Stable Recovery with Analysis Decomposable Priors

Jalal M. Fadili
GREYC
CNRS-ENSICAEN-Univ. Caen
Caen, France

Gabriel Peyré and Samuel Vaiter
CEREMADE
CNRS-Univ. Paris-Dauphine
Paris, France

Charles-Alban Deledalle
IMB
CNRS-Univ. Bordeaux 1
Bordeaux, France

Joseph Salmon
LTCI
CNRS-Télécom ParisTech
Paris, France

***Abstract*—In this paper, we investigate in a unified way the structural properties of solutions to inverse problems. These solutions are regularized by the generic class of semi-norms defined as a decomposable norm composed with a linear operator, the so-called analysis type decomposable prior. This encompasses several well-known analysis-type regularizations such as the discrete total variation (in any dimension), analysis group-Lasso or the nuclear norm. Our main results establish sufficient conditions under which uniqueness and stability to a bounded noise of the regularized solution are guaranteed. Along the way, we also provide a strong sufficient uniqueness result that is of independent interest and goes beyond the case of decomposable norms.**

I. Introduction

A. Problem statement

Suppose we observe

$$y = \Phi x_0 + w, \quad \text{where} \quad \|w\|_2 \leqslant \varepsilon \ ,$$

where Φ is a linear operator from $\mathbb{R}^N$ to $\mathbb{R}^M$ that may have a non-trivial kernel. We want to robustly recover an approximation of x_0 by solving the optimization problem

$$x^\star \in \underset{x\in\mathbb{R}^N}{\mathrm{Argmin}}\ \tfrac{1}{2}\|y - \Phi x\|_2^2 + \lambda R(x) \ , \tag{1}$$

where

$$R(x) := \|L^* x\|_{\mathcal{A}} \ ,$$

with $L : \mathbb{R}^P \to \mathbb{R}^N$ a linear operator, and $\|\cdot\|_{\mathcal{A}} : \mathbb{R}^P \to \mathbb{R}^+$ is a decomposable norm in the sense of [1]. Decomposable regularizers are intended to promote solutions conforming to some notion of simplicity/low complexity that complies with that of $u_0 = L^* x_0$. This motivates the following definition of these norms. Throughout the paper, given a subspace $V \subset \mathbb{R}^P$, we will use the shorthand notation $L_V = L\mathcal{P}_V$, $L_V^* = \mathcal{P}_V L^*$, and $\alpha_V = \mathcal{P}_V \alpha$ for any vector $\alpha \in \mathbb{R}^P$, where $\mathcal{P}_V$ (resp. $\mathcal{P}_{V^\perp}$) is the orthogonal projector on V (resp. on its orthogonal complement $V^\perp$).

Definition 1. *A norm $\|\cdot\|_{\mathcal{A}}$ is* decomposable *at $u \in \mathbb{R}^P$ if:*

(i) there is a subspace $T \subset \mathbb{R}^P$ and a vector $e \in T$ such that

$$\partial\|\cdot\|_{\mathcal{A}}(u) = \left\{\alpha \in \mathbb{R}^P \mid \alpha_T = e \quad \text{and} \quad \|\alpha_{T^\perp}\|_{\mathcal{A}}^* \leqslant 1\right\}$$

(ii) and for any $z \in T^\perp$, $\|z\|_{\mathcal{A}} = \sup_{v\in T^\perp, \|v\|_{\mathcal{A}}^* \leqslant 1} \langle v,\, z\rangle$, where $\|\cdot\|_{\mathcal{A}}^*$ is the dual norm of $\|\cdot\|_{\mathcal{A}}$.

From this definition, it can be easily proved, using Fenchel identity, that $u \in T$ whenever $\|\cdot\|_{\mathcal{A}}$ is decomposable at u. Popular examples covered by decomposable regularizers are the ℓ_1-norm, the ℓ_1-ℓ_2 group sparsity norm, and the nuclear norm [1].

B. Contributions and relation to prior work

In this paper, we give a strong sufficient condition under which (1) admits a unique minimizer. From this, sufficient uniqueness conditions are derived. Then we develop results guaranteeing a stable approximation of x_0 from the noisy measurements y by solving (1), with an ℓ_2-error that comes within a factor of the noise level ε. This goes beyond [1] who considered identifiability under a generalized irrepresentable condition in the noiseless case with $L = \mathrm{Id}$. ℓ_2-stability for a class of decomposable priors closely related to Definition 1, is also studied in [8] for $L = \mathrm{Id}$ and general sufficiently smooth data fidelity. Their stability results require however stronger assumptions than ours (typically a restricted strong convexity which becomes a type of restricted eigenvalue property for linear regression with quadratic data fidelity). The authors in [3] provide sharp estimates of the number of generic measurements required for exact and ℓ_2-stable recovery of models from random partial information by solving a constrained form of (1) regularized by atomic norms. This is however restricted to the compressed sensing scenario. Our results generalize the stability guarantee of [7] established when the decomposable norm is ℓ_1 and L^* is the analysis operator of a frame. A stability result for general sublinear functions R is given in [6]. The stability is however measured in terms of R, and ℓ_2-stability can only be obtained if R is coercive, *i.e.,* L^* is injective.

At this stage, we would like to point out that although we carry out our analysis on the penalized form (1), our results remain valid for the data fidelity constrained version but obviously with different constants in the bounds. We omit these results for obvious space limitations.

II. Uniqueness

A. Main assumptions

We first note that traditional coercivity and convexity arguments allow to show that the set of (global) minimizers of (1) is a non-empty compact set if, and only if, $\ker(\Phi)\cap\ker(L^*) = \{0\}$.

The following assumptions will play a pivotal role in our analysis.

Assumption (SC_x) There exist $\eta \in \mathbb{R}^M$ and $\alpha \in \partial\|\cdot\|_{\mathcal{A}}(L^* x)$ such that the following so-called source (or range) condition is verified:

$$\Phi^* \eta = L\alpha \in \partial R(x) \ .$$

Assumption (INJ_T) For a subspace $T \subset \mathbb{R}^P$, Φ is injective on $\ker(L^*_{T^\perp})$.

It is immediate to see that since $\ker(L^*) \subseteq \ker(L^*_{T^\perp})$, ($\mathrm{INJ}_T$) implies that the set of minimizers is indeed non-empty and compact.

B. Strong Null Space Property

We shall now give a novel strong sufficient uniqueness condition under which problem (1) admits exactly one minimizer.

Theorem 1. *For a minimizer $x^\star$ of* (1)*, let T and e be the subspace and vector in Definition 1 associated to $u^\star = L^* x^\star$, and denote $S = T^\perp$. $x^\star$ is the unique minimizer of* (1) *if*

$$\langle L^*_T h,\, e\rangle < \|L^*_S h\|^*_{\mathcal{A}}, \qquad \forall h \in \ker(\Phi) \setminus \{0\} \ .$$

The above condition is a strong generalization of the Null Space Property well known in ℓ_1 regularization [4].

C. Sufficient uniqueness conditions

1) General case: A direct consequence of the above theorem is the following corollary.

Corollary 1. *For a minimizer $x^\star$ of* (1)*, let T and e be the subspace and vector in Definition 1 associated to $u^\star = L^* x^\star$, and denote $S = T^\perp$. Assume that* ($\mathrm{SC}_{x^\star}$) *is verified with $\|\alpha_S\|^*_{\mathcal{A}} < 1$, and that* ($\mathrm{INJ}_T$) *holds. Then, $x^\star$ is the unique minimizer of* (1).

In fact, it turns out that the above two results are proved without requiring some restrictive implications of Definition 1(ii) of decomposable norms, and are therefore valid for a much larger class of regularizations. This can be clearly checked in the arguments used in the proofs.

2) Separable case:

Definition 2. *The decomposable norm $\|\cdot\|_{\mathcal{A}}$ is separable on the subspace $T^\perp = S = V \oplus W \subset \mathbb{R}^P$ if for any $u \in \mathbb{R}^P$, $\|u_{T^\perp}\|_{\mathcal{A}} = \|u_V\|_{\mathcal{A}} + \|u_W\|_{\mathcal{A}}$.*

Separability as just defined is fulfilled for several decomposable norms such as the ℓ_1 or $\ell_1 - \ell_p$ norms, $1 \leqslant p < +\infty$.

The non-saturation condition on the dual certificate required in Corollary 1 can be weakened to hold only on a subspace $V \subset S$ and the conclusions of the corollary remain valid, and assuming a stronger restricted injectivity assumption. We have the following corollary.

Corollary 2. *Assume that $\|\cdot\|_{\mathcal{A}}$ is also separable, with $S = V \oplus W$, such that* ($\mathrm{SC}_{x^\star}$) *is verified with $\|\alpha_V\|^*_{\mathcal{A}} < 1$, and* ($\mathrm{INJ}_V$) *holds. Then, $x^\star$ is the unique minimizer of* (1).

III. Stability to noise

A. Main result

1) General case: We are now ready to state our main stability results.

Theorem 2. *Let T_0 and e_0 be the subspace and vector in Definition 1 associated to $u_0 = L^* x_0$, and denote $S_0 = {T_0}^\perp$. Assume that* (SC_{x_0}) *is verified with $\|\alpha_{S_0}\|^*_{\mathcal{A}} < 1$, and that* ($\mathrm{INJ}_{T_0}$) *holds. Then, choosing $\lambda = c\varepsilon$, $c > 0$, the following holds for any minimizer $x^\star$ of* (1)

$$\|x^\star - x_0\|_2 \leqslant C\varepsilon \ ,$$

where $C = C_1\left(2 + c\|\eta\|_2\right) + C_2 \frac{(1+c\|\eta\|_2/2)^2}{c\left(1-\|\alpha_{S_0}\|^*_{\mathcal{A}}\right)}$, and $C_1 > 0$ and $C_2 > 0$ are constants independent of η and α.

Remark 1 (Separable case)**.** *When the decomposable norm is also separable (see Corollary 2), the stability result of Theorem 2 remains true assuming that $\|\alpha_V\|^*_{\mathcal{A}} < 1$ for $V \subset S_0$. This however comes at the price of the stronger restricted injectivity assumption* (INJ_V)*. To show this, the only thing to modify is the statement and the proof of Lemma 2 which can be done easily using similar arguments to those in the proof of Corollary 2.*

2) Case of frames: Suppose that L^* is the analysis operator of a frame ($\ker(L^*) = \{0\}$) with lower bound $a > 0$, let $\tilde{L}$ be a dual frame. The following stability bound can be obtained whose proof is omitted for space limitations.

Proposition 1. *Let T_0 and e_0 be the subspace and vector in Definition 1 associated to $u_0 = L^* x_0$, and denote $S_0 = {T_0}^\perp$. Assume that* (SC_{x_0}) *is verified with $\|\alpha_{S_0}\|^*_{\mathcal{A}} < 1$, and that Φ is injective on $\mathrm{Im}(\tilde{L}_{T_0})$. Then, choosing $\lambda = c\varepsilon$, $c > 0$, the following holds for any minimizer $x^\star$ of* (1)

$$\|x^\star - x_0\|_2 \leqslant C'\varepsilon \ ,$$

where $C' = C_1\left(2 + c\|\eta\|_2\right) + C'_2 \frac{(1+c\|\eta\|_2/2)^2}{c\left(1-\|\alpha_{S_0}\|^*_{\mathcal{A}}\right)}$, and $C_1 > 0$ and $C'_2 > 0$ are constants independent of η and α.

Since $\ker(L^*_{S_0}) \subseteq \mathrm{Im}(\tilde{L}_{T_0})$, the required restricted injectivity assumption is more stringent than (INJ_{T_0}). On the positive side, the constant C'_2 is in general better than C_2. More precisely, the constant C_L, see the proof of Theorem 2, is replaced with $\sqrt{a}$. Note also that coercivity of R in this case allows to derive a bound similar to ours from the results in [6]. His restricted injectivity assumption is however different and our constants are sharper.

B. Generalized irrepresentable condition

In the following corollary, we provide a stronger sufficient stability condition that can be viewed as a generalization of the irrepresentable condition introduced in [5] when R is the ℓ_1 norm. It allows to construct dual vectors η and α which obey the source condition and are computable, which in turn yield explicit constants in the bound.

Definition 3. *Let $T \subset \mathbb{R}^P$ and $e \in \mathbb{R}^P$, and denote $S = T^\perp$. Suppose that* (INJ_T) *is verified. Define for any $u \in \ker(L_S)$ and $z \in \mathbb{R}^M$ such that $\Phi^* z \in \mathrm{Im}(L_S)$*

$$\mathrm{IC}_{u,z}(T,e) \;=\; \|\Gamma e + u_S + (L_S)^+ \Phi^* z\|_{\mathcal{A}}^*$$

where

$$\begin{aligned} \Gamma &= (L_S)^+(\Phi^*\Phi\Xi - \mathrm{Id})L_T \\ \Xi : h \mapsto \Xi h &= \underset{x \in \ker(L_S^*)}{\operatorname{argmin}} \tfrac{1}{2}\|\Phi x\|_2^2 - \langle h,\, x\rangle \ , \end{aligned}$$

and M^+ is the Moore-Penrose pseudoinverse of M. Let $\bar{u}$, $\bar{z}$ and $\underline{u}$ defined as

$$\begin{aligned} (\bar{u}, \bar{z}) &= \underset{u \in \ker(L_S), \{z \mid \Phi^* z \in \mathrm{Im}(L_S)\}}{\operatorname{argmin}} \mathrm{IC}_{u,z}(T,e) \\ \textit{and} \quad \underline{u} &= \underset{u \in \ker(L_S)}{\operatorname{argmin}} \ \mathrm{IC}_{u,0}(T,e) \ . \end{aligned}$$

Obviously, we have

$$\mathrm{IC}_{\bar{u},\bar{z}}(T,e) \leqslant \mathrm{IC}_{\underline{u},0}(T,e) \leqslant \mathrm{IC}_{0,0}(T,e) \ .$$

The convex programs defining $\mathrm{IC}_{\bar{u},\bar{z}}(T,e)$ and $\mathrm{IC}_{\underline{u},0}(T,e)$ can be solved using primal-dual proximal splitting algorithms whenever the proximity operator of $\|\cdot\|_{\mathcal{A}}$ can be easily computed [2]. The criterion $\mathrm{IC}_{\underline{u},0}(T,e)$ specializes to the one developed in [10] when $\|\cdot\|_{\mathcal{A}}$ is the ℓ_1 norm. $\mathrm{IC}_{0,0}(T,e)$ is a generalization of the coefficient involved in the irrepresentable condition introduced in [5] when R is the ℓ_1 norm, and to the one in [1] for decomposable priors with $L = \mathrm{Id}$.

Corollary 3. *Assume that* (INJ_{T_0}) *is verified and $\mathrm{IC}_{\bar{u},\bar{z}}(T_0, e_0) < 1$. Then, taking $\eta = \Phi\Xi L_{T_0} e_0 + \bar{z}$, one can construct α such that* (SC_{x_0}) *is satisfied and $\|\alpha_{S_0}\|_{\mathcal{A}}^* < 1$. Moreover, the conclusion of Theorem 2 remains true substituting $1 - \mathrm{IC}_{\bar{u},\bar{z}}(T_0, e_0)$ for $1 - \|\alpha_{S_0}\|_{\mathcal{A}}^*$.*

IV. Proofs

A. Proof of Theorem 1

A key observation is that by strong (hence strict) convexity of $\mu \mapsto \|y - \mu\|_2^2$, all minimizers of (1) share the same image under Φ. Therefore any minimizer of (1) takes the form $x^\star + h$ where $h \in \ker(\Phi)$. Furthermore, it can be shown by arguments from convex analysis that any proper convex function R has a unique minimizer $x^\star$ (if any) over a convex set C if its directional derivative satisfies

$$R'(x^\star; x - x^\star) > 0, \quad x \in C, x \neq x^\star \ .$$

Applying this to (1) with $C = x^\star + \ker(\Phi)$, and using the fact that the directional derivative is the support function of the subdifferential, we get that $x^\star$ is the unique minimizer of (1) if $\forall\, h \in \ker(\Phi) \setminus \{0\}$

$$\begin{aligned} 0 < R'(x^\star; h) &= \sup_{v \in \partial R(x^\star)} \langle v,\, h\rangle \\ &= \sup_{\alpha \in \partial\|\cdot\|_{\mathcal{A}}(L^* x^\star)} \langle \alpha,\, L^* h\rangle \\ &= \langle e,\, L_T^* h\rangle + \sup_{\|\alpha_S\|_{\mathcal{A}}^* \leqslant 1} \langle \alpha_S,\, L_S^* h\rangle \\ &= \langle e,\, L_T^* h\rangle + \|L_S^* h\|_{\mathcal{A}} \ . \end{aligned}$$

We conclude using symmetry of the norm and the fact that $\ker(\Phi)$ is a subspace. ∎

B. Proof of Corollary 1

The source condition ($\mathrm{SC}_{x^\star}$) implies that $\forall\, h \in \ker(\Phi) \setminus \{0\}$

$$\langle h,\, L\alpha\rangle = \langle h,\, \Phi^*\eta\rangle = \langle \Phi h,\, \eta\rangle = 0 \ .$$

Moreover

$$\langle h,\, L\alpha\rangle = \langle L^* h,\, \alpha\rangle = \langle L_T^* h,\, e\rangle + \langle L_S^* h,\, \alpha_S\rangle \ .$$

Thus, applying the dual-norm inequality we get

$$\langle L_T^* h,\, e\rangle \leqslant \|L_S^* h\|_{\mathcal{A}} \|\alpha_S\|_{\mathcal{A}}^* < \|L_S^* h\|_{\mathcal{A}} \ ,$$

where the last inequality is strict since $L_S^* h$ does not vanish owing to (INJ_T), and $\|\alpha_S\|_{\mathcal{A}}^* < 1$. ∎

C. Proof of Corollary 2

We follow the same lines as the proof of Corollary 1 and get

$$\langle L^* h,\, \alpha\rangle = \langle L_T^* h,\, e\rangle + \langle L_V^* h,\, \alpha_V\rangle + \langle L_W^* h,\, \alpha_W\rangle \ .$$

We therefore obtain

$$\begin{aligned} \langle L_T^* h,\, e\rangle &\leqslant \|L_V^* h\|_{\mathcal{A}} \|\alpha_V\|_{\mathcal{A}}^* + \|L_W^* h\|_{\mathcal{A}} \|\alpha_W\|_{\mathcal{A}}^* \\ &< \|L_V^* h\|_{\mathcal{A}} + \|L_W^* h\|_{\mathcal{A}} = \|L_S^* h\|_{\mathcal{A}}^* \ , \end{aligned}$$

where we used that $h \notin \ker(L_V^*)$, $\|\alpha_V\|_{\mathcal{A}}^* < 1$, separability and $\|\alpha_W\|_{\mathcal{A}}^* \leqslant \|\alpha_V\|_{\mathcal{A}}^* + \|\alpha_W\|_{\mathcal{A}}^* = \|\alpha_S\|_{\mathcal{A}}^* \leqslant 1$. ∎

D. Proof of Theorem 2

We first define the Bregman distance/divergence.

Definition 4. *Let $D_s^R(x, x_0)$ be the Bregman distance associated to R with respect to $s \in \partial R(x_0)$,*

$$D_s^R(x, x_0) = R(x) - R(x_0) - \langle s,\, x - x_0\rangle \ .$$

Define $D_\alpha^{\mathcal{A}}(u, u_0)$ as the Bregman distance associated to $\|\cdot\|_{\mathcal{A}}$ with respect to $\alpha \in \partial\|\cdot\|_{\mathcal{A}}(u_0)$.

Observe that by convexity, the Bregman distance is non-negative.

Preparatory lemmata We first need the following key lemmata.

Lemma 1 (Prediction error and Bregman distance convergence rates). *Suppose that* (SC_{x_0}) *is satisfied. Then, for any minimizer $x^\star$ of* (1), *and with $\lambda = c\varepsilon$ for $c > 0$, we have*

$$\begin{aligned} D_{\Phi^*\eta}^R(x^\star, x_0) = D_\alpha^{\mathcal{A}}(L^* x^\star, L^* x_0) &\leqslant \varepsilon \frac{(1 + c\|\eta\|_2/2)^2}{c} \ , \\ \|\Phi x^\star - \Phi x_0\|_2 &\leqslant \varepsilon(2 + c\|\eta\|_2) \ . \end{aligned}$$

The proof follows the same lines as that for any sublinear regularizer, see e.g. [9], where we additionally use the source condition (SC_{x_0}) and $D_{\Phi^*\eta}^R(x, x_0) = D_{L\alpha}^R(x, x_0) = D_\alpha^{\mathcal{A}}(L^* x, L^* x_0)$.

Now since $\|\cdot\|_{\mathcal{A}}$ is a norm, it is coercive, and thus

$$\exists\, C_{\mathcal{A}} > 0 \quad \text{s.t.} \quad \forall x \in \mathbb{R}^P,\ \|x\|_{\mathcal{A}} \geqslant C_{\mathcal{A}} \|x\|_2.$$

We get the following inequality.

Lemma 2 (From Bregman to ℓ_2 bound)**.** *Suppose that* (SC_{x_0}) *holds with* $\|\alpha_{S_0}\|_{\mathcal{A}}^* < 1$. *Then,*

$$\|L_{S_0}^*(x^\star - x_0)\|_2 \leqslant \frac{D_\alpha^{\mathcal{A}}(L^*x^\star, L^*x_0)}{C_{\mathcal{A}}\,(1 - \|\alpha_{S_0}\|_{\mathcal{A}}^*)} ,$$

Proof: Decomposability of $\|\cdot\|_{\mathcal{A}}$ implies that $\exists v \in S_0$ such that $\|v\|_{\mathcal{A}}^* \leqslant 1$ and $\|L_{S_0}^*(x^\star - x_0)\|_{\mathcal{A}} = \langle L_{S_0}^*(x^\star - x_0),\, v\rangle$. Moreover, $v + e_0 \in \partial\|\cdot\|_{\mathcal{A}}(L^*x_0)$. Thus

$$\begin{aligned}
D_\alpha^{\mathcal{A}}(L^*x^\star, L^*x_0) &\geqslant D_\alpha^{\mathcal{A}}(L^*x^\star, L^*x_0) \\
&\qquad - D_{v+e_0}^{\mathcal{A}}(L^*x^\star, L^*x_0) \\
&= \langle v + e_0 - \alpha,\, L^*(x^\star - x_0)\rangle \\
&= \langle v - \alpha_{S_0},\, L_{S_0}^*(x^\star - x_0)\rangle \\
&= \|L_{S_0}^*(x^\star - x_0)\|_{\mathcal{A}} \\
&\qquad - \langle \alpha_{S_0},\, L_{S_0}^*(x^\star - x_0)\rangle \\
&\geqslant \|L_{S_0}^*(x^\star - x_0)\|_{\mathcal{A}}(1 - \|\alpha_{S_0}\|_{\mathcal{A}}^*) \\
&\geqslant C_{\mathcal{A}}\|L_{S_0}^*(x^\star - x_0)\|_2(1 - \|\alpha_{S_0}\|_{\mathcal{A}}^*) .
\end{aligned}$$

■

Proof of the main result

$$\begin{aligned}
\|x^\star - x_0\|_2 &\leqslant \|\mathcal{P}_{\ker(L_{S_0}^*)}(x^\star - x_0)\|_2 \\
&\qquad + \|\mathcal{P}_{\mathrm{Im}(L_{S_0}^*)}(x^\star - x_0)\|_2 \\
&\leqslant {C_\Phi}^{-1}\|\Phi\mathcal{P}_{\ker(L_{S_0}^*)}(x^\star - x_0)\|_2 \\
&\qquad + \|\mathcal{P}_{\mathrm{Im}(L_{S_0}^*)}(x^\star - x_0)\|_2 \\
&\leqslant {C_\Phi}^{-1}\|\Phi(x^\star - x_0)\|_2 \\
&\qquad + (1 + {C_\Phi}^{-1}\|\Phi\|_{2,2})\|\mathcal{P}_{\mathrm{Im}(L_{S_0}^*)}(x^\star - x_0)\|_2 ,
\end{aligned}$$

where we used assumption (INJ_{T_0}), *i.e.,*

$$\exists\, C_\Phi > 0 \quad \text{s.t.} \quad \|\Phi x\|_2 \geqslant C_\Phi\|x\|_2, \quad \forall x \in \ker(L_{S_0}^*) .$$

Since $L_{S_0}^*$ is injective on the orthogonal of its kernel, there exists $C_L > 0$ such that

$$\begin{aligned}
\|x^\star - x_0\|_2 &\leqslant {C_\Phi}^{-1}\|\Phi(x^\star - x_0)\|_2 \\
&\qquad + \tfrac{\|\Phi\|_{2,2} + C_\Phi}{C_L C_\Phi}\|L_{S_0}^*\mathcal{P}_{\mathrm{Im}(L_{S_0}^*)}(x^\star - x_0)\|_2 .
\end{aligned}$$

Noticing that

$$\|L_{S_0}^*(x^\star - x_0)\|_2 = \|L_{S_0}^*\mathcal{P}_{\mathrm{Im}(L_{S_0}^*)}(x^\star - x_0)\|_2,$$

we apply Lemma 2 to get

$$\begin{aligned}
\|x^\star - x_0\|_2 &\leqslant {C_\Phi}^{-1}\|\Phi(x^\star - x_0)\|_2 \\
&\qquad + \tfrac{\|\Phi\|_{2,2} + C_\Phi}{C_L C_\Phi\left(1 - \|\alpha_{S_0}\|_{\mathcal{A}}^*\right)} D_\alpha^{\mathcal{A}}(L^*x^\star, L^*x_0) .
\end{aligned}$$

Using Lemma 1 yields the desired result. ■

E. Proof of Corollary 3

Take $\alpha = e_0 + \Gamma e_0 + \bar{u}_{S_0} + (L_{S_0})^+\Phi^*\bar{z}$. First, $\alpha_{T_0} = e_0$ since $e_0 \in T_0$ and $\mathrm{Im}(\Gamma) \subseteq \mathrm{Im}((L_{S_0})^+) = \mathrm{Im}(L_{S_0}^*)$. Then $\|\alpha_{S_0}\|_{\mathcal{A}}^* = \mathrm{IC}_{\bar{u},\bar{z}}(T_0, e_0) < 1$, whence we get that $\alpha \in \partial\|\cdot\|_{\mathcal{A}}(L^*x_0)$.

Now, we observe by definition of Ξ that $\mathcal{P}_{\ker(L_{S_0}^*)}(\Phi^*\Phi\Xi - \mathrm{Id})L_{T_0} = 0$, which implies that $\mathrm{Im}((\Phi^*\Phi\Xi - \mathrm{Id})L_{T_0})) \subseteq \mathrm{Im}(L_{S_0})$. In turn, $L_{S_0}\Gamma = L_{S_0}(L_{S_0})^+\left((\Phi^*\Phi\Xi - \mathrm{Id})L_{T_0}\right) = \mathcal{P}_{\mathrm{Im}(L_{S_0})}\left((\Phi^*\Phi\Xi - \mathrm{Id})L_{T_0}\right) = (\Phi^*\Phi\Xi - \mathrm{Id})L_{T_0}$. This, together with the fact that $\bar{u} \in \ker(L_{S_0})$ and $\Phi^*\bar{z} \in \mathrm{Im}(L_{S_0})$ yields

$$\begin{aligned}
L_{S_0}\alpha &= (\Phi^*\Phi\Xi - \mathrm{Id})L_{T_0}e_0 + \Phi^*\bar{z} \\
&= \Phi^*\eta - L_{T_0}\alpha \iff \Phi^*\eta = L\alpha ,
\end{aligned}$$

which implies that $\Phi^*\eta = L\alpha \in \partial R(x_0)$. We have just shown that the vectors α and η as given above satisfy the source condition (SC_{x_0}) and the dual non-saturation condition. We conclude by applying Theorem 2 using (INJ_{T_0}). ■

V. Conclusion

We provided a unified analysis of the structural properties of regularized solutions to linear inverse problems through a class of semi-norms formed by composing decomposable norms with a linear operator. We provided conditions that guarantee uniqueness, and also those ensuring stability to bounded noise. The stability bound was achieved without requiring (even partial) recovery of T_0 and e_0. Recovery of T_0 and e_0 for analysis-type decomposable priors and beyond is currently under investigation. Another perspective concerns whether the ℓ_2 bound on $x^\star - x_0$ can be extended to cover more general low complexity-inducing regularizers beyond decomposable norms.

References

[1] E. J. Candès and B. Recht. Simple bounds for recovering low-complexity models. *Mathematical Programming*, pages 1–13, 2012.

[2] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[3] V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational Mathematics*, 12:805–849, 2012.

[4] D. L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 47(7):2845–2862, 2001.

[5] J.-J. Fuchs. On sparse representations in arbitrary redundant bases. *IEEE Trans. Info. Theory*, 50(6):1341–1344, 2004.

[6] M. Grasmair. Linear convergence rates for Tikhonov regularization with positively homogeneous functionals. *Inverse Problems*, 27:075014, 2011.

[7] M. Haltmeier. Stable signal reconstruction via ℓ^1-minimization in redundant, non-tight frames. *IEEE Trans. on Sig. Proc.*, 2012. to appear.

[8] S. Negahban, P. Ravikumar, M. J. Wainwright, and B. Yu. A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers. *Statistical Science*, 27(4):538–557, December 2012.

[9] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Applied Mathematical Sciences. Springer, 1st edition, 2009.

[10] S. Vaiter, G. Peyré, C. Dossal, and M.J. Fadili. Robust sparse analysis regularization. *to appear in IEEE Trans. Inf. Theo.*, 2012.