「『特集名』特集号」 解　説

Plug-and-Play アルゴリズムによる信号復元 – モデルベースとデータドリブンの止揚 –

小野　峻佑*

1. はじめに

計測技術は人類文明の発展を支える不動の礎である．今日では，観測衛星から電子顕微鏡まで，ありとあらゆる時空間スケール・ドメインで高度な計測が可能となっている．しかし，計測技術・対象が先鋭化するほど，計測された生データには物理的に不可避な劣化ーノイズ・外乱・欠損・外れ値ーがつきまとう．さらに，環境的・資源的制約により，連続的な物理量を粗くサンプリングしたものしか得られない状況も多い（例：ハイパースペクトルイメージング [1]）．加えて，計測方式に起因する何らかの変換が避けられないケースもしばしばある（例：CT[2]）．故に，様々な計測データを科学的・社会的に利活用するためには「生データから劣化を取り除きつつ背後の信号情報を復元する技術ー信号復元（signal recovery）ー」が不可欠となる．

信号復元技術には，大きく分けて，モデルベースアプローチとデータドリブンアプローチが存在する．前者は，「対象の信号情報に関する事前知識」と「計測モデルに基づくデータ整合性」をそれぞれ数理的指標（正則化項およびデータ忠実項）として表現し，これを取り入れた最適化問題を解くことで信号を復元する．スパースモデリングなどが代表的であり，事前知識を人の手で適切にモデル化できる場合は，計測モデルを柔軟に反映した説明性・妥当性の高い復元が実現できる．他方，後者は，計測データ（入力）と完全な信号情報（出力）のペアを複数用意し，これを教師データとして信号復元器を学習する．こちらは深層学習に代表され，ブラックボックスかつドメイン・タスク特化であるものの，豊富な教師データから自動的に獲得した事前知識に基づく高精度な復元が可能となる．ここで自然な発想が浮かぶ：両者のメリットを併せ持った信号復元技術を構築できないか？

そのような技術のひとつとして注目を集めているのが，Plug-and-Play（PnP）アルゴリズムである [3,4]．PnP アルゴリズムは，モデルベース信号復元の基盤となっている凸最適化技術である「近接分離アルゴリズム（proximal splitting algorithms）[5]」に，ノイズ除去器を挿入することで実現される．ここで，このノイズ除去器は，対象としている信号情報から人工的に加えたノイズ（多くの場合，白色ガウス性）を除去するように学習されており，その過程で信号情報に関する事前知識を暗に獲得していると想定される．[1]後ほど詳しく説明するが，PnP アルゴリズムにおいて，挿入されたノイズ除去器は正則化項（から導かれる更新式）のような役割を果たす．一方，その他のステップは元の最適化アルゴリズムをそのまま継承しているため，利用したいデータ忠実項に合わせて適切に更新式を設計・変更できる．故に，PnP アルゴリズムは，モデルベースとデータドリブン双方のアプローチのメリット（計測モデルの明示的かつ柔軟な反映とデータからの自動的な事前知識獲得）を有したハイブリッドな信号復元技術となっている．

ここまでの説明で，PnP アルゴリズムについて次のような疑問を抱いた読者も多いだろう．

- どうやって具体的なアルゴリズムを導出するのか？
- 実際の信号復元応用においてどの程度有効なのか？
- 収束性などの理論的保証はどうなっているのか？

これらの疑問に答えながら PnP アルゴリズムについて理解を深めてもらうことが本稿の目的である．そのために，まず，PnP アルゴリズム設計の鍵となっている「近接写像とノイズ除去器の関係」について説明した後に，いくつかの具体的な PnP アルゴリズムの導出と応用について概観する．さらに，PnP アルゴリズムが安定して収束するための数学的条件やその収束先の性質について，最新の研究動向を参考に解説を試みる．

2. 数学的準備

以下，ベクトルのユークリッドノルム（ℓ_2 ノルム）を $\|\cdot\|$，ベクトルや行列の転置を $(\cdot)^\top$ で表す．

関数 $f:\mathbb{R}^N \to (-\infty,\infty]$ が，任意の $\mathbf{x},\mathbf{y}\in\mathbb{R}^N$ と $\alpha\in[0,1]$ に対して $f(\alpha\mathbf{x}+(1-\alpha)\mathbf{y})\leq\alpha f(\mathbf{x})+(1-\alpha)f(\mathbf{y})$ を満たすとき，f を凸関数という．凸関数 f の実効定義域 $\mathrm{dom}(f):=\{\mathbf{x}\in\mathbb{R}^N \mid f(\mathbf{x})<\infty\}$ が空集合でないとき，f を真凸関数と呼ぶ．真凸関数 f のレベル集

[1] 原理的には，PnP アルゴリズムに挿入できるノイズ除去器の種類に制限はないが，理論的にも実用的にも学習によって適切に構成したノイズ除去器が望ましい（詳しくは 4. 節を参照）．

* 東京工業大学　情報理工学院

Key Words: style file, transactions, ISCIE, paper

合 $\mathrm{lev}_{\le\alpha}(f) := \{\mathbf{x}\in\mathbb{R}^N \,|\, f(\mathbf{x})\le\alpha\}$ が任意の $\alpha\in\mathbb{R}$ について閉集合となるとき f を下半連続な真凸関数という．以降，$\mathbb{R}^N$ 上の全ての下半連続な真凸関数の集合を $\Gamma_0(\mathbb{R}^N)$ で表す．関数 $f\in\Gamma_0(\mathbb{R}^N)$ に対して定義される $f^*(\mathbf{y}) := \sup_{\mathbf{x}\in\mathbb{R}^N}\{\langle\mathbf{y},\mathbf{x}\rangle - f(\mathbf{x})\}$（$\langle\cdot,\cdot\rangle$ はユークリッド内積）を f の凸共役関数といい，$f^*\in\Gamma_0(\mathbb{R}^N)$ となる．

集合 $C\subset\mathbb{R}^N$ が任意の $\mathbf{x},\mathbf{y}\in C$ と $\alpha\in[0,1]$ に対して $\alpha\mathbf{x}+(1-\alpha)\mathbf{y}\in C$ を満たすとき，集合 C を凸集合という．また，凸集合が閉集合であるとき，閉凸集合という．空でない閉凸集合 C に対して指示関数（indicator function）$\iota_C:\mathbb{R}^N\to(-\infty,\infty]$ を

$$\iota_C(\mathbf{x}) := \begin{cases} 0, & \text{if } \mathbf{x}\in C, \\ \infty, & \text{otherwise} \end{cases} \tag{1}$$

のように定義すると，$\iota_C\in\Gamma_0(\mathbb{R}^N)$ となる．

任意の $\gamma>0$ に対し，関数 $g\in\Gamma_0(\mathbb{R}^N)$ の近接写像（proximity operator/proximal mapping）は

$$\mathrm{prox}_{\gamma g}(\mathbf{x}) := \arg\min_{\mathbf{y}\in\mathbb{R}^N} g(\mathbf{y}) + \frac{1}{2\gamma}\|\mathbf{x}-\mathbf{y}\|^2 \tag{2}$$

で定義される．また，g^* の近接写像は，以下のように g の近接写像を用いて表すことができる．

$$\mathrm{prox}_{\gamma g^*}(\mathbf{x}) = \mathbf{x} - \gamma\,\mathrm{prox}_{\frac{1}{\gamma}g}(\tfrac{1}{\gamma}\mathbf{x}) \tag{3}$$

定義 (2) において $g := \iota_C$ とすると，任意の $\gamma>0$ に対して，

$$\begin{aligned}\mathrm{prox}_{\gamma\iota_C}(\mathbf{x}) &= \mathop{\mathrm{argmin}}_{\mathbf{y}\in C} \tfrac{1}{2\gamma}\|\mathbf{x}-\mathbf{y}\|^2 \\ &= \mathop{\mathrm{argmin}}_{\mathbf{y}\in C}\|\mathbf{x}-\mathbf{y}\| =: P_C(\mathbf{x})\end{aligned} \tag{4}$$

となり，C への距離射影（metric projection）と呼ばれる操作（P_C で表される）と一致する．

3. PnP アルゴリズムの設計と応用

3.1 近接写像とノイズ除去

近接写像の定義 (2) をみると，$\mathrm{prox}_{\gamma g}(\mathbf{x})$ は，入力 $\mathbf{x}$ と ℓ_2 ノルムの意味で近いベクトルの中から関数 g をできる限り小さくするものを出力する操作になっている．ここで，関数 g が信号情報にとって何らかの望ましい性質（例えばスパース性や滑らかさ等）をモデル化していると仮定する．このとき，“$\mathrm{prox}_{\gamma g}(\mathbf{x})$ は，ノイズを含む入力データ $\mathbf{x}$ から関数 g がモデル化している性質を有する信号情報を推定する操作＝一種のノイズ除去器” であると考えることができる．

統計的推定の観点からより詳細に説明すると，

- $g(\mathbf{y})$：信号情報 $\mathbf{y}$ に関する何らかの事前知識（分布）から導かれる関数
- $\frac{1}{2\gamma}\|\mathbf{x}-\mathbf{y}\|^2$：入力データ $\mathbf{x}$ と信号情報 $\mathbf{y}$ の差であるノイズ成分が平均 0，分散 γ のガウス分布に従うと仮定したときの負の対数尤度

と捉えることで「近接写像は MAP 推定によって入力データから分散 γ の白色ガウス性ノイズを除去している」と解釈できる．これは，逆に言えば，（粗い議論ではあるが）任意のノイズ除去器 $\mathcal{D}_\gamma:\mathbb{R}^N\to\mathbb{R}^N$ を近接写像とみなせることを意味する．つまり，$\mathcal{D}_\gamma$ が想定している事前知識と紐づいた関数 $R_{\mathcal{D}_\gamma}$ の存在を仮定した上で，以下の近似関係が成立すると考えるわけである．

ノイズ除去器の近接写像的解釈

$$\mathcal{D}_\gamma(\mathbf{x}) \approx \mathrm{prox}_{\gamma R_{\mathcal{D}_\gamma}}(\mathbf{x}) \tag{5}$$

本節では以降，この解釈のもと，三種類の代表的な近接分離アルゴリズムを原型としている PnP アルゴリズムについて，それぞれが解いている（と想定している）最適化問題を起点に具体的に解説する．[1]

3.2 近接勾配法･交互方向乗数法（ADMM）型

以下のような線形計測モデルを考える．

$$\mathbf{v} = \mathbf{\Phi}\bar{\mathbf{u}} + \mathbf{n} \tag{6}$$

ここで，$\mathbf{v}\in\mathbb{R}^M$ は計測データ，$\mathbf{\Phi}\in\mathbb{R}^{M\times N}$ は計測過程，$\bar{\mathbf{u}}\in\mathbb{R}^N$ は真の信号情報，$\mathbf{n}\in\mathbb{R}^M$ はノイズを表す．

このモデルのもと，計測データから信号情報を復元するために，次のような最適化問題を考える．

最小二乗 PnP 信号復元問題

$$\min_{\mathbf{u}\in\mathbb{R}^N} \frac{\lambda}{2}\|\mathbf{\Phi}\mathbf{u}-\mathbf{v}\|^2 + R_{\mathcal{D}_\gamma}(\mathbf{u}) \tag{7}$$

ただし，$\lambda>0$ はバランスパラメータである．上記は，典型的な正則化付き最小二乗問題の形をしており，実際に，第二項を ℓ_1 ノルムにすれば圧縮センシングにおけるスパース信号再構成（またはLASSO），信号値の差分の混合 $\ell_{1,2}$ ノルムにすれば画像復元に幅広く使われている全変動最小化になる．そのため，様々なPnP アルゴリズムの研究においても，この定式化を扱っているものが非常に多い．

問題 (7) を対象とした PnP アルゴリズムは，大きく分けて二種類存在する．ひとつめは近接勾配法 [8] に基づくものである．近接勾配法は最も基本的な近接分離アルゴリズムであり，特に ℓ_1 ノルム最小化に適用される際には ISTA[9] と呼ばれる．近接勾配法は以下の凸最適化問題を解くことができる．

$$\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x}) + g(\mathbf{x}) \tag{8}$$

[1] もちろん，本稿で扱わない型の PnP アルゴリズムも存在する（例えば近接分離アルゴリズムではない最適化技術に基づいているもの [6] や，ノイズ除去器を近接写像と異なる操作として解釈しているもの [7] 等）．本稿では，PnP アルゴリズムに関する研究の流れを理解する上で重要となるものに絞って議論を進める．

ただし，$f,g \in \Gamma_0(\mathbb{R}^N)$ であり，f は微分可能かつその勾配 ∇f が β-リプシッツ連続[2]である．近接勾配法の更新式は以下のとおりである（$\gamma > 0$）．

$$\mathbf{x}^{(n+1)} = \mathrm{prox}_{\gamma g}(\mathbf{x}^{(n)} - \gamma \nabla f(\mathbf{x}^{(n)})) \tag{9}$$

近接勾配法型 PnP アルゴリズム（PnP-PG）は，問題 (8) において，$f := \frac{\lambda}{2}\|\mathbf{\Phi}\cdot - \mathbf{z}\|^2$ および $g := R_{\mathcal{D}_\gamma}$ とし，式 (5) を用いることで，直ちに導かれる．

PnP-PG

$$\mathbf{u}^{(n+1)} = \mathcal{D}_\gamma(\mathbf{u}^{(n)} - \gamma\lambda\mathbf{\Phi}^\top(\mathbf{\Phi}\mathbf{u}^{(n)} - \mathbf{v})) \tag{10}$$

もうひとつは交互方向乗数法（ADMM）[10][1]に基づくものである．ADMM は近接分離アルゴリズムの中でも非常にポピュラーであり，信号復元以外にも，画像の成分分離 [12,13]，非凸関数・制約を用いた画像処理 [14,15]，分散推定 [16,17] など様々な応用がある．ADMM は以下の凸最適化問題[2]を対象にしている．

$$\min_{\mathbf{x}\in\mathbb{R}^N, \mathbf{z}\in\mathbb{R}^K} f(\mathbf{x}) + g(\mathbf{z}) \ \text{s.t.}\ \mathbf{z} = \mathbf{A}\mathbf{x} \tag{11}$$

ただし，$f \in \Gamma_0(\mathbb{R}^N)$，$g \in \Gamma_0(\mathbb{R}^K)$，$\mathbf{A} \in \mathbb{R}^{K\times N}$ である．ADMM の更新式は以下で与えられる（$\gamma > 0$）．

$$\mathbf{x}^{(n+1)} = \operatorname*{argmin}_{\mathbf{x}} f(\mathbf{x}) + \tfrac{1}{2\gamma}\|\mathbf{z}^{(n)} - \mathbf{A}\mathbf{x} - \mathbf{y}^{(n)}\|^2 \tag{12}$$
$$\mathbf{z}^{(n+1)} = \mathrm{prox}_{\gamma g}(\mathbf{A}\mathbf{x}^{(n+1)} + \mathbf{y}^{(n)}) \tag{13}$$
$$\mathbf{y}^{(n+1)} = \mathbf{y}^{(n)} + \mathbf{A}\mathbf{x}^{(n+1)} - \mathbf{z}^{(n+1)} \tag{14}$$

ADMM 型 PnP アルゴリズム（PnP-ADMM）は，問題 (11) において，$f := \frac{\lambda}{2}\|\mathbf{\Phi}\cdot - \mathbf{z}\|^2$，$g := R_{\mathcal{D}_\gamma}$，および $\mathbf{A} := \mathbf{I}$ とすることで以下のように導出される．

PnP-ADMM

$$\mathbf{u}^{(n)} \leftarrow \gamma\lambda\mathbf{\Phi}^\top\mathbf{v} + \mathbf{z}^{(n)} - \mathbf{y}^{(n)} \tag{15}$$
$$\mathbf{u}^{(n+1)} = (\gamma\lambda\mathbf{\Phi}^\top\mathbf{\Phi} + \mathbf{I})^{-1}\mathbf{u}^{(n)} \tag{16}$$
$$\mathbf{z}^{(n+1)} = \mathcal{D}_\gamma(\mathbf{u}^{(n+1)} + \mathbf{y}^{(n)}) \tag{17}$$
$$\mathbf{y}^{(n+1)} = \mathbf{y}^{(n)} + \mathbf{u}^{(n+1)} - \mathbf{z}^{(n+1)} \tag{18}$$

PnP アルゴリズムのアイデアが初めて登場した論文 [18] で ADMM が採用されていたため，後に登場する多くの PnP アルゴリズムが ADMM 型となっている [19–22]．他方，近接勾配法型は，元のアルゴリズムが単純であるため，PnP アルゴリズムの収束解析（4. 節で詳しく説明する）を目的とした研究においてしばしば扱われる [23–25]．PnP-PG および PnP-ADMM は，直感的に理解しやすく応用対象もある程度広いが，以下に挙げるふたつの制限を抱えている．

ひとつめは，データ忠実項が二乗誤差に限定されることである．信号復元応用では，しばしば，非ガウス性のノイズが登場する（例えば，低光量環境下でのイメージングではポアソン性，合成開口レーダによるイメージングでは乗算性のノイズがそれぞれ現れる）．このようなノイズを含む計測データを扱う場合，二乗誤差ではなく各々のノイズの統計的性質に基づいた関数をデータ忠実項として用いるべきである．また，ガウス性のノイズとしてモデル化できる場合であっても，二乗誤差項を目的関数（の一部）ではなく制約条件として表現するほうが，パラメータ設定の簡易化や復元結果の妥当性の観点から望ましいことが知られている [26,27]．しかし，このような関数や制約条件（の指示関数）は，線形写像が合成された非可微分凸関数となるため，勾配はもちろんのこと，そのままでは近接写像も効率的に計算できない．結果として，データ忠実項の勾配計算が必須である PnP-PG では対応できず，また，PnP-ADMM も（適用できたとしても）近接写像の計算に内部ループが必要となるため，計算コストや収束性の観点で望ましくない．[3]

ふたつめは，対象の信号情報に関する付加的な制約条件（非負性，Box 制約，ノルム制限等）を利用できないことである．これは，PnP-PG や PnP-ADMM が，「二乗誤差項＋ PnP 正則化項」という定式化に焦点を当てて設計されていることに起因する．[4]

3.3 主-双対近接分離法（PDS）型

前述した制限を解消するために，まずは次のような最適化問題を考える．

一般化 PnP 信号復元問題

$$\min_{\mathbf{u}} R_{\mathcal{D}_{\gamma_1}}(\mathbf{u}) + F_{\mathbf{v}}(\mathbf{\Phi}\mathbf{u}) \ \text{s.t.}\ \mathbf{u} \in C \tag{19}$$

ただし，$F_{\mathbf{v}}(\mathbf{\Phi}\cdot) \in \Gamma_0(\mathbb{R}^N)$ は一般化されたデータ忠実項，$C \subset \mathbb{R}^N$ は付加的な制約条件を表す閉凸集合である．ここで，$F_{\mathbf{v}} \in \Gamma_0(\mathbb{R}^M)$ の近接写像および C への凸射影は効率的に計算できると仮定する．[5]また，便宜上，ノイズ除去器が仮定しているガウス性ノイズの分散を γ ではなく γ_1 と表記している．

---

[2]任意の $\mathbf{x},\mathbf{y} \in \mathbb{R}^N$ に対して，ある $\beta > 0$ が存在し，写像 $T : \mathbb{R}^N \to \mathbb{R}^N$ が $\|T(\mathbf{x}) - T(\mathbf{y})\| \le \beta\|\mathbf{x} - \mathbf{y}\|$ を満たすとき，T は β-リプシッツ連続であるという．

[1]分離ブレグマン法（split Bregman method）[11] と呼ばれることもある．

[2]一般的な ADMM の定式化と比較して，線形制約を若干簡略化している．

[3]実際，文献 [21] ではポアソン性のノイズが重畳した画像の復元問題に対する PnP-ADMM を提案しているが，式 (12) を計算する際に反復法を利用している．

[4]厳密に述べると，PnP-ADMM は，変数分離を工夫することである程度この制限を緩和できる．しかし，その場合でも，内部ループが必要なことは変わらない上に，部分問題 (12) の次元と複雑さが増すため，計算コストや収束性の問題がより顕著になる．

[5]ある関数 $g \in \Gamma_0(\mathbb{R}^N)$ の近接写像が効率的に計算可能であっても，g に線形写像が合成された関数の近接写像は一般に計算可能ではないことに注意されたい．

上記の問題を扱える PnP アルゴリズムを設計するために，汎用的な近接分離アルゴリズムのひとつである主-双対近接分離法（PDS）[28] を紹介する．PDS は，以下の最適化問題を解くことができる．

$$\min_{\mathbf{x}\in\mathbb{R}^N} f(\mathbf{x})+g(\mathbf{x})+h(\mathbf{A}\mathbf{x}) \tag{20}$$

ただし，$f,g\in\Gamma_0(\mathbb{R}^N)$，$h\in\Gamma_0(\mathbb{R}^K)$，$\mathbf{A}\in\mathbb{R}^{K\times N}$ であり，f は微分可能かつその勾配 ∇f が β-リプシッツ連続であるとする（$\beta>0$）．PDS の更新式は以下で与えられる（$\gamma_1,\gamma_2>0$）．

$$\mathbf{x}^{(n+1)}=\mathrm{prox}_{\gamma_1 g}(\mathbf{x}^{(n)}-\gamma_1(\nabla f(\mathbf{x}^{(n)})+\mathbf{A}^\top\mathbf{y}^{(n)})) \tag{21}$$
$$\mathbf{y}^{(n+1)}=\mathrm{prox}_{\gamma_2 h^*}(\mathbf{y}^{(n)}+\gamma_2\mathbf{A}(2\mathbf{x}^{(n+1)}-\mathbf{x}^{(n)})) \tag{22}$$

他の近接分離アルゴリズムにはない PDS の大きな特徴は，最適化問題中の行列 $\mathbf{A}$ に依存する部分問題（ADMM における式 (12)）が完全に回避されていることである．各ステップをみればわかるように，f の勾配降下，g と h^* の近接写像（h^* の近接写像は (3) により h のそれを用いて計算可能），および $\mathbf{A}$ と $\mathbf{A}^\top$ のみでアルゴリズムが構成されている．これにより，$\mathbf{A}$ に依存した部分問題解法のアドホックな設計や計算コストの増加に悩まされることなく，変数分離テクニック（詳しくは文献 [5] を参照）によって発展的な正則化や複数の制約条件を見通しよく扱えるため，PDS は非常に汎用的で効率的な近接分離アルゴリズムになっている．これらの利点は対象の信号情報（の構造）が高次元化・複雑化するほど効いてくるため，そのような応用（衛星リモートセンシングデータの解析や X 線/CT 材料計測データの再構成等）において実際に PDS が有効活用されている [29–34]．また，PDS の利便性をより高めるため，適切なステップサイズを問題構造から自動的に決定する手法も提案されている [35,36]．

本題に戻ろう．PDS 型の PnP アルゴリズム（PnP-PDS）は，問題 (19) において，

$$f(\mathbf{u}):=0,\ g(\mathbf{u}):=R_{\mathcal{D}_{\gamma_1}}(\mathbf{u}) \tag{23}$$
$$h(\mathbf{y}):=F_{\mathbf{v}}(\mathbf{y}_1)+\iota_C(\mathbf{y}_2) \tag{24}$$
$$\mathbf{y}:=\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{pmatrix},\ \mathbf{A}:=\begin{pmatrix}\mathbf{\Phi}\\ \mathbf{I}\end{pmatrix} \tag{25}$$

と設定し，式 (3) および (4) を用いることで導かれる．

PnP-PDS

$$\mathbf{u}^{(n+1)}=\mathcal{D}_{\gamma_1}(\mathbf{u}^{(n)}-\gamma_1(\mathbf{\Phi}^\top\mathbf{y}_1^{(n)}+\mathbf{y}_2^{(n)})) \tag{26}$$
$$\mathbf{y}_1^{(n)}\leftarrow\mathbf{y}_1^{(n)}+\gamma_2\mathbf{\Phi}(2\mathbf{u}^{(n+1)}-\mathbf{u}^{(n)}) \tag{27}$$
$$\mathbf{y}_2^{(n)}\leftarrow\mathbf{y}_2^{(n)}+\gamma_2(2\mathbf{u}^{(n+1)}-\mathbf{u}^{(n)}) \tag{28}$$
$$\mathbf{y}_1^{(n+1)}=\mathbf{y}_1^{(n)}-\gamma_2\mathrm{prox}_{\frac{1}{\gamma_2}F_{\mathbf{v}}}(\tfrac{1}{\gamma_2}\mathbf{y}_1^{(n)}) \tag{29}$$
$$\mathbf{y}_2^{(n+1)}=\mathbf{y}_2^{(n)}-\gamma_2 P_C(\tfrac{1}{\gamma_2}\mathbf{y}_2^{(n)}) \tag{30}$$

PnP-PDS は，前述した PnP-PG および PnP-ADMM の制限を解消するために提案された [37]．この文献で提案された PnP-PDS は，厳密には Condat の PDS[28] ではなく，その前身である Chambolle-Pock の PDS[38] に基づいている．両者の本質的な違いは微分可能項 f の有無だけであるため，どちらから導出しても上記の PnP-PDS が得られる．また，Condat の PDS をベースにした，より一般的かつ収束保証付きの PnP-PDS が最近提案されている [39,40]（後ほど紹介）．

3.4 画像復元への応用

実際の信号復元における PnP アルゴリズムの有効性を確認するために，自然画像の復元問題に適用した例を紹介する．具体的には，線形計測モデル (6) の典型例である以下の画像劣化モデルを考える．

$$\mathbf{v}=\mathbf{B}\mathbf{u}+\mathbf{n} \tag{31}$$

ここで，$\mathbf{v}$ は劣化画像，$\mathbf{B}$ はボケ作用素，$\mathbf{u}$ は真の画像，$\mathbf{n}$ はガウス性ノイズをそれぞれ表す．

本節では二種類の PnP 画像復元問題を扱う．ひとつは問題 (7) において $\mathbf{\Phi}:=\mathbf{B}$ としたものであり，こちらは PnP-PG または PnP-ADMM によって解かれる．もうひとつは以下のように定式化される．

$$\min_{\mathbf{u}} R_{\mathcal{D}_{\gamma_1}}(\mathbf{u})\ \text{s.t.}\ \begin{cases}\|\mathbf{B}\mathbf{u}-\mathbf{v}\|\leq\varepsilon\\ \mathbf{u}\in[0,255]^N\end{cases} \tag{32}$$

ただし，$\varepsilon\geq 0$ は $\mathbf{v}$ との許容誤差を定めるパラメータである．問題 (7) との大きな違いとして，データ忠実性が制約条件で表現されていること，および画素値のダイナミックレンジ制約が課されていることが挙げられる．前者は，問題 (7) における PnP 正則化項との相対的なバランスを調整するパラメータ λ を，ノイズの強度（分散）から直接定められるパラメータ ε に変換する役割を果たしており，パラメータ設定が格段に簡易化される．後者は画像データを扱う上で自然な制約条件であり，アーティファクトの抑制やコントラストの維持などの効果がある．問題 (32) は，問題 (19) において，$F_{\mathbf{v}}:=\iota_{\{\|\cdot-\mathbf{v}\|\leq\varepsilon\}}$，$\mathbf{\Phi}:=\mathbf{B}$，および $C:=[0,255]^N$ とすることで，PnP-PDS によって解くことができる．

図 1 に結果画像の例を示す．各画像の下に記載している数値は PSNR[dB] である．なお，ノイズ除去器は文献 [24] にある手順に従って ImageNet[41] によって学習した DnCNN[42] を用いている．全変動最小化（モデルベース）と比べて，PnP アルゴリズムのほうがより鮮明な画像を推定できている．これは，ノイズ除去器が膨大な学習データから獲得した事前知識を，PnP アルゴリズムが上手く他の画像復元問題へと転用していることを示している．また，PnP アルゴリズム同士の結果を比較すると，PnP-PDS（問題 (32)）のほうが，画像全体のコントラストを高精度に復元でき

原画像　劣化画像（21.29）　全変動（30.53）　PnP-PG（33.47）　PnP-PDS（33.97）

第 1 図　PnP アルゴリズムによる画像復元結果

ている．これは，ダイナミックレンジ制約の恩恵であると考えられる．[1]客観評価指標（PSNR）においても，PnP-PDS が最も優れていることが確認できる．

なお，PnP アルゴリズムは，画像復元以外にも，超解像 [43] や画像合成 [44] など，画像データを中心に幅広く応用されている．原理的にはその他の信号情報へ適用することも可能だが，PnP アルゴリズムが生まれた経緯やノイズ除去器に関する研究の進展が背景にあるため，画像データを扱う応用が群を抜いて多い．

4.　PnP アルゴリズムの収束性

さて，積み残した最も重要な課題である「PnP アルゴリズムの収束条件」について考えていこう．PnP アルゴリズムはノイズ除去器を形式的に最適化アルゴリズムに挿入したものであるため，当然ながらその収束は一般には保証できない．これは，利便性・信頼性の観点で大きな懸念事項である．そこで，「PnP アルゴリズムが安定して収束するために必要なノイズ除去器が満たすべき条件」についての研究が進められてきた．これまで主に議論されてきた条件は（厳密性を多少欠くが）以下の二種類にまとめられる．

- 恒等写像に十分に近い [20,23].
- ある微分可能な凸関数の勾配になっている [19,25].

しかし，これらの条件は数学的に厳しすぎるため，実際に性能が高いノイズ除去器を構成（学習）することを考えるとあまり現実的ではない．加えて，PnP アルゴリズムの収束自体は保証できるものの，その収束先がどのような解であるかについては何も説明できない．

最近になり，Pesquet らの研究 [24] によって，より現実的な条件および PnP-PG の収束先を数学的に表現できることが明らかになった．さらに，文献 [39] において，PnP-PDS についても同様の条件で収束することが示されるとともに，収束先の数学的表現が与えられた．以下では，これらの結果を紹介する．

[1]PnP アルゴリズムによる求解プロセスの途中で暫定解（画像）の画素値が負になった際に，挿入したノイズ除去器が予期せぬ挙動を起こし，出力される復元画像の品質が著しく低下するという問題が著者等の実験により確認されている．このような問題はインペインティングや圧縮センシング再構成等の劣決定な問題で生じやすいが，ダイナミックレンジ制約を入れることでこれを回避できることがわかっている [40].

4.1　ノイズ除去器が満たすべき条件

近接分離アルゴリズムの収束原理において，堅非拡大（firmly nonexpansive）と呼ばれるクラスの写像が重要な役割を果たすことが知られている．これまで何度も登場した近接写像もその一種である．ここから自然に導かれる帰結として，以下に示す「ノイズ除去器が堅非拡大性写像となること」が，収束保証付き PnP アルゴリズム（PnP-PG）を構築するための鍵となる．

定義 1 (堅非拡大なノイズ除去器)**.** ノイズ除去器 $\mathcal{D}_\gamma$ が堅非拡大であるとは，任意の $\mathbf{x},\mathbf{y}\in\mathbb{R}^N$ に対して，以下の不等式が成り立つことである．

$$\|T_{\mathcal{D}_\gamma}(\mathbf{x})-T_{\mathcal{D}_\gamma}(\mathbf{y})\|\leq\|\mathbf{x}-\mathbf{y}\| \tag{33}$$

ただし，$T_{\mathcal{D}_\gamma}:=2\mathcal{D}_\gamma-\mathrm{Id}$ であり，Id は恒等写像を表す．

直ちに次の疑問が生じる：どのようにして堅非拡大なノイズ除去器を構成するのか？ Pesquet らは，この要請に対しひとつの処方箋を与えている．まず，ノイズ除去器を学習によって構成する状況を考え，学習可能なパラメータの集合 Θ によって定まるノイズ除去器を $\mathcal{D}_\gamma^\Theta$ で表す．ここで，$T_{\mathcal{D}_\gamma}^\Theta:=2\mathcal{D}_\gamma^\Theta-\mathrm{Id}$ とすると，$T_{\mathcal{D}_\gamma}^\Theta$ が不等式 (33) を満たすように Θ を学習すればよいことになる．いま，$T_{\mathcal{D}_\gamma}^\Theta$ が Θ について微分可能であると仮定すると，以下の (a) と (b) が必要十分の関係にあることが知られている [24].

(a) $T_{\mathcal{D}_\gamma}^\Theta$ のヤコビ行列の最大特異値が 1 以下である．

(b) 不等式 (33) が成立する（堅非拡大）．

そこで Pesquet らは，深層学習によってノイズ除去器を学習する際の損失関数に，条件 (a) を近似した罰則項を加えることを提案している．[1]

4.2　収束定理と収束先

ノイズ除去器の堅非拡大性を踏まえた上で，まずは PnP-PG の収束定理を紹介する．[2]

定理 1 (PnP-PG の収束定理 [24])**.** ノイズ除去器 $\mathcal{D}_\gamma:\mathbb{R}^N\to\mathbb{R}^N$ が堅非拡大であるとする．このとき，

[1]厳密に条件 (a) を担保する学習方法は困難である．

[2]理解しやすさのため，多少簡略化している．

$\mathcal{D}_\gamma=(\mathrm{Id}+M_{\mathcal{D}_\gamma})^{-1}$ となる極大単調写像（maximally monotone operator）$M_{\mathcal{D}_\gamma}:\mathbb{R}^N\to 2^{\mathbb{R}^N}$ が存在する．[3] ここで，次の零点包含問題（zero inclusion）を考える．

$$\text{find } \mathbf{u}^\star \text{ s.t. } \mathbf{0}\in\lambda\mathbf{\Phi}^\top(\mathbf{\Phi}\mathbf{u}^\star-\mathbf{v})+\frac{1}{\gamma}M_{\mathcal{D}_\gamma}(\mathbf{u}^\star) \quad (34)$$

ただし，λ，$\mathbf{\Phi}$，$\mathbf{v}$ は問題 (6) で定義されているものと同一である．さらに，以下を仮定する．

- 問題 (34) の解が空でない．
- PnP-PG のステップサイズ $\gamma>0$ が $\gamma\in(0,\frac{2}{\lambda\sigma_1(\mathbf{\Phi})^2})$ を満たす（$\sigma_1(\mathbf{\Phi})$ は $\mathbf{\Phi}$ の最大特異値）．

このとき，PnP-PG によって生成される点列 $\{\mathbf{u}^{(n)}\}_{n\in\mathbb{N}}$ は問題 (34) の解のひとつに収束する．

PnP-PDS についても次の収束定理が示されている．[1]

定理 2 (PnP-PDS の収束定理 [39,40])**.** 定理 1 と同様に，堅非拡大なノイズ除去器を $\mathcal{D}_{\gamma_1}$，対応する極大単調写像を $M_{\mathcal{D}_{\gamma_1}}$ とする．次の零点包含問題を考える．

$$\text{find } (\mathbf{u}^\star,\mathbf{y}_1^\star,\mathbf{y}_2^\star) \text{ s.t.}$$
$$\begin{pmatrix}\mathbf{0}\\\mathbf{0}\\\mathbf{0}\end{pmatrix}\in\begin{pmatrix}\frac{1}{\gamma_1}M_{\mathcal{D}_{\gamma_1}}(\mathbf{u}^\star)+\mathbf{\Phi}^\top\mathbf{y}_1^\star+\mathbf{y}_2^\star\\ -\mathbf{\Phi}\mathbf{u}^\star+\partial F_{\mathbf{v}}^*(\mathbf{y}_1^\star)\\ -\mathbf{u}^\star+\partial\iota_C^*(\mathbf{y}_2^\star)\end{pmatrix} \quad (35)$$

ただし，$\mathbf{\Phi}$，$F_{\mathbf{v}}$，C は問題 (19) で定義されているものと同一であり，$\partial(\cdot)$ は劣微分を表す．さらに，

- 問題 (35) の解が空でない．
- PnP-PDS のステップサイズ $\gamma_1,\gamma_2>0$ が $\gamma_1\gamma_2(\sigma_1(\mathbf{\Phi})^2+1)<1$ を満たす．

と仮定する．このとき，PnP-PDS によって生成される点列 $\{\mathbf{u}^{(n)},\mathbf{y}_1^{(n)},\mathbf{y}_2^{(n)}\}_{n\in\mathbb{N}}$ は問題 (35) の解のひとつに収束する．

上記の定理は，PnP-PG および PnP-PDS が収束するために必要な条件を示すだけでなく，収束先がどのような解であるかを数学的に表現している点が大きな特徴である．[2] では，これらの「零点包含問題」と「PnP アルゴリズムを導出する際に“想定”していた問題 ((7) および (19))」は本質的に等しいのだろうか？この疑問に関しては，凸解析の父である Rockafellar の研究 [46] に解答がある：$M_{\mathcal{D}_\gamma}$ が極大周期単調（maximally cyclically monotone）である場合に限り，$M_{\mathcal{D}_\gamma}$ が何らかの下半連続な真凸関数の劣微分になっている（＝ノイズ除去器の背後に $R_{\mathcal{D}_\gamma}$ が存在する）ことが保証できる（必要十分条件）．これは極大単調性に比べて非常に厳しい条件であるため，一般には満たされない．故に，3. 節で紹介してきた「PnP 正則化付きの最適化問題を介した解釈」は直感的で理解しやすかったものの，理論的には成立しないということがわかる．

最後に一点だけ補足したい．これまで議論は，（わかりやすさのため）挿入されるノイズ除去器が想定しているノイズの分散と PnP アルゴリズムのステップサイズが一致している仮定してきた（3. 節を参照）が，前述した収束定理において，実はこの仮定は必要ない．実際に「どのようなノイズの分散を想定してノイズ除去器を学習すべきか」については，発見的な方法が紹介されている [24] が，本質的には未解決である．

[3] 極大単調写像の定義および堅非拡大写像との関係は [45] 等を参照されたい．

[1] こちらも多少簡略化している．

[2] PnP-ADMM については未だ同様の定理が示されていない．これは，ADMM の収束を単調写像理論的に解釈することが近接勾配法や PDS に比べて難しいためであると考えられる．

5. おわりに

近年発展が著しい PnP アルゴリズムによる信号復元技術について，その導出から応用，収束理論などを網羅的に解説した．PnP アルゴリズムは「データドリブンに獲得した事前知識による高い復元性能」と「モデルベースアプローチが有する説明性・信頼性」を兼ね備えているため，まさに両者の“止揚（Aufheben）”とも呼べる新世代の信号復元技術になっている．本稿の締めくくりに，今後の研究の方向性をいくつか挙げる．

- 画像データ以外への適用：対象の信号情報に合わせてノイズ除去器を構成する必要があるが，その方法論が（画像データに比べて）現状乏しい．
- 信号復元以外への展開：PnP アルゴリズムの基本原理は信号復元に限定されないものであるため，最適化問題として記述できる様々なタスクに展開可能である（実例として，ハイパースペクトルデータの異常検知に応用されている [47]）．
- PnP アルゴリズム特化型ノイズ除去器の構成：厳密に非拡大なノイズ除去器を学習する方法は未だ確立されていない．また，PnP アルゴリズムの復元性能を最大化するようなノイズ除去器の条件や構成法も明らかになっていない．
- 汎用性の高い PnP アルゴリズムの開発：現状，PnP-PDS が最もそれに近いが，最適化技術の進歩と共に PnP アルゴリズムの発展も期待される．

本稿が，PnP アルゴリズムに興味を持ち，これから研究に取り組みたいと考えている読者にとって，理解の一助となれば幸いである．

参 考 文 献

[1] 小野峻佑. ハイパースペクトルイメージングと最適化—復元と合成—. 電子情報通信学会 基礎・境界ソサイエティ *Fundamentals Review*, 14(2):138–146, 2020.

[2] 工藤博幸. 低被曝 CT における画像再構成法—統計的画像再構成, 逐次近似画像再構成, 圧縮センシングの基礎—. *Medical Imaging Technology*, 32(4):239–248, 2014.

[3] Rizwan Ahmad, Charles A Bouman, Gregery T Buz-

zard, Stanley Chan, Sizhuo Liu, Edward T Reehorst, and Philip Schniter. Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery. *IEEE Signal Processing Magazine*, 37(1):105–116, 2020.

[4] Ulugbek S Kamilov, Charles A Bouman, Gregery T Buzzard, and Brendt Wohlberg. Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 40(1):85–97, 2023.

[5] 小野峻佑. 近接分離アルゴリズムとその応用: 信号処理・画像処理的観点から. オペレーションズ・リサーチ= *Communications of the Operations Research Society of Japan:* 経営の科学, 64(6):316–325, 2019.

[6] Kai Zhang, Yawei Li, Wangmeng Zuo, Lei Zhang, Luc Van Gool, and Radu Timofte. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):6360–6376, 2021.

[7] Yaniv Romano, Michael Elad, and Peyman Milanfar. The little engine that could: Regularization by denoising (RED). *SIAM Journal on Imaging Sciences*, 10(4):1804–1844, 2017.

[8] Patrick L Combettes and Valérie R Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005.

[9] Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 57(11):1413–1457, 2004.

[10] Jonathan Eckstein and Dimitri P Bertsekas. On the Douglas—Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55:293–318, 1992.

[11] Tom Goldstein and Stanley Osher. The split Bregman method for L1-regularized problems. *SIAM journal on Imaging Sciences*, 2(2):323–343, 2009.

[12] Hayden Schaeffer and Stanley Osher. A low patch-rank interpretation of texture. *SIAM Journal on Imaging Sciences*, 6(1):226–262, 2013.

[13] Shunsuke Ono, Takamichi Miyata, and Isao Yamada. Cartoon-texture image decomposition using blockwise low-rank texture characterization. *IEEE Transactions on Image Processing*, 23(3):1128–1142, 2014.

[14] Shunsuke Ono and Isao Yamada. Color-line regularization for color artifact removal. *IEEE Transactions on Computational Imaging*, 2(3):204–217, 2016.

[15] Shunsuke Ono. L_0 gradient projection. *IEEE Transactions on Image Processing*, 26(4):1554–1564, 2017.

[16] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.

[17] 小野峻佑. 近接分離による分散凸最適化—交互方向乗数法に基づくアプローチを中心として—. 計測と制御, 55(11):954–959, 2016.

[18] Singanallur V Venkatakrishnan, Charles A Bouman, and Brendt Wohlberg. Plug-and-play priors for model based reconstruction. In *Proceedings of IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 945–948. IEEE, 2013.

[19] Suhas Sreehari, S Venkat Venkatakrishnan, Brendt Wohlberg, Gregery T Buzzard, Lawrence F Drummy, Jeffrey P Simmons, and Charles A Bouman. Plug-and-play priors for bright field electron tomography and sparse interpolation. *IEEE Transactions on Computational Imaging*, 2(4):408–423, 2016.

[20] Stanley H Chan, Xiran Wang, and Omar A Elgendy. Plug-and-play ADMM for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98, 2016.

[21] Arie Rond, Raja Giryes, and Michael Elad. Poisson inverse problems by the plug-and-play scheme. *Journal of Visual Communication and Image Representation*, 41:96–108, 2016.

[22] Yu Sun, Zihui Wu, Xiaojian Xu, Brendt Wohlberg, and Ulugbek S Kamilov. Scalable plug-and-play ADMM with convergence guarantees. *IEEE Transactions on Computational Imaging*, 7:849–863, 2021.

[23] Ernest Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, and Wotao Yin. Plug-and-play methods provably converge with properly trained denoisers. In *Proceedings of International Conference on Machine Learning (ICML)*, pages 5546–5557. PMLR, 2019.

[24] Jean-Christophe Pesquet, Audrey Repetti, Matthieu Terris, and Yves Wiaux. Learning maximally monotone operators for image recovery. *SIAM Journal on Imaging Sciences*, 14(3):1206–1237, 2021.

[25] Samuel Hurault, Arthur Leclaire, and Nicolas Papadakis. Proximal denoiser for convergent plug-and-play optimization with nonconvex regularization. In *Proceedings of International Conference on Machine Learning (ICML)*, pages 9483–9505. PMLR, 2022.

[26] Manya V Afonso, José M Bioucas-Dias, and Mário AT Figueiredo. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Transactions on Image Processing*, 20(3):681–695, 2010.

[27] Shunsuke Ono and Isao Yamada. Signal recovery with certain involved convex data-fidelity constraints. *IEEE Transactions on Signal Processing*, 63(22):6149–6163, 2015.

[28] Laurent Condat. A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *Journal of Op-*

timization Theory and Applications, 158(2):460–479, 2013.

[29] Hiroki Ogawa, Shunsuke Ono, Yukihiro Nishikawa, Akihiko Fujiwara, Taizo Kabe, and Mikihito Takenaka. Improving grazing-incidence small-angle X-ray scattering–computed tomography images by total variation minimization. *Journal of Applied Crystallography*, 53(1):140–147, 2020.

[30] Hiroki Ogawa, Shunsuke Ono, Yuki Watanabe, Yukihiro Nishikawa, Shotaro Nishitsuji, Taizo Kabe, and Mikihito Takenaka. Artifact removal in the contour areas of SAXS-CT images by Tikhonov-L1 minimization. *Journal of Applied Crystallography*, 54(6):1784–1792, 2021.

[31] Kazuki Naganuma and Shunsuke Ono. A general destriping framework for remote sensing images using flatness constraint. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–16, 2022.

[32] Shingo Takemoto, Kazuki Naganuma, and Shunsuke Ono. Graph spatio-spectral total variation model for hyperspectral image denoising. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2022.

[33] Kazuki Naganuma and Shunsuke Ono. Towards robust hyperspectral unmixing: Mixed noise modeling and image-domain regularization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pages 8117–8138, 2024.

[34] Ryosuke Isono, Kazuki Naganuma, and Shunsuke Ono. Robust spatiotemporal fusion of satellite images: A constrained convex optimization approach. *IEEE Transactions on Geoscience and Remote Sensing*, 62:1–16, 2024.

[35] Thomas Pock and Antonin Chambolle. Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, pages 1762–1769. IEEE, 2011.

[36] Kazuki Naganuma and Shunsuke Ono. Variable-wise diagonal preconditioning for primal-dual splitting: Design and applications. *IEEE Transactions on Signal Processing*, 71:3281–3295, 2023.

[37] Shunsuke Ono. Primal-dual plug-and-play image restoration. *IEEE Signal Processing Letters*, 24(8):1108–1112, 2017.

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

[39] Yodai Suzuki, Ryosuke Isono, and Shunsuke Ono. A convergent primal-dual deep plug-and-play algorithm for constrained image restoration. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 9541–9545. IEEE, 2024.

[40] Yodai Suzuki, Ryosuke Isono, and Shunsuke Ono. Convergent primal-dual plug-and-play image restoration: A general algorithm and applications. in preparation for submission.

[41] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 248–255. IEEE, 2009.

[42] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.

[43] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Deep plug-and-play super-resolution for arbitrary blur kernels. In *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1671–1681, 2019.

[44] Renwei Dian, Shutao Li, and Xudong Kang. Regularizing hyperspectral and multispectral image fusion by CNN denoiser. *IEEE Transactions on Neural Networks and Learning Systems*, 32(3):1124–1135, 2020.

[45] Heinz H Bauschke and Patrick L Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces.* Springer, 2017.

[46] Ralph Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific Journal of Mathematics*, 17(3):497–510, 1966.

[47] Xiyou Fu, Sen Jia, Lina Zhuang, Meng Xu, Jun Zhou, and Qingquan Li. Hyperspectral anomaly detection via deep plug-and-play denoising CNN regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 59(11):9553–9568, 2021.