such that* $(f+h)(x) \geq (f+h)_{\text{low}}$ *for all* $x \in \mathbb{R}^n$.

3.2. Convergence Analysis. Our first result relates the decrease predicted by the model to the step size.

LEMMA 3.2. *Let Assumption 3.4 hold. Then,*

$$\varphi(0;x_k) + \psi(0;x_k) - (\varphi(s_k;x_k) + \psi(s_k;x_k)) \geq \tfrac{1}{2}(1-\theta_1)\theta_2^{-2}\nu_k^{-1}\|s_k\|^2.$$

Proof. Assumption 3.4, (2.14) and line 10 of Algorithm 3.1 yield

$$\begin{aligned}\varphi(0;x_k) + \psi(0;x_k) - (\varphi(s_k;x_k) + \psi(s_k;x_k)) &\geq (1-\theta_1)\widehat{\xi}_{k,\text{cp}} \\ &\geq \tfrac{1}{2}(1-\theta_1)\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\|^2 \\ &\geq \tfrac{1}{2}(1-\theta_1)\theta_2^{-2}\nu_k^{-1}\|s_k\|^2. \qquad \square\end{aligned}$$

Our next result mirrors [6, Theorem 4.1] and shows that whenever σ_k exceeds a threshold σ_{succ}, iteration k is very successful and σ_{k+1} decreases.

LEMMA 3.3. *Let Assumptions 3.1 to 3.4 and 3.6 be satisfied and define*

$$\sigma_{\text{succ}} := \max\left(\frac{\theta_1\theta_2^2(L + \kappa_B + 2\kappa_h + 4\kappa_f + 2\kappa_\nabla)}{(1-\theta_1)(1-\widehat{\eta}_2)}, \lambda^{-1}\right) > 0.$$

If, at iteration k *of Algorithm 3.1,* $s_k \neq 0$ *and* $\sigma_k \geq \sigma_{\text{succ}}$, *then* $\widehat{\rho}_k \geq \widehat{\eta}_2$, *and iteration* k *is very successful.*

Proof. As in the proof of [6, Theorem 4.1], σ_k increases as long as it is below $\lambda_{x_k}^{-1}$. Thus, we assume that $\sigma_k \geq \lambda^{-1}$. The definitions of $\widehat{\rho}_k$ and φ, Assumption 3.4, the triangle inequality and Lemma 3.2 yield

$$\begin{aligned}
&|\widehat{\rho}_k - 1| \\
&= \frac{|\widehat{f}(x_k + s_k) - \widehat{f}(x_k) - \widehat{\nabla} f(x_k)^T s_k - \frac{1}{2} s_k^T B_k s_k + h(x_k + s_k) - \psi(s_k; x_k)|}{\varphi(0; x) + \psi(0; x) - (\varphi(s_k; x_k) + \psi(s_k; x_k))} \\
&\leq \frac{|\widehat{f}(x_k + s_k) - \widehat{f}(x_k) - \widehat{\nabla} f(x_k)^T s_k| + |\frac{1}{2} s_k^T B_k s_k| + |h(x_k + s_k) - \psi(s_k; x_k)|}{\frac{1}{2}(1 - \theta_1)\theta_2^{-2} \nu_k^{-1} \|s_k\|^2}.
\end{aligned}$$

The triangle inequality along with Assumptions 3.1 and 3.6 bound the first term in the numerator as

$$\begin{aligned}
&|\widehat{f}(x_k + s_k) - \widehat{f}(x_k) - \widehat{\nabla} f(x_k)^T s_k| \\
&\qquad \leq |f(x_k + s_k) - f(x_k) - \nabla f(x_k)^T s_k| + 2\kappa_f \|s_k\|^2 + \kappa_\nabla \|s_k\|^2 \\
&\qquad \leq (\tfrac{1}{2} L + 2\kappa_f + \kappa_\nabla) \|s_k\|^2.
\end{aligned}$$

Assumption 3.2 bounds the second term in the numerator by $\frac{1}{2}\|B_k\|\|s_k\|^2 \leq \frac{1}{2}\kappa_B \|s_k\|^2$. Assumption 3.3 bounds the last term in the numerator by $\kappa_h \|s_k\|^2$. After simplifying by $\|s_k\|^2$ and using $\nu_k \leq \theta_1/\sigma_k$ by definition in Algorithm 3.1, those observations give

$$|\widehat{\rho}_k - 1| \leq \frac{\theta_1 \theta_2^2 (L + \kappa_B + 2\kappa_h + 4\kappa_f + 2\kappa_\nabla)}{(1 - \theta_1)\sigma_k}.$$

Therefore, $\sigma_k \geq \sigma_{\text{succ}}$ implies that $\widehat{\rho}_k \geq \widehat{\eta}_2$. ∎

In Lemma 3.3, we showed that $\sigma_k \geq \sigma_{\text{succ}} \implies \widehat{\rho}_k \geq \widehat{\eta}_2$, which means that there is a decrease in $\widehat{f} + h$. Next, we show that there exists $\eta_1 > 0$ such that $\widehat{\rho}_k \geq \widehat{\eta}_1 \implies \rho_k \geq \eta_1$, and similarly for $\widehat{\eta}_2$. Therefore, a decrease also occurs in $f + h$ every time a step is accepted.

LEMMA 3.4. *Let Assumptions 3.4 and 3.6 hold. At iteration k, denote*

$$\rho_k := \frac{f(x_k) + h(x_k) - (f(x_k + s_k) + h(x_k + s_k))}{\varphi(0; x_k) + \psi(0; x_k) - (\varphi(s_k; x_k) + \psi(s_k; x_k))}$$

the measure of agreement between the actual and predicted decrease in $f + h$. Let $\sigma_{\min}$ be as in Algorithm 3.1 and

$$\eta_1 := \widehat{\eta}_1 - \frac{4\kappa_f \theta_1 \theta_2^2}{(1 - \theta_1)\sigma_{\min}} > 0, \qquad \eta_2 := \widehat{\eta}_2 - \frac{4\kappa_f \theta_1 \theta_2^2}{(1 - \theta_1)\sigma_{\min}} > 0.$$

Then, $\widehat{\rho}_k \geq \widehat{\eta}_1 \implies \rho_k \geq \eta_1$ and $\widehat{\rho}_k \geq \widehat{\eta}_2 \implies \rho_k \geq \eta_2$.

Proof. By definition of $\widehat{\rho}_k$ and ρ_k,

$$\widehat{\rho}_k = \rho_k + \frac{(\widehat{f} - f)(x_k) + (f - \widehat{f})(x_k + s_k)}{(\varphi + \psi)(0; x_k) - (\varphi + \psi)(s_k; x_k)}.$$

Because Algorithm 3.1 enforces $\sigma_k \geq \sigma_{\min} > 0$, we obtain $\nu_k \leq \theta_1/\sigma_k \leq \theta_1/\sigma_{\min}$. Thus, Lemma 3.2 and Assumption 3.6 give

$$|\widehat{\rho}_k - \rho_k| \leq \frac{2\kappa_f \|s_k\|^2}{\frac{1}{2}(1-\theta_1)\theta_2^{-2}\nu_k^{-1}\|s_k\|^2} \leq \frac{4\kappa_f \theta_1 \theta_2^2}{(1-\theta_1)\sigma_{\min}}.$$

Now, if $\widehat{\rho}_k \geq \widehat{\eta}_1$,

$$\rho_k \geq \widehat{\eta}_1 - \frac{4\kappa_f \theta_1 \theta_2^2}{(1-\theta_1)\sigma_{\min}} = \eta_1.$$

The lower bound on $\sigma_{\min}$ ensures $\eta_1 > 0$. The same holds for η_2 because $\widehat{\eta}_2 \geq \widehat{\eta}_1$. ∎

Lemmas 3.3 and 3.4 together imply that $\widehat{\rho}_k \geq \widehat{\eta}_1$ guarantees a decrease in $f + h$.

The next result is classic and considers the case where only a finite number of successful iterations occur.

LEMMA 3.5. *Let Assumptions 3.1 to 3.4 and 3.6 be satisfied. Suppose Algorithm 3.1 generates finitely many successful iterations. Then $x_k = x_\star$ for all k sufficiently large and $x_\star$ is first-order stationary.*

Proof. By assumption, there is $k_0 \in \mathbb{N}$ such that $x_k = x_\star$ for all $k \geq k_0$. If $x_\star$ is not stationary, as of iteration k_0, Algorithm 3.1 repeatedly computes nonzero steps s_k, all of which are rejected, i.e., $\rho_k < \eta_1$. Thus, for all $k \geq k_0$, $\sigma_{k+1} > \sigma_k$. Hence, for sufficiently large k, $\sigma_k > \sigma_{\text{succ}}$, which triggers a successful iteration, and is absurd. ∎

Lemma 3.3 implies that there exists $\sigma_{\max} = \min(\sigma_0, \gamma_2\sigma_{\text{succ}})$ such that $\sigma_k \leq \sigma_{\max}$ for all $k \in \mathbb{N}$. Consequently, Assumption 3.2 yields that for all $k \in \mathbb{N}$,

$$\nu_{\min} \leq \nu_k \leq \nu_{\max}, \quad \nu_{\min} := \theta_1/(\kappa_B + \sigma_{\max}), \quad \nu_{\max} := \theta_1/\sigma_{\min}. \tag{3.5}$$

Let $\epsilon > 0$. We seek a bound on $k_\epsilon := \min\{k \in \mathbb{N} \mid \nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\| < \epsilon\} = |\mathcal{S}(\epsilon)| + |\mathcal{U}(\epsilon)| + 1$, where

$$\mathcal{S}(\epsilon) := \{k \in \mathbb{N} \mid \widehat{\rho}_k \geq \widehat{\eta}_1 \text{ and } k < k_\epsilon\}, \qquad \mathcal{U}(\epsilon) := \{k \in \mathbb{N} \mid \widehat{\rho}_k < \widehat{\eta}_1 \text{ and } k < k_\epsilon\}.$$

LEMMA 3.6. *Let Assumptions 3.1 to 3.4, 3.6 and 3.7 be satisfied. Assume that Algorithm 3.1 generates infinitely many successful iterations. Then,*

$$|\mathcal{S}(\epsilon)| \leq \frac{(f+h)(x_0) - (f+h)_{\text{low}}}{\frac{1}{2}\eta_1(1-\theta_1)\nu_{\min}}\ \epsilon^{-2} := \omega_s \epsilon^{-2},$$

where $\nu_{\min}$ is defined in (3.5).

Proof. Let $k \in \mathcal{S}(\epsilon)$. By definition, $\widehat{\rho}_k \geq \widehat{\eta}_1$, which, by Lemma 3.4, implies that $\rho_k \geq \eta_1$. Assumption 3.4, (3.5), (2.14) and the fact that $k < k_\epsilon$ then imply

$$\begin{aligned}
(f+h)(x_k) - (f+h)(x_k + s_k) &\geq \eta_1((\varphi+\psi)(0;x_k) - (\varphi+\psi)(s_k;x_k)) \\
&\geq \eta_1(1-\theta_1)\widehat{\xi}_{k,\text{cp}} \\
&\geq \tfrac{1}{2}\eta_1(1-\theta_1)\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\|^2 \\
&\geq \tfrac{1}{2}\eta_1(1-\theta_1)\nu_k\epsilon^2 \\
&\geq \tfrac{1}{2}\eta_1(1-\theta_1)\nu_{\min}\epsilon^2.
\end{aligned}$$

The rest of the proof is classic and identical to, e.g., [6, Lemma 4.3]. ∎

It is remarkable that the bound in Lemma 3.6 is identical to that of the standard R2N, which is more apparent when comparing with [6, Lemma 4.3] than with [25, Theorem 6.4]. The extra factor $\frac{1}{2}$ in the denominator of our bound on $|\mathcal{S}(\epsilon)|$ is due to the fact that we use $\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\|$ as stationarity measure instead of $\nu_k^{-1/2}\widehat{\xi}_{k,\text{cp}}^{1/2}$, as in [6].

Finally, we recover a worst-case complexity bound of the same order as in the analysis with exact proximal operator evaluations. The proof is identical to that of, e.g., [6, Theorem 4.5], and is omitted.

THEOREM 3.7. *Let Assumptions 3.1 to 3.4, 3.6 and 3.7 be satisfied. Then,*

$$|\mathcal{S}(\epsilon)| + |\mathcal{U}(\epsilon)| = \left(1 + |\log_{\gamma_1}(\gamma_3)|\right)\omega_s\epsilon^{-2} + \log_{\gamma_1}(\sigma_{\max}/\sigma_0) = O(\epsilon^{-2}),$$

where ω_s is defined in Lemma 3.6.

Theorem 3.7 shows that iR2N brings the measure $\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\|$ below ϵ in $O(\epsilon^{-2})$ iterations. That measure is not a stationarity measure because it includes the inexactness on $\widehat{s}_{k,\text{cp}}$. By Assumption 3.5, there exists an exact Cauchy step $s_{k_\epsilon,\text{cp}}$ such that

$$\nu_k^{-1}\|s_{k_\epsilon,\text{cp}}\| \leq \kappa_s^{-1}\nu_k^{-1}\|\widehat{s}_{k_\epsilon,\text{cp}}\| < \kappa_s^{-1}\epsilon. \tag{3.6}$$

Thus, if $\nu_k^{-1}\|\widehat{s}_{k_\epsilon,\text{cp}}\|$ is small, $\nu_k^{-1}\|s_{k_\epsilon,\text{cp}}\|$ is comparably small. The next result shows that when the latter occurs, we have identified a near stationary point, and marks the impact of κ_s on the analysis.

THEOREM 3.8. *Let Assumptions 3.5 and 3.6 be satisfied. Let $\epsilon > 0$ and assume $\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\| < \epsilon$. There exists $s_{k,\text{cp}} \in \text{prox}_{\nu\psi(\cdot;x_k)}(-\nu_k\widehat{\nabla}f(x_k))$ that satisfies Assumption 3.5 such that $\|s_{k,\text{cp}}\| < \kappa_s^{-1}\nu_{\max}\epsilon$, and $u_k \in \nabla f(x_k) + \partial\psi(s_{k,\text{cp}}; x_k)$ such that*

$$\|u_k\| < \left(\kappa_\nabla\theta_2\nu_{\max} + \kappa_s^{-1}\right)\epsilon. \tag{3.7}$$

Proof. By definition, $s_{k,\text{cp}}$ is an exact minimizer of (2.4), thus

$$\begin{aligned} 0 &\in \widehat{\nabla}f(x_k) + \nu_k^{-1}s_{k,\text{cp}} + \partial\psi(s_{k,\text{cp}}; x_k) \\ &= \nabla f(x_k) + g_k + \nu_k^{-1}s_{k,\text{cp}} + \partial\psi(s_{k,\text{cp}}; x_k), \end{aligned} \tag{3.8}$$

where $g_k := \widehat{\nabla}f(x_k) - \nabla f(x_k)$ and $\|g_k\| \leq \kappa_\nabla\|s_k\| \leq \kappa_\nabla\theta_2\|\widehat{s}_{k,\text{cp}}\|$ from Assumption 3.6 and line 10 of Algorithm 3.1. By (3.5) and $\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\| < \epsilon$, $\|\widehat{s}_{k,\text{cp}}\| \leq \nu_k\epsilon < \nu_{\max}\epsilon$. Thus, $\|g_k\| < \kappa_\nabla\theta_2\nu_{\max}\epsilon$.

On the other hand, Assumption 3.5 gives

$$\|\nu_k^{-1}s_{k,\text{cp}}\| \leq \kappa_s^{-1}\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\| < \kappa_s^{-1}\epsilon.$$

Now, (3.8) implies that

$$u_k := -(g_k + \nu_k^{-1}s_{k,\text{cp}}) \in \nabla f(x_k) + \partial\psi(s_{k,\text{cp}}; x_k).$$

Because $\|u_k\| \leq \|g_k\| + \|\nu_k^{-1}s_{k,\text{cp}}\|$, (3.7) holds. Finally, the same reasoning as above shows that $\|s_{k,\text{cp}}\|$ is bounded as announced. ∎

The following results directly from Theorem 3.7 and mirrors [29, Lemma 3].

LEMMA 3.9. *Under the assumptions of Theorem 3.7 and Assumption 3.5, there exists an infinite index set $N \subseteq \mathbb{N}$ and $\{s_{k,\text{cp}}\}$ where $s_{k,\text{cp}} \in \text{prox}_{\nu\psi(\cdot;x_k)}(-\nu_k \widehat{\nabla} f(x_k))$ for all k such that*

1. *$\{\widehat{s}_{k,\text{cp}}\}_N \to 0$ and $\{s_{k,\text{cp}}\}_N \to 0$,*
2. *$\{s_k\}_N \to 0$*
3. *there exists $u_k \in \nabla f(x_k) + \partial\psi(s_{k,\text{cp}}; x_k)$ such that $\{u_k\}_N \to 0$.*

Proof. Claim 1 follows directly from Theorem 3.7, (3.5) and (3.6). Claim 2 follows from Line 10 of Algorithm 3.1. Claim 3 results from Theorem 3.8. ∎

We close this section with a result stating that every limit point of the sequence $\{x_k\}_N$ generated by Algorithm 3.1 is stationary, where N is defined in Lemma 3.9, under an assumption on the subdifferential of the models $\psi(\cdot; x_k)$.

Recall that for a sequence of sets $\{\mathcal{A}_k\}$ with $\mathcal{A}_k \subseteq \mathbb{R}^n$ for all $k \in \mathbb{N}$, the set $\limsup \mathcal{A}_k$ is the set of limits of all possible convergent sequences $\{a_k\}_N$ with $N \subset \mathbb{N}$ infinite and $a_k \in \mathcal{A}_k$ for all $k \in N$.

THEOREM 3.10. *Under the assumptions of Theorem 3.7, Assumptions 2.1 and 3.5, let $N \subseteq \mathbb{N}$ be as in Lemma 3.9. Assume that $\{x_k\}_N \to \bar{x}$ and that*

$$\limsup_{k \in N} \partial\psi(s_{k,\text{cp}}; x_k) \subseteq \partial\psi(0; \bar{x}). \tag{3.9}$$

Then $\bar{x}$ is stationary for (1.1).

Proof. By our assumptions, Lemma 3.9, continuity of ∇f and Assumption 2.1,

$$0 \in \nabla f(\bar{x}) + \limsup_{k \in N} \partial\psi(s_{k,\text{cp}}; x_k) \subseteq \nabla f(\bar{x}) + \partial\psi(0; \bar{x}) \subseteq \nabla f(\bar{x}) + \partial h(\bar{x}).$$

Thus, $\bar{x}$ is stationary for (1.1). ∎

As Leconte and Orban [29] explain, (3.9) holds in several relevant cases, e.g.,

1. each $\psi(\cdot; x_k)$ and $\psi(\cdot; \bar{x})$ are proper, lsc and convex with $\psi(\cdot; x_k) \to \psi(\cdot; \bar{x})$ in the epigraphical sense, and $0 \in \text{dom } \psi(\cdot; \bar{x})$;
2. $\psi(s; x) := h(x + s)$ and $h(x_k + s_{k,\text{cp}}) \to h(\bar{x})$ as would occur, in particular but not exclusively, when h is continuous.

4. Evaluation of inexact proximal operators. In this section, we discuss the practical implementation of Algorithm 3.1 with focus on computing an approximate solution of (2.8) that satisfies Assumption 3.5. Our approach is simple: assume that an upper bound $M_k > 0$ on $\|s_{k,\text{cp}}\|$ can be determined based on properties of $\psi(\cdot; x_k)$. Assume also that a descent procedure is applied to (2.8) starting from $s = 0$ that generates iterates $\widehat{s}_j$, $j \geq 0$. Then, stopping the procedure as soon as $\|\widehat{s}_j\| \geq \kappa_s M_k$ ensures that Assumption 3.5 holds.

We consider three regularizers whose proximal operators (1.2) are not known analytically and must be computed inexactly:

$$h(x) = \ell_p(x) = \|x\|_p \qquad (1 \leq p < \infty), \tag{4.1}$$

$$h(x) = \text{TV}_p(x) = \Big(\sum_i |x_i - x_{i-1}|^p\Big)^{1/p} \qquad (1 \leq p < \infty), \tag{4.2}$$

$$h(x) = \chi_{p,r}(x) = \begin{cases} 0 & \text{if } \|x\|_p^p \leq r \\ \infty & \text{otherwise} \end{cases} \qquad (0 < p < 1), \tag{4.3}$$

where TV_p is the one-dimensional total-variation operator, and $\chi_{p,r}$ is the indicator of the ℓ_p-pseudo norm "ball" of radius $r^{1/p}$ for $r > 0$.

The next lemmas derive bounds on the norm of solutions to the proximal problems associated with those regularizers.

LEMMA 4.1. *Let h be given by (4.1) and $\psi(s; x_k) := h(x_k + s)$ with $s \in \mathbb{R}^n$. The unique solution $s_{k,\text{cp}}$ of (2.8) is such that*

$$\|s_{k,\text{cp}}\| \leq \begin{cases} \nu_k(\|\widehat{\nabla} f(x_k)\| + n^{1/p-1/2}) & (1 \leq p < 2) \\ \nu_k(\|\widehat{\nabla} f(x_k)\| + 1) & (p \geq 2). \end{cases} \tag{4.4}$$

Proof. Since $\psi(\cdot; x_k)$ is convex, (2.8) is strongly convex and, therefore, has a unique solution $s_{k,\text{cp}}$. The necessary optimality conditions read

$$\widehat{\nabla} f(x_k) + \nu_k^{-1} s_{k,\text{cp}} + u_k = 0, \qquad u_k \in \partial\psi(s_{k,\text{cp}}; x_k).$$

Here, $\partial\psi(s_{k,\text{cp}}; x_k) = \{u \in \mathbb{R}^n \mid \|u\|_q \leq 1 \text{ and } u^T(s_{k,\text{cp}} + x_k) = \|s_{k,\text{cp}} + x\|_p\}$, where q is such that $1/p + 1/q = 1$. By equivalence of norms,

$$\|u_k\| \leq n^{1/2-1/q} \|u_k\|_q \leq n^{1/2-1/q} = n^{1/p-1/2}.$$

When $1 \leq p \leq 2$, the latter bound is attained for $u_k := (n^{-1/q}, n^{-1/q}, \ldots, n^{-1/q})$ with $\|u_k\|_q = 1$. When $p > 2$, the bound simplifies to $\|u_k\| \leq 1$, which is attained for $u_k := (1, 0, \ldots, 0)$. Thus, $\|s_{k,\text{cp}}\| = \nu_k \|\widehat{\nabla} f(x_k) + u_k\| \leq \nu_k(\|\widehat{\nabla} f(x_k)\| + \|u_k\|)$, which yields (4.4). ∎

The next result helps bound solutions of (2.8) when h is given by (4.2), but is more general, which is why it is stated separately.

LEMMA 4.2. *Let $A \in \mathbb{R}^{m \times n}$, $h(x) := \|Ax\|_\bullet$ where $\|\cdot\|_\bullet$ is a norm on $\mathbb{R}^m$, and $\psi(s; x_k) := h(x_k + s)$. The unique solution $s_{k,\text{cp}}$ of (2.8) satisfies*

$$\|s_{k,\text{cp}}\| \leq \nu_k \left(\|\widehat{\nabla} f(x_k)\| + \|A\| \, \|u_k\|\right), \tag{4.5}$$

where $u_k \in \partial\|A(x_k + s_{k,\text{cp}})\|_\bullet$.

Proof. Here again, $s_{k,\text{cp}}$ is unique by strong convexity of (2.8). For $\eta(y) := \|y\|_\bullet$,

$$\partial\eta(y) = \{u \in \mathbb{R}^m \mid \|u\|_\star \leq 1 \text{ and } u^T y = \|y\|_\bullet\},$$

where $\|\cdot\|_\star$ is the dual norm of $\|\cdot\|_\bullet$. By [35, Theorem 23, 9], $\partial\psi(s; x_k) = A^T \partial\eta(A(x_k + s))$. Thus, the first-order optimality conditions of (2.8) imply

$$0 \in \widehat{\nabla} f(x_k) + \nu_k^{-1} s_{k,\text{cp}} + A^T u_k,$$

where $u_k \in \partial\eta(A(x_k + s_{k,\text{cp}}))$. We extract $s_{k,\text{cp}} = -\nu_k(\widehat{\nabla} f(x_k) + A^T u_k)$, and $\|s_{k,\text{cp}}\| \leq \nu_k(\|\widehat{\nabla} f(x_k)\| + \|A^T\| \, \|u_k\|)$, which is (4.5) since $\|A\| = \|A^T\|$. ∎

Lemma 4.2 does not state a bound on $\|u_k\|$ as one would depend on $\|\cdot\|_\bullet$ and the bound $\|u_k\|_\star \leq 1$. The next corollary applies Lemma 4.2 to (4.2).

COROLLARY 4.3. *Let h be as in (4.2) and $\psi(s; x_k) := h(x_k + s)$. The unique solution $s_{k,\text{cp}}$ of (2.8) satisfies*

$$\|s_{k,\text{cp}}\| \leq \begin{cases} \nu_k \left(\|\widehat{\nabla} f(x_k)\| + 2\sin\left(\frac{\pi(n-1)}{2n}\right) n^{1/p-1/2}\right) & (1 \leq p < 2) \\ \nu_k \left(\|\widehat{\nabla} f(x_k)\| + 2\sin\left(\frac{\pi(n-1)}{2n}\right)\right) & (p \geq 2). \end{cases} \tag{4.6}$$

Proof. Apply Lemma 4.2 with $\|\cdot\|_\bullet = \|\cdot\|_p$ and

$$A \coloneqq \begin{bmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(n-1)\times n}.$$

Note that $A^T A$ is the centered finite-difference operator for second derivatives, which is symmetric, tridiagonal and Toeplitz. Its eigenvalues are thus known in closed form, hence the value of $\|A\|$ [40, p. 54]. Finally, $\|u_k\|$ can be bounded as in the proof of Lemma 4.1. ∎

The final lemma derives a bound on the solution of the proximal problem associated to the indicator function in (4.3).

LEMMA 4.4. *Let h be as in (4.3) and $\psi(s; x_k) \coloneqq h(x_k + s)$. Any solution $s_{k,\text{cp}}$ of (2.8) satisfies*

$$\|s_{k,\text{cp}}\| \leq r^{1/p} + \|x_k\|. \tag{4.7}$$

Proof. Because $0 < p < 1$, $t \mapsto t^p$ is concave for $t \geq 0$, and thus subadditive, i.e., $(a+b)^p \leq a^p + b^p$ for any $a, b \geq 0$. Let $u \in \mathbb{R}^n$. By recurrence on n, $\|u\|_p^p = \sum_{i=1}^n |u_i|^p \geq (\sum_{i=1}^n |u_i|)^p$, which states that $\|u\|_p \geq \|u\|_1$. This implies that the unit "ball" in ℓ_p-pseudo-norm is a subset of the unit ℓ_1-norm ball. In turn, the latter is a subset of the unit ℓ_2-norm ball. A scaling argument shows that the same holds with balls of radius $r > 0$. Therefore, because $\|x_k + s_{k,\text{cp}}\|_p \leq r^{1/p}$, we have $\|x_k + s_{k,\text{cp}}\| \leq r^{1/p}$. The triangle inequality yields $\|s_{k,\text{cp}}\| \leq \|x_k + s_{k,\text{cp}}\| + \|x_k\| \leq r^{1/p} + \|x_k\|$. ∎

In (4.4), (4.6) and (4.7), the bound on $\|s_{k,\text{cp}}\|$ depends only on known quantities at iteration k. Thus, we can enforce Assumption 3.5 by stopping the inexact proximal procedure as soon as $\|\widehat{s}_{k,\text{cp}}^{(j)}\|$ exceeds a fixed fraction of said bound.

5. Numerical experiments. In this section, we present numerical experiments indicating that exploiting inexact objective values, gradients and proximal operators can reduce computational cost substantially. We implement Algorithm 3.1 in the Julia language [14] as a modification of the R2N solver [25] in [7].

The implementation of the proximal operator of (4.1) and (4.2), which are both convex, is available from the Julia interface [2] to the `proxTV` library [9]. Both implement iterative methods. The method for (4.1) computes projected quasi-Newton search directions, and performs a backtracking line search to determine the step size. That for (4.2) alternates between gradient projection into the ℓ_p-norm ball and Frank-Wolfe steps. After each update, the primal solution is reconstructed from the dual variable, and a new gradient is computed.

Our implementation of the proximal operator of (4.3) is based on the Iteratively Reweighted ℓ_p-Ball Projection (IRBP) scheme of [43]. At each iteration, IRBP approximates the ℓ_p-"ball" norm via a weighted linearization of the nonconvex set around the current iterate. This results in a convex subproblem describing a projection into a weighted ℓ_1-norm ball, which can be solved efficiently [21]. A smoothing vector is maintained and adaptively updated to avoid numerical instability and improve convergence. The nonconvex nature of $\chi_{p,r}$ implies that there may be non-global minima or saddle points [43]. Therefore, the step output by $\chi_{p,r}$ may not even induce $\widehat{\xi}_{k,\text{cp}} \geq 0$. To the best of our knowledge, there is currently no procedure that is guaranteed to determine a global minimum. In order to mitigate the issue, we

implement a multi-start strategy to increase the odds that $\widehat{s}_{k,\text{cp}}$ be a global solution. Our strategy is not always successful, but nevertheless often results in acceptable steps. Part of future work is to find a procedure that identifies a global minimizer. Our implementation is available from [1].

In each case, inexactness in the proximal operator evaluations is controlled by $0 < \kappa_s \leq 1$ in Assumption 3.5. For $\kappa_s \approx 0$, the expectation on the quality of $\widehat{s}_{k,\text{cp}}$ is at its lowest, i.e., Assumption 3.5 is easiest to satisfy, but (5.1) is harder to reach. Thus the solver may spend less time inside each (cheap) proximal operator evaluation at the cost of potentially performing more (costly) outer iterations. On the other hand, when $\kappa_s \approx 1$, the $\widehat{s}_{k,\text{cp}}$ should be close to an exact solution. In this case, the solver may spend more time than necessary inside each proximal operator evaluation, which may adversely affect the total solution time. In our experiments, we vary the value of κ_s to assess the impact of the inexactness on the performance of iR2N.

Step 9 in Algorithm 3.1 is performed by a special case of Algorithm 3.1 with $B = 0$ in which the proximal step computation is the only subproblem. In effect, that is a variant of the R2 algorithm [3, Algorithm 6.1] extended to the inexact proximal framework. We refer to this variant as iR2. Although iR2 is also allowed to perform inexact evaluations of its smooth objective and gradient, we evaluate the quadratic model $\varphi(s; x_k)$ exactly in our experiments.

Each procedure to solve (4.1)–(4.3) comes with its original stopping condition. We say that we run iR2N in *exact* mode when we use this original stopping condition, independently of Assumption 3.5, and we consider that the resulting proximal operator is then evaluated exactly. By contrast, we run iR2N in *inexact* mode when the iterations of the proximal operator evaluation are terminated as soon as either (i) $\|\widehat{s}_{k,\text{cp}}\| \geq \kappa_s M_k$, where M_k is the upper bound on $\|s_{k,\text{cp}}\|$ given in (4.4), (4.6), or (4.7), or (ii) the original stopping condition of the proximal operator evaluation is met. In proximal operator evaluations, iR2 uses the same value of κ_s as iR2N.

Inequalities (3.6) suggest using $\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\| \leq \kappa_s\epsilon$ as stopping criterion in Algorithm 3.1, since it guarantees that $\nu_k^{-1}\|s_{k,\text{cp}}\| \leq \epsilon$. However, we will see that small values of κ_s yield the best performance but make that stopping condition overly stringent. In addition, the bound M_k given in Lemmas 4.1, 4.2 and 4.4 need not be tight, and could indeed be quite loose. For those reasons, all our experiments use the simple stopping condition

$$\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\| \leq \epsilon. \tag{5.1}$$

In the next sections, we report the performance of iR2N on problems that use the inexact proximal operators described above. In Subsections 5.1 to 5.3, both the objective and gradient are assumed to be evaluated exactly, i.e., only subject to the limits of floating-point operations. In Subsection 5.4, we consider inexact evaluations of the objective and gradient. All our tests are performed in double precision on a 2020 MacBook Air with an M1 chip (8-core CPU, 8 GB unified memory).

Because f in our test problems is based on randomly-generated data, we average the statistics over 10 runs. It is useful to keep in mind that each iR2N and iR2 iteration evaluates a single proximal operator—see Line (8) of Algorithm 3.1. Tables in the next sections use the following headers: "κ_s" is the value of the inexactness parameter in Assumption 3.5, "iR2N" is the average number of outer iterations, "iR2" is the average number of inner iterations per outer iteration, "prox" is the average number of iterations per proximal operator evaluation, and "time (s)" is the average CPU solution time in seconds.

5.1. Basis pursuit denoising problem (BPDN). The BPDN problem is stated as

$$\min_{x\in\mathbb{R}^n} \tfrac{1}{2}\|Ax-b\|_2^2 + \mu\|x\|_p, \tag{5.2}$$

where $\mu = 10^{-1}$, $A \in \mathbb{R}^{200\times 512}$ is random with orthonormal rows, $b = A\bar{x} + \varepsilon$, $\bar{x}$ has 10 nonzeros, and ε is a noise vector from a normal $(0, 1)$ distribution. We use $p = 1.1$ to attempt to recover a sparse solution. In (5.1), we set $\epsilon = 10^{-6}$.

TABLE 5.1
Statistics on (5.2) for several values of κ_s.

κ_s	iR2N	iR2	prox	time (s)
1.00e−07	1.61e+01	1.21e+02	1.02e+02	5.03e+00
1.00e−05	1.57e+01	1.63e+02	1.90e+02	9.80e+00
1.00e−03	1.49e+01	1.33e+01	4.02e+02	1.55e+01
1.00e−02	1.49e+01	1.78e+01	6.02e+02	1.77e+01
1.00e−01	1.45e+01	1.39e+01	5.81e+02	1.32e+01
5.00e−01	1.45e+01	1.37e+01	5.90e+02	1.28e+01
9.00e−01	1.45e+01	1.39e+01	5.80e+02	1.25e+01
9.90e−01	1.46e+01	1.37e+01	5.90e+02	1.38e+01
exact mode	1.45e+01	1.35e+01	5.68e+02	1.20e+01

Table 5.1 shows that the average number of iR2N/iR2 iterations decreases globally as κ_s increases. The proximal operator iterations increase as κ_s increases, as expected. For small values of κ_s, inexact mode yields a substantial reduction in the number of proximal iterations and solution time compared with exact mode at the expense of a modest increase in outer iterations. For large values of κ_s the behavior of iR2N is close to that of exact mode.

Figure 5.1 shows that the solutions produced in exact and inexact mode are essentially identical, and that both recover the sparse support of $\bar{x}$.

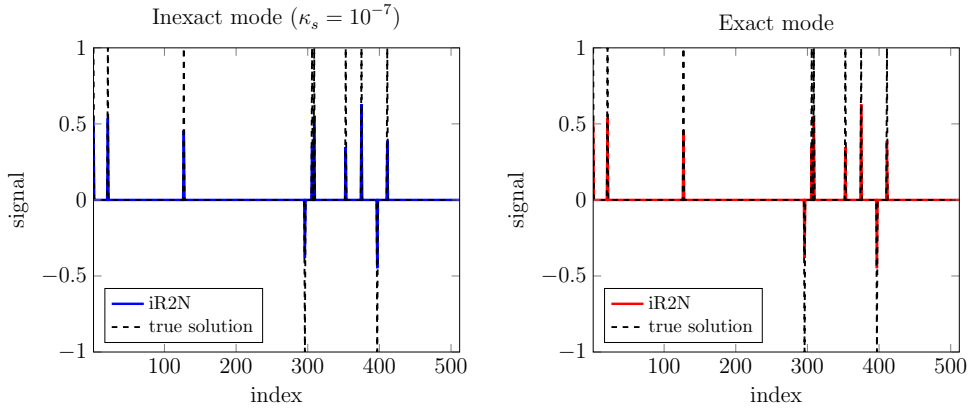

FIG. 5.1. *Components of the solution of (5.2) found by iR2N and of $\bar{x}$.*

5.2. Matrix completion problem. The problem is stated as

$$\min_{x\in\mathbb{R}^n} \tfrac{1}{2}\|P(x-A)\|_F^2 + \mu \mathrm{TV}_p(x), \tag{5.3}$$

where $\mu = 10^{-1}$, $p = 1.1$ and $A \in \mathbb{R}^{10\times 12}$ is a fixed matrix representing an image and the operator P only retains a subset of pixels. In (5.1), $\epsilon = 10^{-3}$.

Table 5.2 gathers our results on (5.3). The benefits of choosing κ_s small are similar to those in Table 5.1. Figure 5.2 shows that the reconstruction error with the solutions of exact and inexact mode are close, as is the discrepancy between the two solutions.

TABLE 5.2
Statistics on (5.3) for several values of κ_s.

κ_s	iR2N	iR2	prox	time (s)
1.00e−07	3.69e+01	3.41e+02	5.88e+02	9.46e+01
1.00e−05	3.72e+01	3.03e+02	8.71e+02	1.42e+02
1.00e−03	3.69e+01	2.09e+02	3.76e+03	3.54e+02
1.00e−02	3.77e+01	2.12e+02	4.06e+03	3.73e+02
1.00e−01	3.41e+01	1.90e+02	4.37e+03	3.25e+02
5.00e−01	3.56e+01	2.19e+02	4.31e+03	3.54e+02
9.00e−01	3.77e+01	1.81e+02	4.49e+03	3.57e+02
9.90e−01	3.55e+01	2.01e+02	4.27e+03	3.54e+02
exact mode	3.18e+01	1.67e+02	4.49e+03	3.36e+02

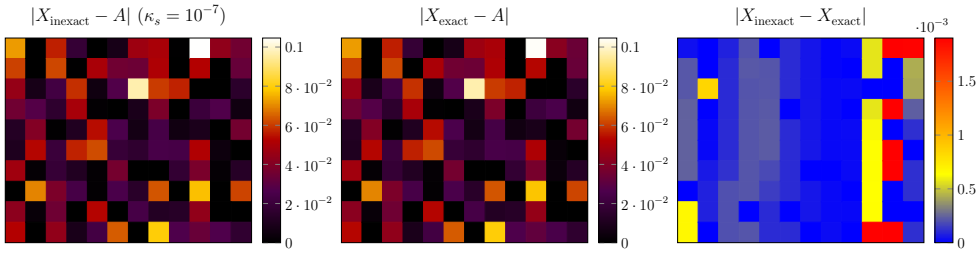

FIG. 5.2. *Left: Heatmap of the difference between the solution X found by iR2N in inexact and exact mode, and A. Right: Difference between the two solutions. The values masked by P are set to zero and shown in black.*

5.3. Fitzhugh-Nagumo inverse problem. The FitzHugh–Nagumo system is a simplified representation of a neuron's action potential modeled by the system of differential equations

$$V'(t) = x_2^{-1}(V(t) - \tfrac{1}{3}V(t)^3 - W(t) + x_1), \qquad W'(t) = x_2(x_3 V(t) - x_4 W(t) + x_5). \tag{5.4}$$

We use initial conditions $V(0) = 2$ and $W(0) = 0$, and generate data $\bar{v}(x), \bar{w}(x)$ by solving (5.4) with $\bar{x} = (0, 0.2, 1, 0, 0)$, which corresponds to the Van der Pol oscillator, to which we add random noise. We then aim to recover $\bar{x}$ by minimizing the misfit while encouraging a sparse solution:

$$\min_{x\in\mathbb{R}^5} \tfrac{1}{2}\|F(x)\|_2^2 + \chi_{p,r}(x), \tag{5.5}$$

where $p = 0.5$, $r = 2$, $F : \mathbb{R}^5 \to \mathbb{R}^{2n+2}$, $F(x) := (v(x) - \bar{v}(x), w(x) - \bar{w}(x))$, and $v(x) = (v_1(x), \ldots, v_{n+1}(x))$ and $w(x) = (w_1(x), \ldots, w_{n+1}(x))$ are sampled values of V and W at $n + 1$ discretization points. We set $\epsilon = 10^{-5}$ in (5.1). Table 5.3 reports our results.

The small number of iterations per proximal call arises from the fact that $\chi_{p,r}$ is an indicator; the projection of a point that already belongs to the set requires zero

TABLE 5.3
Statistics on (5.5) for $p = \frac{1}{2}$ and $r = 2$ with several values of κ_s.

κ_s	iR2N	iR2	prox	time (s)
1.00e−07	5.14e+02	4.90e+02	3.51e−01	5.28e+00
1.00e−05	5.72e+02	4.64e+02	4.62e−01	5.21e+00
1.00e−03	6.31e+02	5.47e+02	5.96e−01	5.56e+00
1.00e−02	5.71e+02	4.81e+02	6.22e−01	5.17e+00
1.00e−01	4.95e+02	4.89e+02	4.11e−01	5.85e+00
5.00e−01	4.90e+02	4.59e+02	1.94e+00	6.42e+00
9.00e−01	5.12e+02	4.98e+02	2.06e+00	6.53e+00
9.90e−01	5.24e+02	5.09e+02	1.91e+00	6.84e+00
exact mode	4.92e+02	5.03e+02	3.92e+01	6.88e+00

iterations. The value of κ_s has little effect on the number of iR2N/iR2 iterations. As in Subsections 5.1 and 5.2, inexact mode yields a reduction in computational cost, though more modest because the smooth objective and its gradient are costlier in (5.5) than in (5.2) or (5.3). Thus, the reduction in proximal evaluations must outweigh the increase in outer iterations. Table 5.4 gives the approximate solution identified by the exact and inexact variants, and the final value of the smooth objective. Although both exact and inexact mode recover a solution that has one more nonzero than $\bar{x}$, the final smooth objective values are close to that at $\bar{x}$. Figure 5.3 plots the simulation of (5.4) with parameters found by iR2N with $\kappa_s = 1.0\text{e}{-07}$ when solving (5.5). The solutions with exact and inexact mode are indistinguishable.

TABLE 5.4
Approximate solutions of (5.5) found by the exact and inexact variants with $\kappa_s = 1.0e{-}07$. The last column shows the smooth objective value at the solution.

	x					$\frac{1}{2}\|F(x)\|^2$
True	0.00e+00	2.00e−01	1.00e+00	0.00e+00	0.00e+00	8.82e−01
Inexact	0.00e+00	2.00e−01	9.98e−01	−1.00e−02	0.00e+00	8.96e−01
Exact	0.00e+00	2.00e−01	9.98e−01	−1.00e−02	0.00e+00	8.96e−01

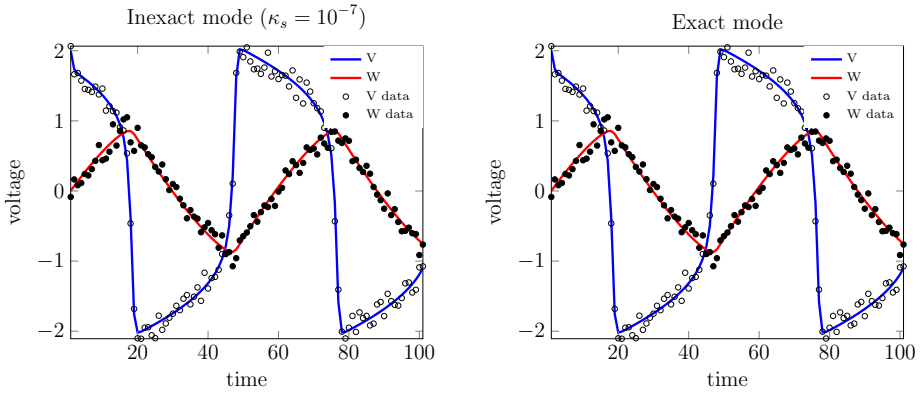

FIG. 5.3. *Simulation of (5.4) with solutions of (5.5) found by iR2N.*

5.4. Inexact objective and gradient evaluations. We now consider inexact evaluations of the smooth objective and its gradient. In (5.5), each evaluation of F involves solving an ODE system numerically, which inherently depends on a stopping

tolerance that introduces an approximation error. We use the Verner [42] 9/8 optimal Runge-Kutta method as implemented in [33]. In our implementation of F, the accuracy of the ODE solve can be adjusted via a parameter $\texttt{prec} > 0$ that sets the absolute and relative stopping tolerances. The gradient is computed via automatic differentiation, and hence, its accuracy also depends on `prec`. Decreasing this tolerance improves the accuracy of the objective and gradient but increases the computational cost. The results of Subsection 5.3 used $\texttt{prec} = 10^{-14}$ as the reference "exact" objective and gradient evaluations.

Because Assumption 3.6 may not be easily verifiable in practice, we propose a heuristic inspired from trust-region methods for derivative-free optimization [23, chapter 10], that consists in adapting the accuracy based on the progress of the algorithm. More precisely, we increase the accuracy on unsuccessful iterations, i.e., $\rho_k < \eta_1$ in Algorithm 3.1. At iteration k, we set `prec` to

$$\texttt{prec}(k) := \max(10^{-3} \exp(\log(10^{-14}/10^{-3})\, n_F/N),\, 10^{-14}), \tag{5.6}$$

where N is a preset maximum number of unsuccessful iterations after which $\texttt{prec} = 10^{-14}$ is always used, and n_F counts the number of unsuccessful iterations. Small values of N lead to a rapid increase in accuracy, whereas larger ones maintain low-accuracy evaluations over more iterations. Though (5.6) may not guarantee Assumption 3.6 at every iteration, the objective and gradient accuracy improves as the algorithm progresses, as required by the assumption.

We focus on (5.5) with the setting of Subsection 5.3 and we use (5.6) for inexact objective and gradient. We vary the value of N with fixed $\kappa_s = 10^{-7}$ in Table 5.5.

TABLE 5.5
Iterations and time on (5.5) with inexact objective and gradient evaluations.

N	fail rate	iter iR2N	iter iR2	prox	time (s)
exact F	0%	5.14e+02	4.90e+02	3.51e−01	5.28e+00
20	0%	5.66e+02	5.10e+02	4.55e−01	5.16e+00
50	20%	6.36e+02	5.07e+02	3.77e−01	4.31e+00
100	30%	6.31e+02	5.08e+02	3.46e−01	3.27e+00
200	80%	6.67e+02	5.47e+02	3.69e−01	2.46e+00

The first line of Table 5.5 reports the number of iterations and the solution time obtained with "exact" objective and gradient. Lines 2–5 use (5.6) for several values of N. As N increases, iR2N spends a larger fraction of its iterations in a *low*-precision regime, making it increasingly likely that Assumption 3.6 is violated. When iR2N operates with insufficient accuracy for too long, the algorithm may eventually stall, cease to make progress, and reach the maximum number of allowed iterations. The second column of Table 5.5 reports the proportion of such failed runs over ten trials. Importantly, the iteration and timing statistics shown in Table 5.5 correspond *only* to the successful runs. The failure rate increases with N, and for $N = 200$ few runs succeed. Moderate values of N yield significant benefits in terms of solution time.

In Table 5.6, we report the performance of Algorithm 3.1 using inexact objective, gradient and proximal operator evaluations following rule (5.6) on (5.5) with $N = 100$. The number of iR2N, iR2 and proximal iterations is globally unaffected by inexact evaluations, but the latter yield significant savings in terms of solution time.

6. Discussion. Method iR2N subsumes R2N [25] by allowing inexact evaluations of the objective, its gradient, and the proximal operator. Under usual global convergence conditions, we showed that inexact evaluations and proximal operators do not

TABLE 5.6

Statistics on (5.5) with increasing accuracy given by (5.6) with $N = 100$ and several values of κ_s. Each entry reports the multiplicative gain or loss compared to the reference values in Table 5.3. A value smaller than 1 indicates a gain.

κ_s	iR2N	iR2	prox	time (s)
1.00e−07	1.23e+00	1.04e+00	9.90e−01	6.20e−01
1.00e−05	1.08e+00	1.02e+00	1.38e+00	4.90e−01
1.00e−03	8.40e−01	7.70e−01	5.50e−01	2.70e−01
1.00e−02	1.00e+00	1.00e+00	1.10e+00	3.60e−01
1.00e−01	1.11e+00	9.60e−01	5.20e−01	3.00e−01
5.00e−01	9.90e−01	9.20e−01	1.19e+00	2.50e−01
9.00e−01	1.03e+00	8.80e−01	1.17e+00	3.00e−01
9.90e−01	9.40e−01	8.30e−01	1.36e+00	2.50e−01
average factor	1.03e+00	9.30e−01	1.03e+00	3.60e−01

deteriorate asymptotic complexity compared to methods that use exact evaluations. Our assumptions on the inexactness of f and ∇f are standard.

Assumption 3.5 on the inexact evaluation of proximal operators differs in nature from Definitions (ii) and (iii) of [38]. Their Definition (i), also used in [36], can be written $\|\widehat{s}_{k,\text{cp}} - s_{k,\text{cp}}\| \leq \epsilon_k$ for at least one $s_{k,\text{cp}}$, where $\{\epsilon_k\}$ is positive and summable. It is equivalent to $\|s_{k,\text{cp}}\| - \epsilon_k \leq \|\widehat{s}_{k,\text{cp}}\| \leq \|s_{k,\text{cp}}\| + \epsilon_k$, which is strictly stronger than Assumption 3.5 in that we only require one of the inequalities. Moreover, we use the specific value $\epsilon_k = (1 - \kappa_s)\|s_{k,\text{cp}}\|$, which need not be summable. Indeed, by the same reasoning as in the proof of Lemma 3.6, for any successful iteration k, there exists a Cauchy step $s_{k,\text{cp}}$ such that

$$\begin{aligned}(f+h)(x_k) - (f+h)(x_k+s_k) &\geq \tfrac{1}{2}\eta_1(1-\theta_1)\nu_k^{-1}\|\widehat{s}_{k,\text{cp}}\|^2 \\ &\geq \tfrac{1}{2}\eta_1(1-\theta_1)\nu_{\max}^{-1}\|\widehat{s}_{k,\text{cp}}\|^2 \\ &\geq \tfrac{1}{2}\eta_1(1-\theta_1)\nu_{\max}^{-1}\kappa_s^2\|s_{k,\text{cp}}\|^2.\end{aligned}$$

Therefore, if we sum those inequalities over the set $\mathcal{S}$ of all successful iterations and use Assumption 3.7, we obtain

$$(f+h)(x_0) - (f+h)_{\text{low}} \geq \tfrac{1}{2}\eta_1(1-\theta_1)\nu_{\max}^{-1}\kappa_s^2 \sum_{k\in\mathcal{S}} \|s_{k,\text{cp}}\|^2.$$

A similar inequality holds for $\widehat{s}_{k,\text{cp}}$. Thus, both $\{\widehat{s}_{k,\text{cp}}\}$ and $\{s_{k,\text{cp}}\}$ are square summable. However, showing that they are summable appears to require the stronger Kurdyka-Łojasiewicz assumption [15, Theorem 1], which is not used in our analysis.

iR2N naturally generalizes the special cases R2 [3] with $B(x) = 0$, R2DH [25] with $B(x)$ diagonal, and LM [6] when f is a squared residual norm and $B(x) = J(x)J(x)^T$, where $J(x)$ is the residual Jacobian. It stands to reason that the same mechanisms can be used to extend the trust-region variants (TR [3], TRDH [30], and LMTR [6]) to inexact evaluations and proximal operators with minimal modifications.

Numerical experiments confirm that iR2N provides substantial flexibility in contexts where exact evaluations are expensive or unavailable, and demonstrate that controlled inexactness can be leveraged to reduce computational cost without compromising convergence behavior.

In the context of trust-region methods for (1.1), Aravkin et al. [3, 6] give procedures based on the solution of a nonlinear equation to obtain an element of (2.4) with the additional constraint $\|s\|_\infty \leq \Delta$, where $\Delta > 0$, or, equivalently, with the additional

term $\chi(s \mid \Delta\mathbb{B}_\infty)$ in the objective, where $\mathbb{B}_\infty$ is the ℓ_∞-norm unit ball and χ is the indicator of a set. They do so for two choices of ψ. Our results apply directly to both regularizers, and indeed to any regularizer combined with a trust-region constraint. Here, $\mathbb{B}_2 \subset \mathbb{B}_\infty$, and hence, $\|s_{k,\text{cp}}\|_2 \leq \Delta$. Thus, we may use the stopping condition $\|\widehat{s}_{k,\text{cp}}\| \geq \kappa_s \Delta$.

Future work will focus on allowing inexact evaluations of the quadratic model (2.5), particularly regarding B_k, which itself may be computed inexactly—for instance, when represented in reduced numerical precision or when linear systems involving B_k are solved approximately.

REFERENCES

[1] N. Allaire and D. Orban. IRBP.jl: An efficient numerical algorithm for computing the Euclidean projection of a vector onto the nonconvex ℓ_p-"ball", September 2024.

[2] N. Allaire, D. Orban, and A. Montoison. ProxTV.jl: Wrapper for general total variation regularization, September 2024.

[3] A. Aravkin, R. Baraldi, and D. Orban. A proximal quasi-Newton trust-region method for nonsmooth regularized optimization. *SIAM J. Optim.*, (2):900–929, 2022.

[4] A. Aravkin, R. Baraldi, G. Leconte, and D. Orban. Corrigendum: A proximal quasi-Newton trust-region method for nonsmooth regularized optimization. Cahier G-2021-12-SM, GERAD, Montréal, QC, Canada, 2024.

[5] A. Y. Aravkin, R. Baraldi, and D. Orban. A proximal quasi-Newton trust-region method for nonsmooth regularized optimization. *arXiv*, (2103.15993v1), 2021. Preliminary report.

[6] A. Y. Aravkin, R. Baraldi, and D. Orban. A Levenberg-Marquardt method for nonsmooth regularized least squares. *SIAM J. Sci. Comput.*, 46(4):A2557–A2581, 2024.

[7] R. Baraldi, G. Leconte, and D. Orban. RegularizedOptimization.jl: Algorithms for regularized optimization, 2024.

[8] R. J. Baraldi and D. P. Kouri. A proximal trust-region method for nonsmooth optimization with inexact function and gradient evaluations. *Math. Program.*, 201(1):559–598, 2023.

[9] A. Barbero and S. Sra. Modular proximal optimization for multidimensional total-variation regularization. *J. Mach. Learn Res.*, 19(56):1–82, 2018.

[10] M. Barré, A. B. Taylor, and F. Bach. Principled analyses and design of first-order methods with inexact proximal operators. *Math. Program.*, 201(1):185–230, 2023.

[11] A. Beck. *First-Order Methods in Optimization*. SIAM, Philadelphia, USA, 2017.

[12] S. Becker, J. Fadili, and P. Ochs. On quasi-Newton forward-backward splitting: Proximal calculus and convergence. *SIAM J. Optim.*, 29(4):2445–2481, 2019.

[13] Y. Bello-Cruz, M. L. Gonçalves, and N. Krislock. On inexact accelerated proximal gradient methods with relative error rules. *arXiv*, (2005.03766), 2020.

[14] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A fresh approach to numerical computing. *SIAM Rev.*, 59(1):65–98, 2017.

[15] J. Bolte, S. Sabach, and M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Math. Program.*, (146):459–494, 2014.

[16] S. Bonettini, S. Rebegoldi, and V. Ruggiero. Inertial variable metric techniques for the inexact forward–backward algorithm. *SIAM J. Sci. Comput.*, 40(5):A3180–A3210, 2018.

[17] O. Burdakov, L. Gong, S. Zikrin, and Y.Yuan. On efficiently combining limited-memory and trust-region techniques. *Math. Program. Comp.*, 9(1):101–134, 2017.

[18] Z. Chen, P. Yu, and H. Huang. Zeroth-order methods for stochastic nonconvex nonsmooth composite optimization. *arXiv*, (2510.04446), 2025.

[19] G. Chierchia, E. Chouzenoux, P. Combettes, and J.-C. Pesquet. *The proximity operator repository*, 2020. http://proximity-operator.net/download/guide.pdf.

[20] L. Condat. A direct algorithm for 1-d total variation denoising. *IEEE Signal Process. Lett.*, 20 (11):1054–1057, 2013.

[21] L. Condat. Fast projection onto the simplex and the ℓ_1-ball. *Math. Program.*, 158(1):575–585, 2016.

[22] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust Region Methods*, volume 1 of *MPS-SIAM Series on Optimization*. SIAM, Philadelphia, USA, 2000.

[23] A. R. Conn, K. Scheinberg, and L. N. Vicente. *Introduction to Derivative-Free Optimization*. MPS-SIAM Series on Optimization. SIAM, 2009.

[24] Y. Dai and D. P. Robinson. Inexact proximal-gradient methods with support identification.

arXiv, (2211.02214), 2022.

[25] Y. Diouane, M. L. Habiboullah, and D. Orban. A proximal modified quasi-Newton method for nonsmooth regularized optimization. Cahier G-2024-64, GERAD, Montréal, QC, Canada, 2024. To appear in Math. Program.

[26] M. Fukushima and H. Mine. A generalized proximal point algorithm for certain non-convex minimization problems. *Int. J. Syst. Sci.*, 12:989–1000, 1981.

[27] B. Gu, D. Wang, Z. Huo, and H. Huang. Inexact proximal gradient methods for non-convex and non-smooth optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[28] P. D. Khanh, B. S. Mordukhovich, V. T. Phat, and D. B. Tran. Inexact proximal methods for weakly convex functions. *J. Global. Optim.*, 91(3):611–646, 2025.

[29] G. Leconte and D. Orban. Complexity of trust-region methods with unbounded Hessian approximations for smooth and nonsmooth optimization. Cahier G-2023-65, GERAD, Montréal, QC, Canada, 2023. To appear in Math. Program.

[30] G. Leconte and D. Orban. The indefinite proximal gradient method. *Comput. Optim. Appl.*, 91 (2):861–903, 2025.

[31] D. Monnet and D. Orban. A multi-precision quadratic regularization method for unconstrained optimization with rounding error analysis. *Comput. Optim. Appl.*, 91:997–1031, 2025.

[32] J.-J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. Fr.*, 93:273–299, 1965.

[33] C. Rackauckas and Q. Nie. DifferentialEquations.jl – A performant and feature-rich ecosystem for solving differential equations in Julia. *J. Open Source Softw.*, 5(1), 2017.

[34] H. Robbins and S. Monro. A stochastic approximation method. *Ann. Math. Stat.*, 22(3):400–407, 1951.

[35] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[36] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.*, 14(5):877–898, 1976.

[37] R. T. Rockafellar and R. Wets. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag, 2009.

[38] S. Salzo and S. Villa. Inexact and accelerated proximal point algorithms. *J. Convex Anal.*, 19 (4):1167–1192, 2012.

[39] M. Schmidt, N. Roux, and F. Bach. Convergence rates of inexact proximal-gradient methods for convex optimization. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011.

[40] G. D. Smith. *Numerical solution of partial differential equations: finite difference methods*. Number 1 in Oxford Applied Mathematics and Computing Science Series. Oxford University Press, Oxford, England, third edition, 1985.

[41] S. Sra. Scalable nonconvex inexact proximal splitting. In F. Pereira, C. Burges, L. Bottou, and K. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012.

[42] J. H. Verner. Numerically optimal Runge–Kutta pairs with interpolants. *Numer. Algor.*, 53 (2–3):383–396, 2010.

[43] X. Yang, J. Wang, and H. Wang. Towards an efficient approach for the nonconvex ℓ_p ball projection: algorithm and analysis. *J. Mach. Learn Res.*, 23(101):1–31, 2022.