Chapter 12

Efficient, Quantitative Numerical Methods for Statistical Image Deconvolution and Denoising

D. Russell Luke, C. Charitha, Ron Shefi and Yura Malitsky

Abstract We review the development of efficient numerical methods for statistical multi-resolution estimation of optical imaging experiments. In principle, this involves constrained linear deconvolution and denoising, and so these types of problems can be formulated as convex constrained, or even unconstrained, optimization. We address two main challenges: first of these is to quantify convergence of iterative algorithms; the second challenge is to develop efficient methods for these large-scale problems without sacrificing the quantification of convergence. We review the state of the art for these challenges.

2010 Mathematics Subject Classification: Primary 49J52 · 49M20 · 90C26 · Secondary 15A29 · 47H09 · 65K05 · 65K10 · 94A08.

12.1 Introduction

In this chapter we review progress towards addressing two main challenges in scientific image processing. The first of these is to quantify convergence of iterative algorithms for image processing to solutions (as opposed to optimal values) to the underlying variational problem. The second challenge is to develop efficient methods for these large-scale problems without sacrificing the quantification of convergence.

D. R. Luke (✉) · R. Shefi · Y. Malitsky
Institute for Numerical and Applied Mathematics, Universität Göttingen,
37083 Göttingen, Germany
e-mail: r.luke@math.uni-goettingen.de

R. Shefi
e-mail: ronshefi@gmail.com

Y. Malitsky
e-mail: y.malitskyi@math.uni-goettingen.de

C. Charitha
Indian Institute of Technology Indore, Indore, India
e-mail: charithac@iiti.ac.in

The techniques surveyed here were studied in [1–3]. We present only the main results from these studies, in the context that hindsight provides.

Scientific images are often processed with software that accomplishes a number of tasks like registration, denoising and deblurring. Implicit in the processing is that some systematic error is being corrected to bring the image closer to the truth. This presumption is more complicated for denoising and deblurring. These are often accomplished by filtering or by solving some variational problem such as minimizing the variance of an image. For applications requiring speedy processing, such as audio and video communication, this is sufficient. But the recent development of nanoscale photonic imaging modalities such as STED and RESOLFT featured in Chaps. 1, 7 and 9 has shifted the focus of image denoising and deconvolution from qualitative to quantitative models.

Quantitative approaches to image processing are the subject of Chap. 11 where statistical multiscale estimation is discussed (see Sect. 11.3). Here, the recovered image comes with statistical statements about how far the processed image is, in some statistical sense, from the truth. The estimators are almost exclusively variational, that is, they can be characterized as the solution to an optimization problem. It is important to emphasize that the value of the optimization problem is meaningless. This stands in stark contrast to many conventional applications in economics and operations research, where the value of the optimal solution is related to profit or cost, and so is of principal interest.

The focus on optimal solutions rather than optimal values places heavy demands on the structure of model formulations and the algorithms for solving them. Unless the numerical method allows one to state how far a computed iterate is to the solution of the underlying variational problem, then the scientific significance of the iterate is lost.

The leading computational approaches for solving imaging problems with multi-resolution statistical estimation criterion are based on iterated proximal operators. Most of the analysis for first-order iterative proximal methods is limited to statements about rates of convergence of function values, if rates are discussed at all (see for instance [4–7]). First-order methods have slow convergence in the worst case scenario. A common assumption to guarantee linear convergence of the iterates is strong convexity, but this is far more than is necessary, and in particular it is not satisfied for the Huber function (12.35). It was shown in [8] that metric subregularity is necessary for local linear convergence. Aspelmeier, Charitha and Luke [1] showed that the popular alternating directions method of multipliers algorithm (ADMM) applied to optimization problems with piecewise linear-quadratic objective functions (e.g. the Huber function), together with linear inequality constraints generically satisfies metric subregularity at isolated critical points; hence linear convergence of the iterates for this algorithm can be expected without further ado. More recently, in [3] it was shown that the primal iterates of a modification of the PAPC algorithm (Algorithm 2) converge R-linearly for any *quadratically supportable* objective function (for instance, the Huber function). Conventional results without metric subregularity obtain a convergence rate of $O(1/k)$ with respect to the function values. In settings like qualitative image processing or machine learning, such results are acceptable,

but in the setting of statistical image processing these statements do not contain any scientific content. We present in this chapter efficient iterative first-order methods that offer some hope of quantitative guarantees about the distance of the iterates to optimal solutions.

12.2 Problem Formulation

We limit our scope to the real vector space $\mathbb{R}^n$ with the norm generated from the inner product. The closed unit ball centered on the point $y \in \mathbb{R}^n$ is denoted by $\mathbb{B}(y)$. The positive orthant (resp. negative orthant) in $\mathbb{R}^n$ is denoted by $\mathbb{R}_+^n$ (resp. $\mathbb{R}_-^n$). The domain of an extended real-valued function $\varphi : \mathbb{R}^n \rightarrow \overline{\mathbb{R}} \equiv (-\infty, +\infty]$ is $\text{dom}(\varphi) \equiv \{z \in \mathbb{R}^n : \varphi(z) < +\infty\}$. The *Fenchel conjugate* of φ is denoted by φ^* and is defined by $\varphi^*(u) \equiv \sup_{z \in \mathbb{R}^n} \{\langle z, u \rangle - \varphi(z)\}$. The set of symmetric $n \times n$ positive (semi)-definite matrices is denoted by $\mathbb{S}_{++}^n$ ($\mathbb{S}_+^n$). The notation $A \succ 0$ ($A \succeq 0$) denotes a positive (semi)definite matrix A . For any $z \in \mathbb{R}^n$ and any $A \in \mathbb{S}_+^n$, we denote the semi-norm $\|z\|_A^2 := \langle z, Az \rangle$. The operator norm is defined by $\|A\| = \max_{u \in \mathbb{R}^n} \{\|Au\| : \|u\| = 1\}$ and coincides with the spectral radius of A whenever A is symmetric. If $A \neq 0$, $\sigma_{\min}(A)$ denotes its smallest nonzero singular value. For a sequence $\{z^k\}_{k \in \mathbb{N}}$ converging to z^* , we say the convergence is Q -linear if there exists $c \in (0, 1)$ such that $\frac{\|z^{k+1} - z^*\|}{\|z^k - z^*\|} \leq c$ for all k ; convergence is R -linear if there exists a sequence η_k such that $\|z^k - z^*\| \leq \eta_k$ and $\eta_k \rightarrow 0$ Q -linearly [9, Chap. 9].

We limit our discussion to proper (nowhere equal to $-\infty$ and finite at some point), lower semi-continuous (lsc), extended-valued (can take the value $+\infty$) functions. We will, in fact, limit our discussion to *convex* functions, but convexity is not the central property governing quantitative convergence estimates. By the *subdifferential* of a function φ , denoted $\partial\varphi$, we mean the collection of all *subgradients* that can be written as limits of sequences of *Fréchet subgradients* at nearby points; a vector v is a (*Fréchet*) *subgradient* of φ at y , written $v \in \widehat{\partial}\varphi(y)$, if

$$\liminf_{x \rightarrow y, x \neq y} \frac{\varphi(x) - \varphi(y) - \langle v, x - y \rangle}{\|x - y\|} \geq 0. \quad (12.1)$$

The functions of interest for us are *subdifferentially regular* on their domains, that is, the epigraphs of the functions are *Clarke regular* at points where they are finite [10, Definition 7.25]. For our purposes it suffices to note that, for a function φ that is subdifferentially regular at a point y , the subdifferential is nonempty and all subgradients are Fréchet subgradients, that is, $\partial\varphi(y) = \widehat{\partial}\varphi(y) \neq \emptyset$. Convex functions, in particular, are subdifferentially regular on their domains and the subdifferential has the particularly simple representation as the set of all vectors v where

$$\varphi(x) - \varphi(y) - \langle v, x - y \rangle \geq 0 \quad \forall x. \quad (12.2)$$

A mapping $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be β -inverse strongly monotone [10, Corollary 12.55] if for all $x, x' \in \mathbb{R}^n$

$$\langle v - v', x - x' \rangle \geq \beta \|v - v'\|^2, \quad \text{whenever } v \in \Phi(x), v' \in \Phi(x'). \quad (12.3)$$

The mapping Φ is said to be *polyhedral* (or piecewise polyhedral [10]) if its graph is the union of finitely many sets that are polyhedral convex in $\mathbb{R}^n \times \mathbb{R}^n$ [11]. Polyhedral mappings are generated by the subdifferential of piecewise linear-quadratic functions (see Proposition 12.9).

Definition 12.1 (*piecewise linear-quadratic (plq) functions*) A function $f : \mathbb{R}^n \rightarrow [-\infty, +\infty]$ is called *piecewise linear-quadratic* if $\text{dom} f$ can be represented as the union of finitely many polyhedral sets, relative to each of which $f(x)$ is given by an expression of the form $\frac{1}{2}\langle x, Ax \rangle + \langle a, x \rangle + \alpha$ for some scalar $\alpha \in \mathbb{R}$ vector $a \in \mathbb{R}^n$, and symmetric matrix $A \in \mathbb{R}^{n \times n}$.

Closely related to plq functions is quadratically supportable functions.

Definition 12.2 (*pointwise quadratically supportable (pqs)*) A proper, extended-valued function $\varphi : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ is said to be *pointwise quadratically supportable* at y if it is subdifferentially regular there and there exists a neighborhood V of y and a constant $\mu > 0$ such that

$$(\forall v \in \partial\varphi(y)) \quad \varphi(x) \geq \varphi(y) + \langle v, x - y \rangle + \frac{\mu}{2} \|x - y\|^2, \quad \forall x \in V. \quad (12.4)$$

If for each bounded neighborhood V of y there exists a constant $\mu > 0$ such that (12.4) holds, then the function φ is said to be *pointwise quadratically supportable at y on bounded sets*. If (12.4) holds with one and the same constant $\mu > 0$ on all neighborhoods V , then φ is said to be *uniformly pointwise quadratically supportable* at y .

For more on the relationship between pointwise quadratic supportability, coercivity, strong monotonicity and strong convexity see [3].

We denote the *resolvent* of Φ by $\mathcal{J}_\Phi \equiv (\text{Id} + \Phi)^{-1}$ where Id denotes the identity mapping and the inverse is defined as

$$\Phi^{-1}(y) \equiv \{x \in \mathbb{R}^n \mid y \in \Phi(x)\}. \quad (12.5)$$

The corresponding *reflector* is defined by $R_{\eta\Phi} \equiv 2\mathcal{J}_{\eta\Phi} - \text{Id}$. One of the more prevalent examples of resolvents is the *proximal map*. For $\varphi : \mathbb{R}^n \rightarrow (-\infty, \infty]$ a proper, lsc and convex function and for any $u \in \mathbb{R}^n$ and $Q \in \mathbb{S}_{++}^n$, the proximal map associated with φ with respect to the weighted Euclidean norm is uniquely defined by:

$$\text{prox}_{Q, \varphi}(u) = \underset{z}{\operatorname{argmin}} \{ \varphi(z) + \frac{1}{2} \|z - u\|_Q^2 : z \in \mathbb{R}^n \}.$$

When $Q = c^{-1}\text{Id}$, $c > 0$, we simply use the notation $\text{prox}_{c,\varphi}(u)$. We also recall the fundamental Moreau proximal identity [12], that is, for any $z \in \mathbb{R}^n$

$$z = \text{prox}_{Q,\varphi}(z) + Q\text{prox}_{Q^{-1},\varphi^*}(Q^{-1}(z)), \quad (12.6)$$

where Q^{-1} is the inverse of $Q \in \mathbb{S}_{++}^n$.

Notions of *continuity* of set-valued mappings have been thoroughly developed over the last 40 years. Readers are referred to the monographs [10, 11, 13] for basic results. A key property of set-valued mappings that we will rely on is *metric subregularity*, which can be understood as the property corresponding to a Lipschitz-like continuity of the inverse mapping relative to a specific point. It is a weaker property than *metric regularity* which, in the case of an $n \times m$ matrix for instance, is equivalent to surjectivity. Our definition follows the characterization of this property given in [11, Exercise 3H.4].

Definition 12.3 (*metric subregularity*) The mapping $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is called *metrically subregular at $\bar{x}$ for $\bar{y}$ relative to $W \subset \mathbb{R}^n$* if $(\bar{x}, \bar{y}) \in \text{gph}\Phi$ and there is a constant $c > 0$ and neighborhoods $\mathcal{O}$ of $\bar{x}$ such that

$$\text{dist}(x, \Phi^{-1}(\bar{y}) \cap W) \leq c \text{dist}(\bar{y}, \Phi(x)) \quad \forall x \in \mathcal{O} \cap W. \quad (12.7)$$

The constant c measures the stability under perturbations of inclusion $\bar{y} \in \Phi(\bar{x})$. An important instance where metric subregularity comes for free is for polyhedral mappings.

Proposition 12.4 (polyhedrality implies metric subregularity) *Let $W \subset V$ be an affine subspace and $T : W \rightrightarrows W$. If T is polyhedral and $\text{Fix } T \cap W$ is an isolated point, $\{\bar{x}\}$, then $T - \text{Id} : W \rightrightarrows (W - \bar{x})$ is metrically subregular at $\bar{x}$ for 0 relative to W .*

Proof Polyhedrality and isolated fixed points in fact imply *strong* metric subregularity. See [11, Propositions 3I.1 and 3I.2].

A notion related to metric regularity is that of *weak-sharp solutions*. This will be used in the development of error bounds (Theorem 12.6).

Definition 12.5 (*weak sharp minimum* [14]) The solution set $\text{argmin} \{f(x) \mid x \in \Omega\}$ for a nonempty closed convex set Ω , is weakly sharp if, for $\bar{p} = \inf_{\Omega} f$, there exists a positive number α (sharpness constant) such that

$$f(x) \geq \bar{p} + \alpha \text{dist}(x, S_f) \quad \forall x \in \Omega.$$

Similarly, the solution set S_f is weakly sharp of order $\nu > 0$ if there exists a positive number α (sharpness constant) such that, for each $x \in \Omega$,

$$f(x) \geq \bar{p} + \alpha \text{dist}(x, S_f)^\nu \quad \forall x \in \Omega.$$

12.2.1 Abstract Problem

The generic problem in which we are interested is

$$\begin{aligned} & \underset{x \in \mathbb{R}^n}{\text{minimize}} && f(x) \\ & \text{subject to} && g_i(A_i x) \leq 0 \quad (i = 1, \dots, M) \end{aligned} \quad (\mathcal{P}_0)$$

The following blanket assumptions on the problem's data hold throughout:

Assumption 1

- (i) The set of optimal solutions for problem $(\mathcal{P}_0)$, denoted $\mathcal{S}^*$, is nonempty.
- (ii) the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is proper lsc and convex and coercive, and for $i = 1, \dots, M$ the functions $g_i : \mathbb{R}^{m_i} \rightarrow (-\infty, +\infty]$ are proper, lsc, and convex;
- (iii) the mappings $A_i : \mathbb{R}^n \rightarrow \mathbb{R}^{m_i}$, $i = 1, \dots, M$ ($m_i \leq n$) are linear and full rank, that is, $\sigma_{\min}^2(A_i) = \lambda_{\min}(A_i A_i^T) > 0$.

Assumption (i) implies that the optimal value of $(\mathcal{P}_0)$ is finite. Assumption (ii) implies that the constraint structure is convex. Assumption (iii) implies that the mapping $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is linear and full rank, where

$$\mathcal{A} \equiv [A_0^T, A_1^T, \dots, A_M^T]^T \in \mathbb{R}^m \times \mathbb{R}^n$$

so that $\mathcal{A}x = y$ where $y = (y_1, \dots, y_M) \in \mathbb{R}^m$ for $m = \sum_{i=1}^M m_i$. The challenge of statistical multi-resolution estimation lies in the feature that the dimension of the constraint structure, m , is much greater than the dimension of the unknowns, n , and grows superlinearly with respect to the number of unknowns.

The above constrained optimization problem is often formulated as an unconstrained-looking problem via the introduction of a (nonsmooth) penalty term enforcing the constraints:

$$\min_{x \in \mathbb{R}^n} f(x) + g(\mathcal{A}x) \quad (\mathcal{P})$$

where

$$g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}} \equiv \rho\theta \quad (12.8)$$

for ρ a positive scalar and

$$\theta : \mathbb{R}^m \rightarrow [0, +\infty] \text{ proper, lsc and convex with } \theta(y) = 0 \text{ if and only if } y \in \mathcal{C} \quad (12.9)$$

where

$$\mathcal{C} \equiv \{y = (y_1, y_2, \dots, y_M) \mid y \in \text{range } \mathcal{A} \text{ and } g_i(y_i) \leq 0, i = 1, 2, \dots, M\}. \quad (12.10)$$

The requirements on the function θ align this penalty term with *exact* penalization [15], that is, a relaxation of the constraints where, for all parameters ρ large enough, the constraints are exactly satisfied.

The following assumptions are used to guarantee the exact correspondence between solutions to $(\mathcal{P}_0)$ and $(\mathcal{P})$.

Assumption 2

- (i) The set $S_\infty \equiv \operatorname{argmin} \{ \theta(Ax) \mid x \in \mathbb{R}^n \}$ is nonempty and weakly sharp (Definition 12.5) of order $\nu \geq 1$.
- (ii) The lower level set $\operatorname{lev}_{\leq \alpha} f$ is bounded for each $\alpha \in \mathbb{R}$ and $\inf_{x \in \mathbb{R}^n} f > -\infty$.

Proposition 12.6 *Suppose Assumption 2 holds. Then the set of solutions to $(\mathcal{P})$ with g defined by (12.8) is bounded and for all ρ large enough, the solutions to $(\mathcal{P}_0)$ and $(\mathcal{P})$ coincide.*

Proof This is a distillation of [1, Theorem 3.4].

In $(\mathcal{P}_0)$ and $(\mathcal{P})$ the function f is often smooth, but not prox friendly. In applications it is most often a smooth regularization or a fidelity term. For the ADMM/DR method reviewed in Sect. 12.3 smoothness is not required.

It is assumed that the functions g_i ($i = 1, 2, \dots, M$) are prox friendly and that they enjoy some structure that makes g also prox friendly. For instance, if the constraints are *separable*, then the function

$$g(Ax) \equiv \rho \sum_{i=1}^M g_i(A_i x) \quad (12.11)$$

is also prox-friendly as is the function

$$g(Ax) \equiv \rho \max\{g_1(A_1 x), g_2(A_2 x), \dots, g_M(A_M x), 0\}. \quad (12.12)$$

The functions $g_i \circ A_i$ can be regularizing functions (like total variation) or hard inequality constraints. For example, hard inequality constraints are modeled by the use of indicator functions for g_i in $(\mathcal{P}_0)$: $g_i(A_i x) = \iota_{\gamma_i \mathbb{B}}(A_i x - b_i)$ where, for a subset $\Omega \subset \mathbb{R}^n$,

$$\iota_\Omega(x) \equiv \begin{cases} 0 & x \in \Omega \\ +\infty & \text{else.} \end{cases}$$

12.2.2 Saddle Point and Dual Formulations

The saddle point formulation is derived by viewing the function g in $(\mathcal{P})$ as the *image* of a function g^* under Fenchel conjugation, that is, $g(x) = (g^*)^*$. Writing this explicitly into $(\mathcal{P})$ yields

$$\min_{x \in \mathbb{R}^n} \left\{ \max_{y \in \mathbb{R}^m} f(x) + \langle \mathcal{A}x, y \rangle - g^*(y) \right\}. \quad (\mathcal{S})$$

The bifunction in the saddle point formulation is

$$L(x, y) := f(x) + \langle \mathcal{A}x, y \rangle - g^*(y). \quad (12.13)$$

Contrast this with the Lagrangian for the extended problem

$$\underset{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m}{\text{minimize}} \quad f(x) + g(y), \quad \text{subject to } \mathcal{A}x = y. \quad (\mathcal{P}_{\mathcal{L}})$$

The Lagrangian is

$$\mathcal{L}(x, y, z) = f(x) + g(y) + \langle z, \mathcal{A}x - y \rangle, \quad (12.14)$$

and the augmented Lagrangian $\tilde{\mathcal{L}}$ for $(\mathcal{P}_{\mathcal{L}})$ is given by

$$\tilde{\mathcal{L}}(x, y, z) = f(x) + g(y) + \langle z, \mathcal{A}x - y \rangle + \frac{\eta}{2} \|\mathcal{A}x - y\|^2, \quad (12.15)$$

where $z \in \mathbb{R}^m$, and $\eta > 0$ is a fixed penalty parameter.

Assumption 1(i) guarantees that the mapping $L(\cdot, \cdot)$ has a saddle point, that is, there exists $(\hat{x}, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^m$ such that

$$L(\hat{x}, y) \leq L(\hat{x}, \hat{y}) \leq L(x, \hat{y}) \quad \forall x \in \mathbb{R}^n, y \in \mathbb{R}^m.$$

The existence of a saddle point corresponds to zero duality gap for the induced optimization problems

$$p(x) = \sup_y \{L(x, y) : y \in \mathbb{R}^m\} \quad q(y) = \inf_x \{L(x, y) : x \in \mathbb{R}^n\}.$$

By weak duality, we have $\inf_{x \in \mathbb{R}^n} p(x) \geq \sup_{y \in \mathbb{R}^m} q(y)$.

This can be viewed as a *partial dual* to problem $(\mathcal{P})$. The full dual problem involves the Fenchel conjugate of the entire objective function. For $(\mathcal{P})$ the dual problem is

$$\sup_{y \in \mathbb{R}^m} -f^*(\mathcal{A}^T y) - g^*(-y). \quad (\mathcal{D}')$$

Instead of working with this dual, it is more convenient to work with the following equivalent formulation via the change of variable $y \rightarrow -y$:

$$\inf_{y \in \mathbb{R}^m} f^*(-\mathcal{A}^T y) + g^*(y). \quad (\mathcal{D})$$

Under standard constraint qualifications (e.g., [16, Theorem 2.3.4]), $(\hat{x}, \hat{y})$ is a saddle point of L if and only if $\hat{x}$ is an optimal solution of the primal problem $(\mathcal{P}_0)$, and $\hat{y}$ is an

optimal solution of the dual problem $(\mathcal{D})$. The following two inclusions characterize the solutions of the problems $(\mathcal{P}_0)$ and $(\mathcal{D})$ respectively:

$$\begin{aligned} 0 &\in \partial f(\bar{x}) + \partial(g \circ \mathcal{A})(\bar{x}); \\ 0 &\in \partial(f^* \circ (-\mathcal{A}^T))(\bar{y}) + \partial g^*(\bar{y}). \end{aligned}$$

In both cases, one has to solve an inclusion of the form

$$0 \in (B + D)(x), \quad (12.16)$$

for general set-valued mappings B and D .

12.2.3 Statistical Multi-resolution Estimation

Statistical multi-resolution estimation (SMRE) discussed in Sect. 11.2.7 of Chap. 11 is specialized here for the case of imaging systems with Gaussian noise. Let $\mathcal{G} = \{1, \dots, N\} \times \{1, \dots, N\}$ denote a grid of $N^2 = n$ points. Denote by $\mathcal{V}$ a collection of subsets of $\mathcal{G}$. Suppose there are M such subsets, that is $M = |\mathcal{V}|$, each of these subsets V_i consisting of $m_i \leq n$ grid points with $\sum_{i=1}^M m_i = m \geq n$.

The variational model for statistical multi-resolution estimation with Gaussian noise takes the form

$$\begin{aligned} &\underset{x \in \mathbb{R}^n}{\text{minimize}} && f(x) \\ &\text{subject to} && \left| \sum_{j \in V_i} w_i(j) ((Ax)_j - b_j) \right| \leq \gamma_i \quad \forall i = 1, 2, \dots, M. \end{aligned} \quad (\mathcal{P}_{SMRE})$$

Here $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a regularization functional, which incorporates a priori knowledge about the unknown signal $\hat{x}$ such as smoothness, w_i is a weighting function for the grid points in the subset V_i , and $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the linear imaging operator that models the experiment. The constant γ_i has an interpretation in terms of the *quantile* of the estimator.

In the context of the general model $(\mathcal{P}_0)$,

$$g_i(A_i x) = \left| \sum_{j \in V_i} w_i(j) ((Ax)_j - b_j) \right| - \gamma_i, \quad i = 1, 2, \dots, M. \quad (12.17)$$

Here the affine mapping $A_i x \equiv \sum_{j \in V_i} w_i(j) ((Ax)_j - b_j)$ is an averaging operator that accounts for sampling at different resolutions of the image. Note that the observation b need not be in the range of the imaging operator A - all that is assumed is that this mapping is injective, not surjective. This means that, in applications, practitioners need to be careful not to make the constraint γ_i too small, otherwise the

optimization problem might be *infeasible*. If the algorithms presented below appear to be diverging for a particular instance of $(\mathcal{P}_{SMRE})$, it is because the problem is infeasible; increasing the constants γ_i should solve the problem.

12.3 Alternating Directions Method of Multipliers and Douglas Rachford

In this section we survey the main results (without proofs) from [1]. For proofs of the statements, readers are referred to that article. A starting point for most of the main approaches to solving $(\mathcal{P}_0)$ is the alternating directions method of multipliers (ADMM) (primary sources include [17–21]). This method is one of many *splitting methods* which are the principal approach to handling the computational burden of large-scale, separable problems [22]. ADMM belongs to a class of *augmented Lagrangian methods* whose original motivation was to regularize Lagrangian formulations of constrained optimization problems. The ADMM algorithm for solving $(\mathcal{P}_{\mathcal{L}})$ follows. The penalty parameter η need not be a constant, and indeed evidence indicates that the choice of η can greatly impact the complexity of the algorithm. For simplicity we keep this parameter fixed.

Algorithm 1: ADMM for $(\mathcal{P}_{\mathcal{L}})$ as in [1, Algorithm 2.1].

Parameters : Set $\eta > 0$

Initialization: Set $(x^0, y^0, z^0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$

1 for $k = 1, 2, \dots$ **do**

2 $x^{k+1} \in \operatorname{argmin}_x \{ f(x) + \langle z^k, \mathcal{A}x \rangle + \frac{\eta}{2} \|\mathcal{A}x - y^k\|^2 \};$

3 $y^{k+1} \in \operatorname{argmin}_y \{ g(y) - \langle z^k, y \rangle + \frac{\eta}{2} \|\mathcal{A}x^{k+1} - y\|^2 \};$

4 $z^{k+1} = z^k + \eta(\mathcal{A}x^{k+1} - y^{k+1}).$

We do not specify how the argmin in Algorithm 1 should be calculated, and indeed, the analysis that follows assumes that these can be computed exactly.¹ One problem that should be immediately apparent is that this algorithm operates on a space of dimension $n + 2m$. Since one of the two challenges we address is high dimension, this expansion in the dimension of the problem formulation should be troubling. Nevertheless, we show with this algorithm how the first challenge, namely *quantification of convergence* is achieved.

The connection between the ADMM algorithm and the Douglas–Rachford algorithm introduced in Chap. 6, (6.30) was first discovered by Gabay [19] (see also the thesis of Eckstein [17]). For any $\eta > 0$, the Douglas–Rachford algorithm [23, 24] for solving (12.16) is given by

¹This is not true in practice and remains an unresolved issue in numerical variational analysis.

$$z^{k+1} \in T' z^k \quad (k \in \mathbb{N}), \quad (12.18)$$

$$\text{for } T' \equiv \mathcal{J}_{\eta D} (\mathcal{J}_{\eta B} (\text{Id} - \eta D) + \eta D), \quad (12.19)$$

where $\mathcal{J}_{\eta D} \equiv (\text{Id} + \eta D)^{-1}$ and $\mathcal{J}_{\eta B} (\text{Id} + \eta B)^{-1}$ are the *resolvents* of ηD and ηB respectively.

Given z^0 and $y^0 \in Dz^0$, following [25], define the new variable $\xi^0 \equiv z^0 + \eta y^0$ so that $z^0 = \mathcal{J}_{\eta D} \xi^0$. We thus arrive at an alternative formulation of the Douglas–Rachford algorithm (12.18):

$$\xi^{k+1} \in T \xi^k \quad (k \in \mathbb{N}), \quad (12.20)$$

$$\text{for } T \equiv \frac{1}{2} (R_{\eta B} R_{\eta D} + \text{Id}) = \mathcal{J}_{\eta B} (2\mathcal{J}_{\eta D} - \text{Id}) + (\text{Id} - \mathcal{J}_{\eta D}), \quad (12.21)$$

where $R_{\eta D}$ and $R_{\eta B}$ are the reflectors of the respective resolvents. This is the form of Douglas–Rachford considered in [26].

Specializing this to our application yields

$$B \equiv \partial (f^* \circ (-\mathcal{A}^T)) \quad \text{and} \quad D \equiv \partial g^*, \quad (12.22)$$

and so the resolvent mappings are the proximal mappings of the convex functions $(f^* \circ (-\mathcal{A}^T))$ and g^* respectively, and hence the resolvent mappings and corresponding fixed point operator T are single-valued [12].

Proposition 12.7 (Proposition 2.2 [1]) *Let $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be proper, lsc and convex. Let $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be linear and suppose there exists a solution to $0 \in (B + D)(x)$ for B and D defined by (12.22). For fixed $\eta > 0$, given any initial points ξ^0 and $(y^0, z^0) \in \text{gph } D$ such that $\xi^0 = y^0 + \eta z^0$, the sequences $(z^k)_{k \in \mathbb{N}}$, $(\xi^k)_{k \in \mathbb{N}}$ and $(y^k)_{k \in \mathbb{N}}$ defined respectively by (12.18), (12.20) and $y^k \equiv \frac{1}{\eta} (\xi^k - z^k)$ converge to points $\bar{z} \in \text{Fix } T'$, $\bar{\xi} \in \text{Fix } T$ and $\bar{y} \in D(\text{Fix } T')$. The point $\bar{z} = \mathcal{J}_{\eta D} \bar{\xi}$ is a solution to $(\mathcal{D})$, and $\bar{y} = \frac{1}{\eta} (\bar{\xi} - \bar{z}) \in D\bar{z}$. If, in addition, $\mathcal{A}$ has full column rank, then the sequence $(y^k, z^k)_{k \in \mathbb{N}}$ corresponds exactly to the sequence of points generated in steps 2–3 of Algorithm 1 and the sequence $(\xi^{k+1})_{k \in \mathbb{N}}$ converges to $\bar{\xi}$, a solution to $(\mathcal{P}_0)$.*

The correspondence between Douglas–Rachford and ADMM in the proposition above means that if quantitative convergence can be established for one of the algorithms, it is automatically established for the other. Linear convergence of Douglas–Rachford under the assumption of strong convexity and Lipschitz continuity of f was already established by Lions and Mercier [26]. Recent published work in this direction includes [27, 28]. Local linear convergence of the iterates to a *solution* was established in [29] for linear and quadratic programs using spectral analysis. In Proposition 12.8, two conditions are given that guarantee linear convergence of the ADMM iterates to a solution. The first condition is classical and follows Lions and Mercier [26]. The second condition, based on [30], is much more prevalent in applications and generalizes the results of [29].

Proposition 12.8 (Theorem 2.3 of [1]) *Let $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be proper, lsc and convex. Suppose there exists a solution to $0 \in (B + D)(x)$ for B and D defined by (12.22) where $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is an injective linear mapping. Let $\widehat{\xi} \in \text{Fix } T$ for T defined by (12.21). For fixed $\eta > 0$ and any given triplet of points (ξ^0, y^0, z^0) satisfying $\xi^0 \equiv z^0 + \eta y^0$, with $y^0 \in Dz^0$, generate the sequence $(y^k, z^k)_{k \in \mathbb{N}}$ by Steps 2–3 of Algorithm 1 and the sequence $(\xi^k)_{k \in \mathbb{N}}$ by (12.20).*

- (i) *Let $\mathcal{O} \subset \mathbb{R}^n$ be a neighborhood of $\widehat{\xi}$ on which g is strongly convex with constant μ and ∂g is β -inverse strongly monotone for some $\beta > 0$. Then, for any $(\xi^0, y^0, z^0) \in \mathcal{O}$ satisfying $\xi^0 \equiv z^0 + \eta y^0 \in \mathcal{O}$, the sequences $(\xi^k)_{k \in \mathbb{N}}$ and $(y^k, z^k)_{k \in \mathbb{N}}$ converge linearly to the respective points $\bar{\xi} \in \text{Fix } T$ and $(\bar{y}, \bar{z})$ with rate at least $K = \left(1 - \frac{2\eta\beta\mu^2}{(\mu+\eta)^2}\right)^{\frac{1}{2}} < 1$.*
- (ii) *Suppose that $T : W \rightarrow W$ for some affine subspace $W \subset \mathbb{R}^n$ with $\widehat{\xi} \in W$. On the neighborhood $\mathcal{O}$ of $\widehat{\xi}$ relative to W , that is $\mathcal{O} \cap W$, suppose there is a constant $\kappa > 0$ such that*

$$\|\xi - \xi^+\| \geq \sqrt{\kappa} \text{dist}(\xi, \text{Fix } T) \quad \forall \xi \in \mathcal{O} \cap W, \quad \forall \xi^+ \in T\xi. \quad (12.23)$$

Then the sequences $(\xi^k)_{k \in \mathbb{N}}$ and $(y^k, z^k)_{k \in \mathbb{N}}$ converge linearly to the respective points $\bar{\xi} \in \text{Fix } T \cap W$ and $(\bar{y}, \bar{z})$ with rate bounded above by $\sqrt{1 - \kappa}$.

In either case, the limit point $\bar{z} = \mathcal{J}_{\eta D} \bar{\xi}$ is a solution to $(\mathcal{D})$, $\bar{y} \in D\bar{z}$ and the sequence $(x^k)_{k \in \mathbb{N}}$ of Step 1 of Algorithm 1 converges to $\bar{x}$, a solution of $(\mathcal{P}_0)$.

The strong convexity assumption (i) of Theorem 12.8 fails in many applications of interest, and in particular for feasibility problems (minimizing the sum of indicator functions). By [31, Theorem 2.2], case (ii) of Theorem 12.8, in contrast, holds in general for mappings T for which $T - \text{Id}$ is *metrically subregular* and the fixed point sets are *isolated points* with respect to an affine subspace to which the iterates are confined. The restriction to the affine subspace W is a natural generalization for the Douglas–Rachford algorithm, where the iterates are known to stay confined to affine subspaces orthogonal to the fixed point set [32, 33]. We show that metric subregularity with respect to this affine subspace holds in many applications.

Proposition 12.9 (polyhedrality of the Douglas–Rachford operator) *Let $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be proper, lsc and convex. Suppose, in addition, that f and g are piecewise linear-quadratic. The operator $T : \mathbb{R}^m \rightarrow \mathbb{R}^m$ defined by (12.21) with $\eta > 0$ fixed, is polyhedral for B and D given by (12.22) where $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear mapping.*

Proof This is Proposition 2.6 of [1].

Proposition 12.10 (local linear convergence, Theorem 2.7 of [1]) *Let $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ and $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ be proper, lsc, convex, piecewise linear-quadratic functions (see Definition 12.1). Define the operator $T : \mathbb{R}^m \rightarrow \mathbb{R}^m$ by (12.21) with $\eta > 0$ fixed and B and D given by (12.22) where $\mathcal{A} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear mapping. Suppose that*

there exists a solution to $0 \in (B + D)(x)$, that $T : W \rightarrow W$ for W some affine subspace of $\mathbb{R}^m$ and that $\text{Fix } T \cap W$ is an isolated point $\{\bar{\xi}\}$. Then there is a neighborhood $\mathcal{O}$ of $\bar{\xi}$ such that, for all starting points (ξ^0, y^0, z^0) with $\xi^0 \equiv z^0 + \eta y^0 \in \mathcal{O} \cap W$ for $y^0 \in D(z^0)$ so that $\mathcal{J}_{\eta D} \xi^0 = z^0$, the sequence $(\xi^k)_{k \in \mathbb{N}}$ generated by (12.20) converges Q -linearly to $\bar{\xi}$ where $\bar{z} \equiv \mathcal{J}_{\eta D} \bar{\xi}$ is a solution to $(\mathcal{D})$. The rate of linear convergence is bounded above by $\sqrt{1 - \kappa}$, where $\kappa = c^{-2} > 0$, for c a constant of metric subregularity of $T - \text{Id}$ at $\bar{\xi}$ for the neighborhood $\mathcal{O}$. Moreover, the sequence $(y^k, z^k)_{k \in \mathbb{N}}$ generated by steps 2–3 of Algorithm 1 converges linearly to $(\bar{y}, \bar{z})$ with $\bar{y} = \frac{1}{\eta} (\bar{x} - \bar{z})$, and the sequence $(x^k)_{k \in \mathbb{N}}$ defined by Step 1 of Algorithm 1 converges to a solution to $(\mathcal{P}_0)$.

12.3.1 ADMM for Statistcal Multi-resolution Estimation of STED Images

The theoretical results above are demonstrated with an image $b \in \mathbb{R}^n$ (Fig. 12.1) generated from a Stimulated Emission Depletion (STED) microscopy experiment [34, 35] conducted at the Laser-Laboratorium Göttingen examining tubulin, represented as the “object” $x \in \mathbb{R}^n$. The imaging model is simple linear convolution. The measurement b , shown in Fig. 12.1, is *noisy* or otherwise inexact, and thus an exact solution is not desirable. Although the noise in such images is usually modeled by

Fig. 12.1 Recorded STED image used in [1]. Inset: area to be processed ($640 \times 640 \text{ nm}^2$). The scale bar (white) is $1 \mu\text{m}$

Fig. 12.2 **a** Numerical reconstruction via Algorithm 1 from the imaging data shown in Fig. 12.1 for $\rho = 4096$. **b** Iterates of Algorithm 1: primal and dual step sizes, as well as constraint violation $g(\mathcal{A}x)$ given by (12.12) with $\rho = 4096$. Reproduced from [1]

Poisson noise, a Gaussian noise model with constant variance suffices as the photon counts are of the order of 100 per pixel and do not vary significantly across the image. We calculate the numerically reconstructed tubulin density $\bar{x}$ shown in Fig. 12.2a via Algorithm 1 for the problem $(\mathcal{P})$ with the qualitative regularization

$$f(x) \equiv \alpha \|x\|^2 \quad (12.24)$$

and exact penalty $g(\mathcal{A}x)$ given by (12.12) with g_i given by (12.17).

For an image size of $n = 64 \times 64$ with three resolution levels the resulting number of constraints is $m = 12161$ (that is, 64^2 constraints at the finest resolution, $4 * 32^2$ constraints at the next resolution and $9 * 21^2$ constraints at the lowest resolution). The constant $\alpha = 0.01$ in (12.24) is used to balance the contributions of the individual terms to make the most of limited numerical accuracy (double precision). The constant γ_i is chosen so that the model solution would be no more than 3 standard deviations from the noisy data on each interval of each scale.

Since this is experimental data, there is no “truth” for comparison - the constraint, together with the error bounds on the numerical solution to the model solution provide statistical guarantees on the numerical reconstruction [36]. In Fig. 12.2b the iteration is shown with the value of $\rho = 4096$ for which the constraints are exactly satisfied (to within machine precision), indicating the correspondence of the computed solution of problem $(\mathcal{P})$ to a solution to the exact model problem $(\mathcal{P}_0)$.

The only assumption from Proposition 12.10 that cannot be verified for this implementation is the assumption that the algorithm fixed point is a singleton; all other assumptions are satisfied *automatically* by the problem structure. We *observe*, however, starting from around iteration 1500 in Fig. 12.2b, behavior that is *consistent* with

(i.e. does not contradict) linear convergence. From this, the observed convergence rate is $c = 0.9997$, which yields an a posteriori upper estimate of the pixel-wise error of about $8.9062e^{-4}$, or 3 digits of accuracy at each pixel.²

12.4 Primal-Dual Methods

The ADMM method presented above suffers from the extreme computational cost of computing the prox-operator in step 1. The results of the previous section required several days of cpu time on a 2016-era laptop. In this section we present a method studied in [3] that can achieve results in about 30 s on the same computer architecture. In this section we survey the main results (without proofs) from [1]. There is one subtle difference in the present survey over [3] that has major implications for the application and implementation of the main Algorithm 2.

In this section we consider exclusively functions g in problem $(\mathcal{P})$ of the form (12.11). The algorithm we revisit is the proximal alternating predictor-corrector (PAPC) algorithm proposed in [37] for solving $(\mathcal{S})$. It consists of a predictor-corrector gradient step for handling the smooth part of L in (12.13) and a proximal step for handling the nonsmooth part.

Algorithm 2: Extended Proximal Alternating Predictor-Corrector (EPAPC) for $(\mathcal{S})$.

Parameters : Set $\eta > 0$ and choose the parameters τ and σ to satisfy

$$\tau \in \left(0, \frac{1}{K_f}\right), \quad 0 < \tau\sigma < \frac{1}{\|\mathcal{A}\mathcal{A}^T\|}, \quad \text{and} \\ (\forall w \in (\ker \mathcal{A}^T)^\perp)(\forall w_0 \in \ker \mathcal{A}^T) \quad \frac{1}{\sigma}\|w_0\|^2 \geq g^*(w) - g^*(w + w_0).$$

Initialization: Choose $(x^0, y^0) \in \mathbb{R}^n \times (\ker \mathcal{A}^T)^\perp$

```

1 for  $k = 1, 2, \dots$  do
2    $p^k = x^{k-1} - \tau(\nabla f(x^{k-1}) + \mathcal{A}^T y^{k-1});$ 
3   for  $i = 1, \dots, M$  do
4      $w_i^k = \operatorname{argmax}_{w_i \in \mathbb{R}^{m_i}} \{ \langle A_i p^k, w_i \rangle - g_i^*(w_i) - (1/2\sigma) \|w_i - y_i^{k-1}\|^2 \}$ 
5      $\equiv \operatorname{prox}_{\sigma, g_i^*}(y_i^{k-1} + \sigma A_i p^k);$ 
6    $y^k = P_{(\ker \mathcal{A}^T)^\perp} w^k;$ 
7    $x^k = x^{k-1} - \tau(\nabla f(x^{k-1}) + \mathcal{A}^T y^k).$ 

```

At each iteration the algorithm computes one gradient and a prox-mapping corresponding to the nonsmooth function, both of which are assumed to efficiently implementable. We suppose these can be evaluated exactly, though this does not take into account finite precision arithmetic. The dimension of the iterates of the EPAPC

²This statement does not take into account finite precision evaluation of the steps in Algorithm 1.

algorithm is on the same order of magnitude as with the ADMM/Douglas-Rachford method, but the individual steps can be run in parallel and, with the exception of the projection in Step 6, are much less computationally intensive to execute.

For quantitative convergence guarantees of primal-dual methods, additional assumptions are required.

Assumption 3

- (i) *The function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex and continuously differentiable with Lipschitz continuous gradient ∇f (constant K_f), that is for all $x, x' \in \mathbb{R}^n$, we have*

$$\|\nabla f(x) - \nabla f(x')\| \leq K_f \|x - x'\|; \quad (12.25)$$

- (ii) *The function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is pointwise quadratically supportable (Definition 12.2) at each $\hat{x}$ in the solution set $\mathcal{S}^*$.*

- (iii) *There exists a $\sigma > 0$ such that*

$$(\forall w \in (\ker \mathcal{A}^T)^\perp) (\forall w_0 \in \ker \mathcal{A}^T) \quad \frac{1}{\sigma} \|w_0\|^2 \geq g^*(w) - g^*(w + w_0).$$

The assumption of Lipschitz continuous gradients Assumption 3(i), is standard, but stronger than one might desire in general. The assumption is included mainly to guarantee boundedness of the iterates. Lipschitz continuity of the gradients is enough for our purposes, however. By the standing Assumption 1 the mapping $\mathcal{A}$ is injective and when $m \leq n$, then $\mathcal{A}$ has full row rank, and $\mathcal{A}\mathcal{A}^T$ is invertible. When $m > n$, $\mathcal{A}$ is still injective but $\mathcal{A}\mathcal{A}^T$ has a nontrivial kernel and care must be taken that the conjugate function g^* does not decrease too fast in the direction of the kernel of $\mathcal{A}^T$. This is assured by Assumption 3(iii). This assumption comes into play in Lemma 12.1.

Step (3) of Algorithm 2 can be written more compactly when $g(w) := g(w_1, \dots, w_M) = \sum_{i=1}^M g_i(w_i)$. In this case, the convex conjugate of a separable sum of functions is the sum of the individual conjugates: $g^*(w) := \sum_{i=1}^M g_i^*(w_i)$. Defining the matrix $S = \sigma^{-1} I_m$ we immediately get that for any point $\zeta_i \in \mathbb{R}^{m_i}$, $i = 1, \dots, M$,

$$\text{prox}_{S, g^*}(\zeta) = (\text{prox}_{\sigma, g_1^*}(\zeta_1), \text{prox}_{\sigma, g_2^*}(\zeta_2), \dots, \text{prox}_{\sigma, g_M^*}(\zeta_M)).$$

Thus Step (3) of Algorithm 2 can be written in vector notation by $w^k = \text{prox}_{S, g}(y^{k-1} + \sigma \mathcal{A}p^k)$. It is possible to use different proximal step constants σ_i , $i = 1, \dots, M$, see details in [37]. The choice $\sigma_i = \sigma$ for $i = 1, \dots, M$ is purely for simplicity of exposition. The projection onto $(\ker \mathcal{A}^T)^\perp$ in (6) is carried out by applying the pseudo inverse:

$$P_{(\ker \mathcal{A}^T)^\perp} = \mathcal{A}(\mathcal{A}^T \mathcal{A})^{-1} \mathcal{A}^T.$$

When $m \leq n$ and A_i is full rank for all $i = 1, 2, \dots, M$, then $\ker \mathcal{A}^T = \{0\}$ and the above operation is not needed. But an unavoidable feature of multi-resolution

analysis, our motivating application, is that $m > n$, so some thought must be given to efficient computation of $(\mathcal{A}^T \mathcal{A})^{-1}$.

The next technical result, which is new, establishes a crucial upper bound on the growth of the Lagrangian with respect to the primal variables.

Lemma 12.1 *Let Assumption 3 hold and let $((p^k, y^k, x^k))_{k \in \mathbb{N}}$ be the sequence generated by the EPAPC algorithm. Then for every $k \in \mathbb{N}$ and every $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$,*

$$\begin{aligned} L(x^k, y^k) - L(x, y^k) &\leq \frac{1}{2\tau} \left(\|x - x^{k-1}\|^2 - \|x - x^k\|^2 \right) \\ &\quad - \frac{1}{2} \left(\frac{1}{\tau} - K_f \right) \|y^k - y^{k-1}\|^2. \end{aligned} \quad (12.26)$$

and

$$L(x^k, y) - L(x^k, y^k) \leq \frac{1}{2} \left(\|y - y^{k-1}\|_G^2 - \|y - y^k\|_G^2 - \|y^k - y^{k-1}\|_G^2 \right) \quad (12.27)$$

where

$$G := \sigma^{-1} I_m - \tau \mathcal{A} \mathcal{A}^T. \quad (12.28)$$

Note that for the choice of τ given in the parameter initialization of Algorithm 2, $G > 0$.

Proof The proof of (12.26) follows exactly as in the proof of [37, Lemma 3.1(i)]. The second inequality (12.27) follows exactly as the proof of [37, Lemma 3.1(ii)] in the case that $\ker \mathcal{A}^T = \{0\}$. The case where $\ker \mathcal{A}^T$ is nontrivial requires more care. The proof of this is suppressed in the interest of brevity.

The next intermediate result establishes pointwise quadratic supportability (Definition 12.2) on bounded sets at all saddle points under Assumptions 1 and 3.

Proposition 12.11 ([3]) *Let $((p^k, y^k, x^k))_{k \in \mathbb{N}}$ be the sequence generated by the EPAPC algorithm. If Assumptions 1 and 3 are satisfied, then for any primal solution $\hat{x}$ to the saddle point problem (S), there exists a $\mu > 0$ such that*

$$f(x^k) \geq f(\hat{x}) + \langle \nabla f(\hat{x}), x^k - \hat{x} \rangle + \frac{1}{2} \mu \|x^k - \hat{x}\|^2 \quad \forall k. \quad (12.29)$$

The constant μ in Proposition 12.11 depends on the choice of (x^0, y^0) and so depends implicitly on the distance of the initial guess to the point in the set of saddle point solutions.

Convergence of the primal-dual sequence is with respect to a weighted norm on the primal-dual product space built on G in (12.28).

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad H = \begin{pmatrix} \tau^{-1} I_n & 0 \\ 0 & G \end{pmatrix} \quad (12.30)$$

where by the assumptions on the choice of τ given in Algorithm 2, $G \succ 0$. We can then define an associated norm using the positive definite matrix H , $\|u\|_H^2 := \frac{1}{\tau} \|x\|^2 + \|y\|_G^2$.

We are now ready to state the main result and corollaries, whose proofs can be found in [3].

Theorem 12.4.1 *Let $((p^k, x^k, y^k))_{k \in \mathbb{N}}$ be the sequence generated by the EPAPC algorithm. Let $\lambda_{\min+}(\mathcal{A}\mathcal{A}^T)$ denote the smallest nonzero eigenvalue of $\mathcal{A}\mathcal{A}^T$. If Assumptions 1 and 3 are satisfied, then there exists a saddle point solution for $L(\cdot, \cdot)$, the pair $\hat{u} = (\hat{x}, \hat{y})$, with $\hat{y} \in (\ker \mathcal{A}^T)^\perp$, and for any $\alpha > 1$ and for all $k \geq 1$, the sequence $(u^k = (x^k, y^k))_{k \in \mathbb{N}}$ satisfies*

$$\|u^k - \hat{u}\|_H^2 \leq \frac{1}{1 + \delta} \|u^{k-1} - \hat{u}\|_H^2, \quad (12.31)$$

where

$$\delta = \min \left\{ \frac{(\alpha - 1)\tau\sigma(1 - \tau L_f)\lambda_{\min+}(\mathcal{A}\mathcal{A}^T)}{\alpha}, \frac{\mu\tau\sigma\lambda_{\min+}(\mathcal{A}\mathcal{A}^T)}{\alpha\tau L_f^2 + \sigma\lambda_{\min+}(\mathcal{A}\mathcal{A}^T)} \right\} \quad (12.32)$$

is positive and $\mu > 0$ is the constant of pointwise quadratic supportability of f at $\hat{x}$ depending on the distance of the initial guess to the point $(\hat{x}, \hat{y})$ in the solution set S^* . In particular, $((x^k, y^k))_{k \in \mathbb{N}}$ is Q -linearly convergent with respect to the H -norm to a saddle-point solution.

An unexpected corollary of the result above is that the set of saddle points is a singleton.

Corollary 12.12 (Unique saddle point) *Under Assumptions 1 and 3 the solution set S^* is a singleton.*

The above theorem yields the following estimate on the number of iterations required to achieve a specified distance to a saddle point.

Corollary 12.13 *Under Assumptions 1 and 3, let $\bar{u} = (\bar{x}, \bar{y})$ be the limit point of the sequence generated by the EPAPC algorithm. In order obtain*

$$\|x^k - \bar{x}\| \leq \epsilon \quad (\text{resp. } \|y^k - \bar{y}\|_G \leq \epsilon), \quad (12.33)$$

it suffices to compute k iterations, with

$$k \geq \frac{2 \log \left(\frac{C}{L_f \epsilon} \right)}{\delta} \quad \left(\text{resp. } k \geq \frac{2 \log \left(\frac{C}{\epsilon} \right)}{\delta} \right), \quad (12.34)$$

where $C = \|u^0 - \bar{u}\|_H = \left(\frac{1}{\tau} \|x^0 - \bar{x}\|^2 + \|y^0 - \bar{y}\|_G^2 \right)^{1/2}$, and δ is given in (12.32).

12.4.1 EPAPC for Statistcal Multi-resolution Estimation of STED Images

An efficient computational strategy for evaluating or at least approximating the projection $P_{(\ker \mathcal{A}^T)^\perp}$ in Step 6 of Algorithm 2 has not yet been established. We report here preliminary computational results of Algorithm 2 *without* computing Step 6. Our results show that the method is promising, though error bounds to the solution to (S) are not justified without computation of $P_{(\ker \mathcal{A}^T)^\perp}$.

In our numerical experiments, the constraint penalty in (S) takes the form $g(y) = \sum_i \iota_{C_i}(y)$ where each $C_i = \{y : |\sum_{j \in V_i} \omega_j(y_j - b_j)| \leq \gamma_i\}$. This is an exact penalty function, and so solutions to (S) correspond to solutions to (P_0) . Using Moreau's identity (12.6), the prox-mapping is evaluated explicitly in (6) for each constraint by

$$\begin{aligned} y_i^k &= \text{prox}_{\sigma, \iota_{C_i}^*}(y_i^{k-1} + \sigma A_i p^k) \\ &= y_i^{k-1} + \sigma A_i p^k - \sigma P_{C_i} \left(\frac{y_i^{k-1} + \sigma A_i p^k}{\sigma} \right), \quad i = 1, \dots, M. \end{aligned}$$

The proximal parameter is a function of τ and given by $\sigma = 1/(\tau \|\mathcal{A} \mathcal{A}^T\|_2)$. More details in [37, Sect. 4.1].

Here, we also consider the smooth approximation of the L^1 -norm as the qualitative objective. The L^1 -norm is non-smooth at the origin, thus in order to make the derivative-based methods possible we consider a smoothed approximation of this, known as the Huber approximation.

The Huber loss function is defined as follows:

$$\|x\|_{1,\alpha} = \sum_{i,j} \phi_\alpha(x_{i,j}), \quad \phi_\alpha(t) = \begin{cases} \frac{t^2}{2\alpha} & \text{if } |t| \leq \alpha \\ |t| - \frac{\alpha}{2} & \text{if } |t| > \alpha, \end{cases} \quad (12.35)$$

where $\alpha > 0$ is a small parameter defining the trade-off between quadratic regularization (for small values) and L^1 regularization (for larger values). The function ϕ is smooth with $\frac{1}{\alpha}$ -Lipschitz continuous derivative and its derivative is given by

$$\phi'_\alpha(t) = \begin{cases} \frac{t}{\alpha} & \text{if } |t| \leq \alpha \\ \text{sgn}(t) & \text{if } |t| > \alpha. \end{cases} \quad (12.36)$$

Pointwise quadratically supportability of this function at solutions is not unreasonable but still must be assumed.

We demonstrate our reconstruction of the image inset shown in Fig. 12.1 of size $n = 64^2$ with the same SMRE model as the demonstration in Sect. 12.3.1. The confidence level γ_i was set to $0.25 * i$ at each resolution level ($i = 1, 2, 3$). Figure 12.3(top left) shows the reconstruction with the L^2 function $f(x) = 0.01\|x\|^2$ (compare to Fig. 12.2). Figure 12.3(top right) shows the reconstruction with the Huber function

Fig. 12.3 Results of image denoising and deconvolution via Algorithm 2 for statistical multi-resolution estimation with three levels at iteration 4000 with **a** $f(x) = 0.01 * \|x^2\|^2$ and **b** $f(x) = \|x\|_{1,0.25}$, the Huber loss function. Frame **c** shows the step length $\|u^{k+1} - u^k\|_H$ where $u^k = (x^k, y^k)$ of the associated sequence as function of the number of iterations

$\|x\|_{1,\alpha}$, where $\alpha = 0.25$. Figure 12.3(bottom) shows the step size of the primal-dual pair for each of these regularized problems as a function of iteration. The model with quadratic regularization achieves a better average rate of convergence, but for both objective functions the algorithm appears to exhibit R-linear convergence (not Q-linear). What is not evident from these experiments is the computational effort required per iteration. Without computation of the pseudo-inverse in step 6, the EPAPC algorithm computes these results in about 30 s on a 2018-era laptop, compared to several days for the results shown in Fig. 12.2.

12.5 Randomized Block-Coordinate Primal-Dual Methods

The previous sections reviewed numerical strategies and structures that yield quantitative estimates of the distance of an iterate to the solution of the underlying variational problem. In this section we examine *implementation* strategies for dealing

with high dimensional problems. These are implementation strategies because they do not involve changing the optimization model. Instead, we select at random a smaller subset of variables or constraints in the computation of an update in the full-dimensional iterative procedure. This is the principal strategy for handling problems that, due to their size, must be distributed across many processing and storage units (see for instance [26, 38–40] and references therein). We survey here a randomized primal-dual technique proposed and analyzed in [2]. The main theoretical question to resolve with such approaches is whether, and in what sense, iterates converge to a solution to the original problem. We can determine whether the iterates converge, but obtaining an estimate of the distance to the solution remains an open problem.

The algorithm below is a primal-dual method like the algorithms reviewed above, with the exception that it solves an extension of the dual problem $(\mathcal{D})$:

$$\begin{aligned} & \underset{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m}{\text{minimize}} && f^*(x) + g^*(y) \\ & \text{subject to} && x = -\mathcal{A}^T y \end{aligned} \quad (\tilde{\mathcal{D}})$$

The main prox operation is computed on the dual objective in $(\mathcal{D})$, that is $f^*(x) + g^*(y)$ with respect to the variables $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. The dimension of the basic operations is unchanged from the previous approaches, but the structure of the sum of functions allows for efficient evaluation of the prox mapping. Implicit in this is that the function f is prox friendly. In the algorithm description below it is convenient to use the convention $f \equiv g_0$, $A_0 \equiv \text{Id}$. The algorithm is based in part on [39].

Algorithm 3: Random Block-coordinate Primal-Dual Algorithm (RBPd) for $(\tilde{\mathcal{D}})$

Parameters : Choose $\tau > 0$, $\sigma = (\sigma_0, \sigma_1, \dots, \sigma_M) \in \mathbb{R}_{++}^M$.

Initialization: Choose $y^1 = 0 \in \mathbb{R}^m$, and set $x^1 = u^1 = -\mathcal{A}^T y^1 = 0 \in \mathbb{R}^n$.

```

1 for  $k = 1, 2, \dots$  do
2   Choose  $i \in \{0, 1, 2, \dots, M\}$  uniformly at random;
3    $y_i^{k+1} = \text{prox}_{\sigma_i, g_i^*}(y_i^k - \sigma_i A_i x^k)$ ;
4    $\delta^{k+1} = A_i^T (y_i^{k+1} - y_i^k)$ ;
5    $u^{k+1} = u^k + \frac{\tau}{m} \delta^{k+1}$ ;
6    $x^{k+1} = u^{k+1} + x^k + \tau \delta^{k+1}$ ;
```

Notice that each iteration of Algorithm 3 requires only two small matrix-vector multiplications: $A_i(\cdot)$ and $A_i^T(\cdot)$. The methods of the previous sections, in contrast, worked with full matrix $\mathcal{A} = [A_1^T, \dots, A_m^T]^T$. This means that all iterations involve full vector operations. For some applications this might be not feasible, at least on standard desktop computers due to the size of problems. Algorithm 3 uses only blocks A_i of $\mathcal{A}$, therefore each iteration requires fewer floating point operations, at the cost of having less information available for choosing the next step. This reduction in the effectiveness of the step is compensated for through larger block-wise steps. Computation of the step size is particularly simple. This follows the same

hybrid step-length approach developed in [41] for the nonlinear problem of blind ptychography. In particular, we use step sizes adjusted to each block A_i :

$$\tau \sigma_i \lambda_{\max}(A_i^T A_i) < 1 \quad \forall i = 0, 1, \dots, m.$$

Proposition 12.14 (Theorem 1 of [2]) *Suppose Assumption 1 holds and let $\tau \sigma_i \|A_i\|^2 < 1$. Then (x^k, y^k) generated by Algorithm 3 converges to a solution to $(\mathcal{D})$. In particular, the sequence (x^k) converges almost surely to a solution to $(\mathcal{P})$.*

The statement above concerns part (i) of Theorem 1 of [2]. No smoothness is required of the qualitative regularization f . Instead, it is assumed that this function is prox-friendly. This opens up the possibility of using the 1-norm as a regularizer, promoting, in some sense, sparsity in the image. No claim is made on the *rate* of convergence, though the numerical experiments below indicate that, for regular enough functions f , convergence might be locally linear. This remains to be proved.

12.5.1 RBPD for Statistcal Multi-resolution Estimation of STED Images

Despite many open questions regarding convergence, random methods offer a way to handle extremely large problems. To make a comparison with the deterministic approaches above, cycles of the RBPD Algorithm 3 are counted in terms of *epochs*. An epoch contains the number of passes through steps 1–6 of Algorithm 3 required before each block is chosen at least once. After k epochs, therefore, the i -th coordinate of x will be updated, on average, the same number of times for the randomized algorithm as for the deterministic methods. In other words, an epoch for a randomized block-wise method is comparable to an iteration of a deterministic method. As the RBPD updates only one block per *iteration*, each iteration is less computationally intensive than the the deterministic counterparts. However, in our case this efficient iteration still requires one to evaluate two (possibly) expensive convolution products (embedded in $A_i x$ and $A_i^T y$). Thus, if these operations are relatively expensive, the efficiency gain will be marginal. Nevertheless, because of the ability to operate on smaller blocks, the randomized method requires, per epoch, approximately half the time required per iteration of the deterministic methods. Although the quantitative convergence analysis remains open, our numerical experiments indicate that the method achieves a comparable step-residual to the EPAPC Algorithm 2 after the same number of epochs/iterations.

As with the experiments in the previous Sections, we use three resolutions, which results in one block at the highest resolution, four blocks at the next resolution (four possible shifts of 2×2 pixels), and nine blocks at the third resolution (nine different shifts of 3×3 pixels). We applied Algorithm 3 with different regularization f in $(\mathcal{P}_0)$: the 1-norm $f(x) = \|x\|_1$, Huber loss function $f(x) = \|x\|_{1,\alpha}$ given by (12.35) ($\alpha = 0.25$) and the squared Euclidean norm. As with the EPAPC experiments, the

Fig. 12.4 Results of image denoising and deconvolution using RBPD (Algorithm 3) for statistical multi-resolution estimation of a 64×64 pixel image with three levels at epoch 4000 with **a** $f(x) = 0.01 * \|x\|^2$, **b** $f(x) = \|x\|_{1,0.25}$, the Huber loss function and **c** $f(x) = \|x\|_1$. **d** The step length $\|u^{k+1} - u^k\|_H$ where $u^k = (x^k, y^k)$ of the associated sequence as function of the number of iterations

function g is given by (12.11) with g_i given by (12.17) for the parameter $\gamma_i = 0.25 * i$ for $i = 1, 2, 3$. All of these functions are prox-friendly and have closed-form Fenchel conjugates. The gain in efficiency over the deterministic EPAPC method proposed above (without computation of the pseudo-inverse) is a factor of 2.

Figure 12.4a–c shows the reconstructions on the same 64×64 image data used in the previous sections. The numerical performance of the algorithm is shown in Fig. 12.4(d). What the more efficient randomization strategy enables is for the full 976×976 pixel image to be processed. The result for regularization with the 1-norm is shown in Fig. 12.5.

Fig. 12.5 Left: original image. Right: image denoising and deconvolution using RBPD (Algorithm 3) for statistical multi-resolution estimation with three levels at epoch 4000 with $f(x) = \|x\|_1$

References

1. Aspelmeier, T., Charitha, C., Luke, D.R.: Local linear convergence of the ADMM/Douglas-Rachford algorithms without strong convexity and application to statistical imaging. *SIAM J. Imaging Sci.* **9**(2), 842–868 (2016)
2. Luke, D.R., Malitsky, Y.: Block-coordinate primal-dual method for the nonsmooth minimization over linear constraints. In: Giselsson, P., Rantzer, A. (eds.) *Distributed and Large-Scale Optimization*. Springer (2018)
3. Luke, D.R., Shefi, R.: A globally linearly convergent method for pointwise quadratically supportable convex-concave saddle point problems. *J. Math. Anal. Appl.* **457**(2), 1568–1590 (2018). <https://doi.org/10.1016/j.jmaa.2017.02.068>
4. Burger, M., Sawatzky, A., Steidl, G.: First order algorithms in variational image processing. In: *Splitting Methods in Communication, Imaging, Science, and Engineering*, pp. 345–407. Springer (2016)
5. Chambolle, A., Pock, T.: An introduction to continuous optimization for imaging. *Acta Numerica* **25**, 161–319 (2016)
6. Combettes, P.L., Pesquet, J.C.: Proximal splitting methods in signal processing. In: *Fixed Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212 (2011)
7. Komodakis, N., Pesquet, J.C.: Playing with duality: an overview of recent primal? Dual approaches for solving large-scale optimization problems. *IEEE Signal Proc. Mag.* **32**(6), 31–54 (2015)
8. Luke, D.R., Teboulle, M., Thao, N.H.: Necessary conditions for linear convergence of iterated expansive, set-valued mappings. *Math. Program. A* **180**, 1–31 (2020). <https://doi.org/10.1007/s10107-018-1343-8>
9. Ortega, J.M., Rheinboldt, W.C.: *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York (1970)
10. Rockafellar, R.T., Wets, R.J.: *Variational Analysis*. Grundlehren Math. Wiss. Springer, Berlin (1998)
11. Dontchev, A.L., Rockafellar, R.T.: *Implicit Functions and Solution Mappings*, 2nd edn. Springer, New York (2014)
12. Moreau, J.J.: Proximité et dualité dans un espace Hilbertien. *Bull. de la Soc. Math. de France* **93**(3), 273–299 (1965)

13. Aubin, J.P., Frankowska, H.: Set-valued Analysis. Birkhäuser, Boston (1990)
14. Burke, J.V., Ferris, M.C.: Weak sharp minima in mathematical programming. *SIAM J. Control. Optim.* **31**, 1340–1359 (1993)
15. Friedlander, M.P., Tseng, P.: Exact regularization of convex programs. *SIAM J. Optim.* **18**(4), 1326–1350 (2007). <https://doi.org/10.1137/060675320>
16. Borwein, J., Vanderwerff, J.: Convex Functions: Constructions, Characterizations and Counterexamples, Encyclopedias in Mathematics, vol. 109. Cambridge University Press, New York (2010)
17. Eckstein, J.: Splitting methods for monotone operators with applications to parallel optimization. Ph.D. thesis, MIT, Cambridge, MA (1989)
18. Eckstein, J., Bertsekas, D.P.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.* **55**, 293–318 (1992)
19. Gabay, D.: Applications of the method of multipliers to variational inequalities. In: *Augmented Lagrangian Methods: Applications to the Solution of Boundary-Value Problems*, pp. 299–331. North-Holland, Amsterdam (1983)
20. Glowinski, R., Marroco, A.: Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité, d’une classe de problèmes de dirichlet non linéaires. *Revue Française d’Automatique, Informatique et Recherche Opérationnelle* **9**(R-2), 41–76 (1975)
21. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. *SIAM J. Control. Optim.* **14**, 877–898 (1976)
22. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **3**(1), 1–122 (2011)
23. Douglas Jr., J., Rachford Jr., H.H.: On the numerical solution of heat conduction problems in two or three space variables. *Trans. Am. Math. Soc.* **82**, 421–439 (1956)
24. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16**, 964–979 (1979)
25. Svaiter, B.F.: On weak convergence of the Douglas-Rachford method. *SIAM J. Control. Optim.* **49**(1), 280–287 (2011)
26. Strohmer, T., Vershynin, R.: A randomized kaczmarz algorithm with exponential convergence. *J. Fourier Anal. Appl.* **15**(2), 262–278 (2009)
27. He, B., Yuan, X.: On non-ergodic convergence rate of Douglas-Rachford alternating direction method of multipliers. *Numerische Mathematik* **130**(3), 567–577 (2015)
28. He, B., Yuan, X.: On the $o(1/n)$ convergence rate of the Douglas-Rachford alternating direction method. *SIAM J. Numer. Anal.* **50**(2), 700–709 (2012)
29. Boley, D.: Local linear convergence of the alternating direction method of multipliers on quadratic or linear programs. *SIAM J. Optim.* **23**(4), 2183–2207 (2013). <https://doi.org/10.1137/120878951>
30. Hesse, R., Luke, D.R.: Nonconvex notions of regularity and convergence of fundamental algorithms for feasibility problems. *SIAM J. Optim.* **23**(4), 2397–2419 (2013)
31. Luke, D.R., Thao, N.H., Tam, M.K.: Quantitative convergence analysis of iterated expansive, set-valued mappings. *Math. Oper. Res.* **43**(4), 1143–1176 (2018). <https://doi.org/10.1287/moor.2017.0898>
32. Hesse, R., Luke, D.R., Neumann, P.: Alternating projections and Douglas-Rachford for sparse affine feasibility. *IEEE Trans. Signal. Process.* **62**(18), 4868–4881 (2014). <https://doi.org/10.1109/TSP.2014.2339801>
33. Phan, H.: Linear convergence of the Douglas-Rachford method for two closed sets. *Optimization* **65**, 369–385 (2016)
34. Hell, S.W., Wichmann, J.: Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**(11), 780–782 (1994). <https://doi.org/10.1364/OL.19.000780>
35. Klar, T.A., Jakobs, S., Dyba, M., Egner, A., Hell, S.W.: Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc. Natl. Acad. Sci.* **97**(15), 8206–8210 (2000). <https://doi.org/10.1073/pnas.97.15.8206>

36. Frick, K., Marnitz, P., Munk, A.: Statistical multiresolution dantzig estimation in imaging: fundamental concepts and algorithmic framework. *Electron. J. Stat.* **6**, 231–268 (2012)
37. Drori, Y., Sabach, S., Teboulle, M.: A simple algorithm for a class of nonsmooth convex-concave saddle-point problems. *Oper. Res. Lett.* **43**(2), 209–214 (2015)
38. Defazio, A., Bach, F., Lacoste-Julien, S.: SAGA: A fast incremental gradient method with support for non-strongly convex composite objectives. In: *Advances in Neural Information Processing Systems*, pp. 1646–1654 (2014)
39. Fercoq, O., Richtárik, P.: Accelerated, parallel, and proximal coordinate descent. *SIAM J. Optim.* **25**(4), 1997–2023 (2015)
40. Johnson, R., Zhang, T.: Accelerating stochastic gradient descent using predictive variance reduction. In: *Advances in Neural Information Processing Systems*, pp. 315–323 (2013)
41. Hesse, R., Luke, D.R., Sabach, S., Tam, M.: The proximal heterogeneous block implicit-explicit method and application to blind ptychographic imaging. *SIAM J. Imaging Sci.* **8**(1), 426–457 (2015)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.